**Doctoral Dissertation**

# Towards Flexible Argumentation with Conversational Agents

**Niklas Rach**
Program of Information Science and Engineering
Graduate School of Science and Technology
Nara Institute of Science and Technology

Supervisor: Professor Keiichi Yasumoto
Ubiquitous Computing Systems Lab. (Division of Information Science)

Submitted on December 20, 2021

A Doctoral Dissertation
submitted to Graduate School of Science and Technology,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of Engineering

Niklas Rach

Thesis Committee:

Professor Keiichi Yasumoto (Supervisor)
Professor Satoshi Nakamura (Co-supervisor)
Associate Professor Hirohiko Suwa (Co-supervisor)
Assistant Professor Yuki Matsuda (Co-supervisor)
Professor Wolfgang Minker (Ulm University)

## Acknowledgements

The work in this thesis was only possible with the constant support and help of multiple people. First of all, I would like to express my sincere and deepest gratitude to my supervisor at Ulm University Prof. Wolfgang Minker. His advice and guidance throughout the whole research and writing process were crucial for its completion and frequently went beyond scientific matters. In particular, I am grateful for the opportunity of conducting this work in the scope of a joint PhD program with Japan, which made this thesis not just a scientific but also a personal journey.

A special thanks also goes to my co-supervisor at Ulm University Dr.-Ing. Stefan Ultes for his valuable feedback and advice as well as the constant critical questions that helped me getting the best out of each research step. Especially in the early stages of the work, his support and experience were an essential help in working in a field that was completely new to me at the time.

I extend my gratitude to my supervisor at NAIST Prof. Keiichi Yasumoto, first and foremost for his experienced guidance through an unfamiliar research area but also for his constant efforts and patience that made this joint PhD thesis possible. In addition, I am especially grateful to my co-supervisor at NAIST Prof. Yuki Matsuda for the fruitful and extensive discussions as well as his exhaustive assistance in managing both research and life in Japan. Over time, he has become as much a mentor as a friend. Moreover, both supervisors deserve a special thanks for their warm welcome at NAIST as well as the patient and gentle introduction into the fascinating Japanese culture.

An additional special thanks goes to my reviewer Prof. Philipp Cimiano for his interest in my work as well as the valuable discussions in the scope of the DFG Ratio meetings.

I would also like to thank my cooperation partners Prof. Elisabeth André for her expertise and advice as well as Klaus Weber for the continuous discussions about all the relevant details that made a successful cooperation possible. In addition, I thank Dr.-Ing. Johannes Daxenberger for a fruitful and fascinating cooperation that enabled the combination of two complex technologies.

Next, I want to express my thanks towards my colleges from the Dialogue Systems Group at Ulm University and the Ubiquitous Computing Systems Lab at NAIST for a supportive and productive atmosphere. Although it is not possible to mention everyone by name, I am sincerely grateful for the exceptional working environment at both institutes. Special thanks go to my co-authors Louisa Pragst, Matthias Kraus, Nicolas Wagner, Annalena Aicher, Waheed Abro and Isabel Feustel for countless discussions and mutual motivation. In addition, I thank my bachelor and master students Saskia Langhammer, Carolin Schindler, Yuchi Yang and Pavan Shigehalli.

On a personal note, a special thanks goes to my 'partner in crime' and office colleague at Ulm University Juliana Miehle, whose honest and direct approach to both research and matters of life in general were always refreshing, helpful and motivating. In addition, I thank my colleague at NAIST Hyuckjin Choi for his reliable support in practical matters as well as countless discussions about almost everything, which helped in regaining focus.

Finally, I would like to thank my parents Alexander and Gisela Rach as well as my brother Manuel Rach for their constant support and their pragmatic view that always brought me back on track. In addition, I thank Barbara and Manuel Straessle for always having my back. Most importantly, I thank my partner Tirzah Straessle for her infinite patience and loving support that was indispensable for the completion of this thesis.

# Abstract

Reasoning over conflicting information, i.e. to *argue* is at the core of many types of conversation such as persuasion, negotiation and deliberation. The multitude of setups that require this capacity makes it an interesting, yet challenging task for conversational agents. It is required in any scenario in which the corresponding system has to operate on conflicting information or positions, as for example contradicting online reviews or opposing points of view. However, current systems mostly focus on a specific application scenario and flexible approaches that can be applied in multiple systems are comparatively scarce. The present thesis addresses this issue in pursuit of the goal of *enabling flexible argumentation with conversational agents*. To account for the complexity of the topic, the work is divided into three sub-tasks: The development of challenging agent strategies, the acquisition of arguments and the adaptation to the interlocutor. All tasks are addressed one after the other with an emphasis on the *flexibility* of the proposed methods.

First, an agent-agent setup is introduced to assess the suitability of the utilized formal frameworks, evaluate the proposed methods and identify pending issues. Therein, argumentative dialogue is modelled as a dialogue game for argumentation and a general tree structure to represent arguments and their relations is used. The sub-task of challenging agent strategies is then addressed through multi-agent reinforcement learning based on a general reformulation of dialogue games as markov games. The approach provides the desired flexibility as it does not depend on a specific dialogue game instantiation, a corpus of conversational training data or pre-defined opponent strategies. In addition, a modification of the dialogue game for argumentation utilized in the experimental setup is introduced to enable a more complex and natural interaction. All steps (experimental setup, strategy optimization and dialogue game modification) are assessed separately in suitable evaluations.

For the flexible acquisition of arguments, the use of argument search engines in the context of conversational agents is investigated. To enable this combination of technologies, the arguments retrieved by different search engines are first assessed with respect to their suitability for conversational systems in a user study. Subsequently, a mapping of search results into argument structures that are required in many systems is proposed. The approach is highly flexible and enables a corresponding system to discuss any topic on which the search engine can find suitable arguments. An assessment in a crowd-sourcing study indicates the general feasibility of the method but also the expected room for improvement in comparison to an annotated structure. Complementing this highly flexible approach, a robust semi-supervised method to extract argument structures from reviews is also introduced to retrieve argument structures for evaluation studies.

The task of user-adaptive argumentation is then approached through the introduction of a fine-grained user model and a discussion of corresponding application scenarios. In addition, an implicit recognition of subjective user opinions based on non-verbal cues is proposed. All investigated methods are again evaluated separately to verify their functionality and their suitability for solving the task. Both, the preference model and the approach to recognize subjective argument quality aspects from social signals are quite flexible, as they can be directly applied in different system architectures. In a final step, the application and combination of the proposed methods in practical systems are discussed at the example of three different argumentative dialogue systems to demonstrate the applicability of the proposed methods.

# Contents

# List of Figures

# List of Tables

# 1 Introduction

The capacity of a computer system or program to converse with humans is a core aspect of artificial intelligence (AI). It is central to the idea of the imitation game (later referred to as Turing test) introduced by Turing (2009) to determine whether or not a machine can think in a way similar to humans. Consequently, the development of conversational systems (also referred to as dialogue systems or conversational agents) has gained a lot of interest in recent years and corresponding applications have increased in both popularity and complexity. Natural language interfaces are now part of everyday life in the form of command-based personal assistants like Apple Siri or Amazon Alexa. In addition, speech-based applications are available for multiple scenarios where hands-free interaction is required or beneficial, as for example during driving. At the same time, research has addressed more complex dialogical capacities like user-adaptation, as for example in (Janarthanam and Lemon, 2010; Miehle et al., 2018) and pro-activity, e.g. (Kraus et al., 2020). In particular, the study of systems that can reason over conflicting information or positions i.e. are able to *argue*, has shown a lot of progress in recent years. The relevance of argumentation for dialogue systems can be attributed to its essential role in human communication (Atkinson et al., 2017) and the resulting range of possible applications. Similar to humans, who argue in a variety of scenarios, conversational agents require this ability for multiple tasks such as debating (Slonim et al., 2021) or persuasion (Rosenfeld and Kraus, 2016). The present thesis contributes to this line of research and investigates the question of how to *flexibly* model the general capacity to argue in conversational agents. In particular, it addresses the task of *enabling conversational agents to argue about various topics in a way that is perceived as both natural and challenging by humans*. To this end, approaches to model core aspects of argumentation are proposed, namely challenging argumentation strategies, the ability to discuss various topics and an adaptation to the opinion of human users.

The development of these approaches inevitably crosses the boundaries between research disciplines, as argumentation is an active research area on its own with ties to multiple fields (Prakken, 2018). The study of argumentation in the context of AI is usually referred to as *artificial* or *computational argumentation* (Atkinson et al., 2017) and has also yielded a lot of progress recently. This includes (but is not limited to) formal models to represent and evaluate arguments in computer programs (Bentahar et al., 2010), models of argumentative dialogue (Prakken, 2006) as well as argument mining methods to automatically extract arguments from natural language sources (Lawrence and Reed, 2020). The herein proposed approaches build on this strong background and aim at bridging the gap between computational argumentation and dialogue systems technology.

## 1.1 Motivation

The capacity to argue enables a dialogue system to present conflicting aspects of a topic incrementally while taking the perspective of their interlocutor into account. Argumentation, then, is not so much a clearly defined task as the ability to operate on conflicting data. It hence enables applications in various domains, ranging from customer support (Galitsky, 2019) to completely argumentation oriented scenarios like persuasion (Rosenfeld and Kraus, 2016). The varying natures of the corresponding system tasks, as well as the respective requirements for the system, are illustrated by the following two examples. In the first one, the system (S) engages with the user (U) in a discussion about the controversial topic *animal testing*, thereby addressing the typical argumentative task of persuasion:

U: *Animal testing is generally a good thing.*

S: *I don't think so. Animal testing is mostly used for developing cosmetics. That is cruel and unnecessary.*

U: *Well, animal testing is also important for developing new medicine. In this case, it is essential to save human lives.*

S: *Although this is a better reason than testing cosmetics, there are several alternative procedures to test new medicine.*

In the second example, the user is looking for a suitable restaurant and the system goal is to provide a recommendation based on online reviews. Consequently, both sides cooperate to solve the task by exchanging information:

U: *I'm looking for a nice restaurant. What can you tell me about this new place on Baker Street?*

S: *Hm, several reviews are complaining about the high prices. I'm not sure if I would generally recommend it.*

U: *Well, since it is a special occasion today, I care more about the quality of the food than the pricing. Even if it is just for once.*

S: *In this case it might be an option for you. There are multiple positive reviews about the food. One of them says: They served the best steak I ever had.*

Although the topics and tasks differ between the two examples, the system relies in both cases on knowledge about suitable arguments and the ability to appropriately react to the utterances of the user. The range of these examples shows the high potential of argumentation in dialogue systems but also indicates that their development cannot be associated with a specific set of domains or scenarios. Consequently, the present thesis investigates application-independent (i.e. flexible) approaches to modelling argumentation in conversational agents.

In the discussion of the IBM project debater, this modelling is referred to as *composite AI* (Slonim et al., 2021), as it is composed of multiple sub-task that have to be handled by different modules. In contrast to *narrow AI*, which solves clearly specified individual tasks, end-to-end solutions are not feasible for this composition due to the varying nature of the sub-tasks that need to be addressed. Although the referenced work is concerned with the domain of debates instead of the herein investigated argumentative dialogues, the overall complexity of the task is similar in both cases. Therefore, the present thesis follows the approach of composite AI by dividing the addressed task into three sub-tasks, namely the development of *challenging agent strategies*, the *diversity of topics* and *adaptation to users*. All three are specified and motivated with pending challenges in existing works in the following.

The agent strategy is responsible for the utterance selection during the dialogue and hence essential for the task completion of every argumentative dialogue system. Consequently, a reasonable strategy depends on the individual task, as can be seen in the above examples. In addition, it is related to the utilized dialogue model, i.e. the representation of the interaction in the computer program, as this model defines the options the strategy can choose from. Due to these dependencies, current works on argumentative dialogue systems pursue mostly application-specific approaches to develop agent strategies. This includes a strict limitation of the system response to arguments that address the most recent user argument (Rakshit et al., 2019) and heuristics for the utilized dialogue model (Yuan et al., 2008). In addition, optimization techniques like Monte-Carlo planning (Rosenfeld and Kraus, 2016) or Q-learning (Alahmari et al., 2019) were investigated to optimize the system strategy in a particular scenario. However, they also depend on a corpus of training data or pre-defined opponent strategies to train against. The task of providing flexibility in the development of agent strategies hence requires approaches that abstract from one particular model and do not rely on additional resources like topic-specific training data or external knowledge. In addition, the study of flexible dialogue models facilitates challenging agent strategies in the respective domain. The approach pursued herein provides a formal framework for optimizing the system strategy in a class of models, namely dialogue games for argumentation. This choice is motivated by the argument that every argumentative dialogue between agents can be formalized as such a game (Atkinson et al., 2017). In addition, the training of the individual strategy is conducted in self-play, following the approach in (Barlier et al., 2015) and hence does not require any knowledge about the opponent strategy or additional training data.

The functionality of an argumentative dialogue system depends directly on knowledge about relevant arguments and their relations, as also illustrated in the above examples. A flexible model of argumentation hence has to include an approach to acquire this knowledge for various topics, i.e. to provide topic flexibility. The majority of current argumentative dialogue systems rely in their functionality on formal representations of arguments and their relations to each other. However, the acquisition of these *argument structures* is frequently time-consuming and includes the recording and annotation of human discussions (Rosenfeld and Kraus, 2016), crowd-sourcing experiments (Chalaguine and Hunter, 2019) or a generation by hand (Higashinaka et al., 2010). Although these methods ensure high-quality data, the corresponding system is strictly limited to the acquired resources. On the other hand, the field of argument mining has shown remarkable progress in the automatic extraction of arguments (and relations between them) from natural language sources (Lawrence and Reed, 2020). A flexible data-driven acquisition of arguments is hence not just desired but appears to be feasible. The approach pursued in the scope of this thesis builds on the current state of the art in the field of argument mining by combining argumentative dialogue systems with argument search engines. In doing so, it allows a system to discuss any topic on which the search engine can find suitable arguments.

The third herein addressed sub-task of user adaptation is concerned with adjusting the system strategy to user feedback. It is necessary for an argumentative system in order to respond to individual users accordingly, as again demonstrated in the above examples. However, explicit user feedback is not always feasible and can hinder the interaction. Corresponding approaches hence have to address two aspects, namely the (implicit) recognition of the user feedback during interaction and a corresponding adaptation of the system strategy. Flexibility in this regard can mainly be provided on the recognition side, as the adaptation approach depends inherently on the pursued goal of the system and hence on the specific application domain. User-adaptive argumentation is in comparison to the other two aspects less frequently considered in recent works and approaches so far focus on types of concerns (Chalaguine and Hunter, 2020) that can be recognized from the user utterance or assessed prior to the interaction (Hadoux and Hunter, 2017). In addition, the modelling of beliefs was considered (Hunter et al., 2019) and the use of emotional speech to improve a systems persuasiveness was discussed (Asai et al., 2020). The main contribution of this thesis to the task of user adaptation is an approach to automatically assess user opinion in terms of general argument quality aspects that can be utilized in different domains. In addition, a formal model of user preferences is discussed in the context of different application scenarios.

## 1.2 Contribution

The present thesis was conducted in the scope of a joint PhD program between Ulm University and Nara Institute of Science and Technology (NAIST), Japan. The corresponding work was hence distributed over the two involved research institutes, namely the Dialogue Systems Group at Ulm University and the Ubiquitous Computing Systems Lab at NAIST. In addition, it includes work done in cooperation with researchers from the Ubiquitous Knowledge Processing Lab at TU Darmstadt and the Chair for Human-Centered Artificial Intelligence at Augsburg University. The corresponding research institute at which the work was done and involved cooperation partners are explicitly named at the beginning of each chapter. In particular, approaches that do not contribute to this thesis and are included for the sake of consistency are highlighted as such. The contributions of this thesis to the three sub-tasks of *challenging agent strategy*, *diversity of topics* and *adaptation to users* are summarized in the following. This includes the general approach as well as its evaluation in an explicit scenario.

In the first step, preliminaries are defined that include the utilized dialogue system architecture, a specific dialogue model and the general structural representation of arguments. Based on these choices, a general agent-agent scenario is defined as a testbed for the approaches developed in later chapters. It is evaluated in a pilot study to identify specific pending issues and to confirm the suitability of the model choices.

For the case of challenging agent strategies, the main contribution is the reformulation of the dialogue games for argumentation introduced in (Prakken, 2005) as markov games (Littman, 1994). It allows for the use of reinforcement learning (RL) techniques to train the corresponding strategy in self-play and with respect to a formal winning criterion of the individual framework. The approach is validated in the previously defined evaluation setup and a proof-of-principle scenario for which the optimal strategy is known. Moreover, a modified winning criterion for the utilized dialogue game is proposed and applied in additional simulations. The results indicate the feasibility of the approach, although the utilized learning algorithms gradually reach their limits in scenarios with increasing complexity. As previously discussed, the development of strategies is directly related to the dialogue model. Since the preliminary evaluation indicates a shortcoming of the utilized dialogue game regarding the naturalness of the resulting dialogues, a modification of this dialogue game is introduced. It proposes extensions to the game formalism to increase the freedom of choices for the players and thereby enables more natural discussions. The desired effect on the perceived naturalness of dialogues generated with this framework is again confirmed by a user study in the previously defined evaluation setup.

The topic flexibility is approached through the use of argument search engines (Ajjour et al., 2019) that retrieve a list of arguments on a given search query from a multitude of web pages. As online sources are prone to include various types of errors, the first step in this direction is an assessment of the suitability of the retrieved arguments for their use in dialogue systems. To this end, an evaluation system is developed that presents the retrieved arguments (and their polarity towards the topic) to users via synthetic speech. In addition, a baseline argument search system is proposed and compared to two state-of-the-art search engines in an extensive user study. As the results indicate the general applicability of argument search engines in argumentative dialogue systems, the next step is then to map the retrieved arguments into the considered formal argument representation. This is approached through supervised learning models trained on a relation classification data set to estimate relations between the retrieved arguments. The resulting structures are then again assessed in the evaluation setup and through an extensive user study with respect to dialogue coherence. Although the comparison with the annotated argument structure suggests room for improvement, the general feasibility of the approach is indicated by the results. In addition, a semi-automatic approach to retrieve argument structures from reviews annotated with sentiment analysis labels is proposed. This additional approach addresses the requirements of evaluation setups for *non-opinion based topics*, i.e. topics for which a minimal bias of the user can be assumed.

As for adaptation to users, the contribution is again two-fold. The first one includes the definition of a formal preference model which enables a system to incrementally assess the opinion of an individual user. It is formally introduced and different application systems are discussed. The second and main contribution to user-adaptive argumentation is an approach to automatically recognize subjective argument quality aspects from non-verbal cues shown by users during interaction with a corresponding system. To this end, multimodal data collected during the evaluation of argument search engines is utilized to estimate the explicit user ratings from the respective non-verbal reactions. The investigated aspects are the users *interest* in an argument as well as its *convincingness*. A comparison with human annotators indicates that the task is complex but also that the automatic approaches perform similarly to humans in the considered setup.

Finally, several of the above-discussed approaches are integrated into three versions of a multimodal argumentative dialogue system. The first version enables a human user to assume the role of an agent in the experimental setup and to play the persuasive dialogue game against the virtual interlocutor through a drop-down interface. The system includes a virtual avatar presenting the system utterances via synthetic speech and a dialogue strategy optimized in the proposed markov game framework. The second version of the system considers agent-agent persuasion to display different persuasion strategies to a human audience. This version again utilizes the original dialogue game and a strategy optimized via reinforcement learning. In the final version, the system uses the modified dialogue game and includes, in addition to the optimized strategy, a hybrid strategy that allows the system to adapt the selection of utterances to the user feedback based on the emotional wording of the arguments. For the pre-training of the strategy, a data-efficient deep actor-critic algorithm is used in this version to address the previously identified scaling issues.

## 1.3 Outline

The remainder of the thesis is as follows: Chapter 2 summarizes the relevant background, including concepts from the field of dialogue systems, formal background on computational argumentation as well as utilized machine learning techniques. Subsequently, Chapter 3 provides an overview of related work on argumentative dialogue systems, argument quality assessment and affective computing. The architecture of the evaluation setup, including the utilized formal frameworks and the conducted pilot study are covered in Chapter 4. The approaches to challenging agent strategies are included in Chapter 5, starting with the reformulation of dialogue games as markov games and a numerical evaluation. In addition, the modifications to the utilized game framework and the corresponding evaluation are included. Chapter 6 then addresses the task of providing topic flexibility, starting with the assessment of argument search engines in the context of argumentative dialogue systems. Subsequently, the mapping of the retrieved arguments into a tree structure and the corresponding evaluation are discussed. The final part of the chapter includes the semi-automatic approach to extract argument structures from reviews. Approaches for user adaptation are then covered in Chapter 7, starting with the introduction of the proposed preference model in different scenarios. Its evaluation and application to adaption are reported subsequently. The final part of the chapter covers the automatic estimation of subjective argument quality aspects from non-verbal cues. Finally, the implementation of the application systems is included in Chapter 8, followed by a conclusion of the complete thesis and an outlook on future research directions in Chapter 9.

Although the thesis contributions are distributed over three sub-tasks, the individual approaches build on each other. The corresponding dependencies and the structuring of the contributing chapters is shown in Figure 1.1. The details on all dependencies and the sections in which they are used are discussed in detail in the respective chapters.

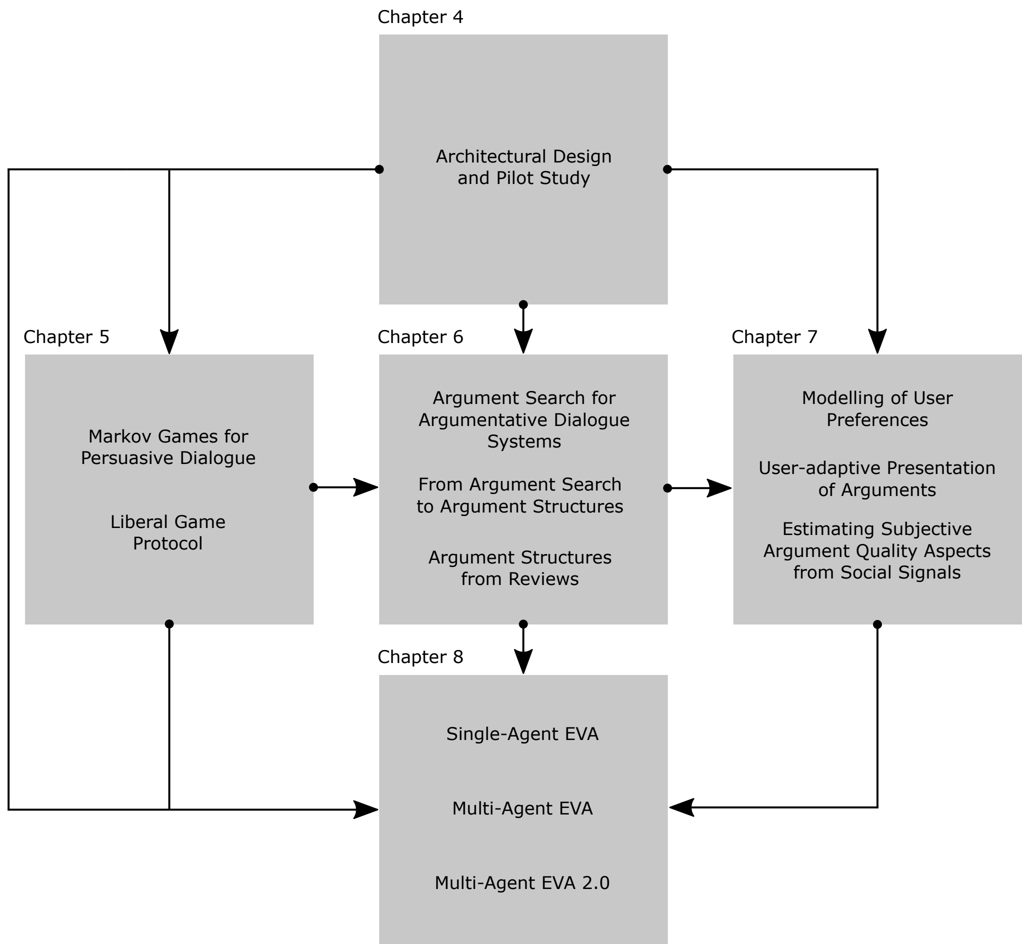


Figure 1.1: Structural overview of the contributing chapters. Arrows indicate dependencies between the corresponding approaches.

# 2 Background

This chapter summarizes relevant background for the present thesis, including theoretical frameworks and techniques. Due to the interdisciplinary nature of the topic, this summary will touch upon several different fields and the relevant context of the respective area is provided when necessary. The first part of the chapter addresses conversational agents, including their architecture and an overview of approaches to their evaluation. In the second part, models from the field of computational argumentation are introduced, starting with the formal representation of arguments and the formal modelling of argumentative dialogue. Subsequently, the theory of argument quality and the background on argument mining are summarized. In the last part of the chapter, machine learning techniques that are utilized within this thesis are introduced alongside evaluation metrics and agreement measures.

## 2.1 Conversational Agents

Computer systems that converse with humans are typically referred to as *dialogue systems* or, sometimes used interchangeably, as *conversational agents*. However, a strict definition of these terms is not represented in literature and they often refer to different types of systems (Deriu et al., 2020). In the scope of this thesis, both are used in a similar manner to describe a system that uses natural language to solve a dialogical task. In the following, established types of dialogue systems and corresponding architectures are summarized. The aim of this overview is to place argumentative systems in the context of current dialogue systems technology. In addition, the discussed conceptual and technological differences are required to motivate the architectural design in Chapter 4. Complementing this discussion, evaluation approaches with relevance for the present thesis are summarized subsequently.

### 2.1.1 Overview

Dialogue systems are frequently categorized based on their specific task or application domain. Following the discussion in (Deriu et al., 2020), three broad groups of dialogue systems can be identified:

- *Task-oriented systems* assist users in solving a specific task, for example finding a suitable restaurant, hotel or means of transportation (Ultes et al., 2017). Systems of this kind frequently serve as a natural language interface between the user and an application, for example a scheduler, and are often developed for a specific domain. In a more liberal sense, this category also includes systems that pursue a specific goal throughout the interaction. This goal is not necessarily in line with the goal of the user, as for example in the case of negotiation (Georgila and Traum, 2011).

- *Chat-based systems* engage with human users in a natural language conversation without a specific task to solve. Systems of this kind are designed to simulate (aspects of) human conversations and are hence not inherently restricted to a specific domain.
- *Question-answering (QA) systems* are built to, as the name suggests, provide answers to user queries. In their original form, they can be seen as a natural language interface between users and a database, but more recent work has investigated aspects like handling follow-up questions and multi-domain QA (Choi et al., 2018).

Systems concerned with argumentation are frequently associated with *task-oriented systems* due to the specific task they are built to solve, for example winning a debate (Slonim et al., 2021) or persuade the user of a certain position (Rosenfeld and Kraus, 2016). However, chat-based argumentative systems were also introduced (Le et al., 2018; Rakshit et al., 2019) and are denoted as *argumentative chatbots* in the remainder of this work. In addition, dialogue systems can also be built to handle additional modalities like gestures or emotions. This includes their recognition, interpretation and simulation in the system output. Systems with this capacity are referred to as *multimodal dialogue systems*.

From a technical perspective, two main approaches to develop dialogue systems have emerged. In the 'traditional' approach, the system is comprised of multiple modules that handle the different sub-tasks. Systems built in this manner are referred to as *pipeline systems*. In contrast, more recent approaches aim at learning the response to a given user utterance from data, thereby combining the different sub-tasks in the system to a single machine learning problem. These systems are consequently called *end-to-end systems*. The relevant background on both approaches is discussed in detail in the following.

### 2.1.2 Architectures

A pipeline system capable of handling spoken natural language is comprised of five different modules, as depicted in Figure 2.1. The automatic speech recognition (ASR) handles the speech input by the user and transfers it into a textual representation by identifying the words most likely spoken in the speech signal. Technically, this is realized through the use of statistical approaches like hidden markov models or, more recently, artificial neural networks.

The natural language understanding (NLU) module is then concerned with mapping the retrieved textual representation of the user utterance into a more abstract representation that can be handled by the dialogue manager (DM). In a typical task-oriented system, this representation is divided into a general *intent* like requesting information and an *entity* (or value) that specifies the intent, like requesting information about the weather in Berlin. Approaches of this kind are called slot-filling approaches but others are possible, depending on the specific system and the corresponding domain. Technically, recent NLU modules build on the neural network-based approaches developed for a variety of natural language processing tasks. A promising and popular approach in this regard is the class of transformer models (Bunk et al., 2020). Also, a variety of out-of-the-box solutions to natural language understanding were provided. A detailed discussion of these tools, including a comparison of their performance in the context of question answering systems, can be found in (Braun et al., 2017).

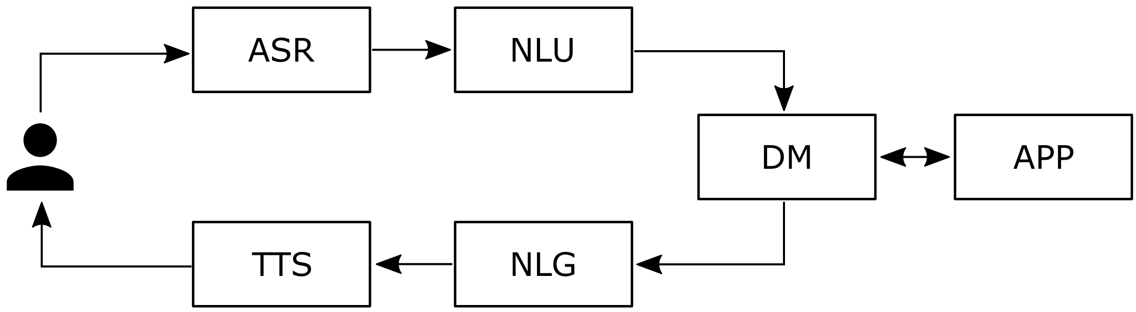


Figure 2.1: Pipeline architecture of a spoken dialogue system, including the five modules automatic speech recognition (ASR), natural language understanding (NLU), dialogue manager (DM), natural language generation (NLG) and text-to-speech (TTS), as well as a possible application (APP).

Given the (abstract) representation of the NLU, the task of the DM is then to update the dialogue state accordingly and select the next system utterance. In the case of *task-oriented systems* serving as an interface, the DM module is also responsible for the communication with the respective application or database. The selection of the next utterance was originally approached through handcrafting topic-specific rules. However, due to their inflexibility and their incapability of compensating errors from the previously discussed modules, statistical approaches to dialogue state tracking and utterance selection were introduced (Young, 2007). To this end, the decision making is formalized as a partially observable markov decision process (POMDP) in which the current state of the interaction is modelled as a distribution over possible dialogue states (so-called belief state). The system policy that selects the next utterance based on the current belief state is then optimized through reinforcement learning (Young et al., 2013). Also along the line of reinforcement learning-based approaches, Barlier et al. (2015) proposed a formulation of the dialogue management as markov game and therefore training the system in self-play.

The system utterance selected by the dialogue manager is again an abstract representation that has to be transformed into a natural language utterance before presentation to the user. This is handled by the natural language generation (NLG) which maps the output of the dialogue manager to text. The basic approach to NLG is based on pre-defined templates with fixed formulations. However, it has been argued that dialogue systems should be capable of adapting their formulations in order to be more natural (Miehle et al., 2018) and more flexible neural network-based approaches to NLG were proposed (Wen et al., 2015).

Finally, the text-to-speech (TTS) module or speech generation transforms the output of the NLG into synthetic speech, inversely to the ASR. The task is hence to generate an acoustic signal. Technically, this step is again approached by means of statistical models like hidden markov models or artificial neural networks.

The pipeline architecture is frequently associated with *task-oriented systems* as the ‘traditional’ form of conversational agents serving as an interface between the user and an application (Deriu et al., 2020). Arguments recently made against this approach are the required time-consuming handcrafting and/or the large amount of training data that is necessary to train a model for each sub-task separately (Wen et al., 2017). Alternatively, *end-to-end systems* address the development of a dialogue system through training a single model that combines the individual sub-tasks. This can be approached in a supervised manner, for example through combining multiple neural networks (Wen et al., 2017) or by means of reinforcement learning (Li et al., 2017). Also along the line of end-to-end trainable systems, approaches considered primarily in the development of *chat-based systems* utilize a trained encoder-decoder neural network architecture to generate the next system utterance from the current user input (Vinyals and Le, 2015). Consequently, they are referred to as *generative approaches* in the following. In addition, *retrieval-based* approaches that predict the next system utterance in a corpus from the user input for example utilizing similarity measures (Rakshit et al., 2019) were proposed.

A general advantage of end-to-end systems is that no domain-specific modelling is used and that it hence, given enough training data, can learn any kind of conversation. On the downside, the corresponding training also frequently requires large amounts of suitable training data (Deriu et al., 2020). In the case of argumentative systems, end-to-end systems are mostly unexplored due to the complex nature of argumentative tasks. Exceptions in this regard are the two argumentative chatbots mentioned earlier (Le et al., 2018; Rakshit et al., 2019).

### 2.1.3 Evaluation

With argumentation being a relatively new field for dialogue systems, the question of evaluation naturally arises. However, the evaluation of dialogue systems in general is still an open issue that is frequently disputed over (Deriu et al., 2020) and for which several approaches were introduced. In the following, a selection of those that bear relevance for the present thesis is introduced and discussed. For task-oriented systems, a straightforward metric of their success is the frequency with which they are able to complete their assigned task, i.e. their *task-success rate* (Schatzmann et al., 2005). Conversely, a chat-oriented system can be evaluated with respect to the overall length of a dialogue, i.e. how long it engaged in a conversation. However, these measures do not include information on the human perception of the dialogue. Especially in the task-oriented case, users are likely to continue the interaction up to a point where they reach the desired outcome, for example get the information they are looking for, even if they are not satisfied with the functionality of the system.

To address this issue, an approach to evaluate dialogue systems based on *user satisfaction* was considered and included in an evaluation framework (Walker et al., 1997). As intended, this measure of success is highly subjective and the evaluation depends on explicit feedback through the individual user. However, for an assessment during the interaction, for example with the goal to adapt the system behaviour accordingly, explicit feedback is impractical. Consequently, the automatic assessment of user satisfaction was explored for example using n-gram models in (Hara et al., 2010) and hidden markov models in (Engelbrecht et al., 2009). Although the proposed models were able to beat the random baseline, a problem indicated by (Higashinaka et al., 2010) is the inherently subjective nature of the ratings. To bridge the gap between objective and subjective evaluation metrics, an approach based on expert ratings, the so-called *interaction quality* was introduced (Ultes et al., 2013). The underlying idea is that expert annotations of transcribed dialogues indicate the overall quality of an interaction turn and that these ratings can be predicted with higher accuracy by machine learning methods than an individual subjective rating. The correlation between user satisfaction and interaction quality was investigated in (Schmitt and Ultes, 2015) and several approaches to estimate the interaction quality automatically were proposed (Rach et al., 2017a; Ultes et al., 2014).

An additional evaluation metric has evolved in the context of end-to-end systems, where the coherence of the system responses is assessed. This is necessary because the retrieval of responses from data can lead to utterances that are inconsistent with earlier ones. In (Venkatesh et al., 2017), a measure of coherence is approached through annotating dialogue utterances as *incorrect*, *inappropriate* and *irrelevant*. The overall coherence of a system is then determined based on the ratio of coherent and overall utterances.

As for argumentative systems, no unified evaluation approach has emerged yet. However, it can be argued that each of the above-discussed approaches bears relevance in this case. The task success of an argumentative system, for example the frequency with which a persuasive system is able to convince a human interlocutor, is arguably an indicator for its functionality. However, the success of such a system is, in contrast to an information system, subjective and hence needs to take into account the user perspective. The approaches based on user satisfaction on the other hand include this subjectivity but the satisfaction of the user in a competitive setup is not necessarily an indicator for success. The coherence finally is also important as a system should be consistent in its argumentation. However, the above-stated criteria can be hard to assess in the case of arguments as for example the relevance of an argument is again subjective.

## 2.2 Computational Argumentation

This section addresses the background of computational argumentation with an emphasis on formal frameworks to model arguments and the (dialogical) process of argumentation. In addition, a theoretical taxonomy for argument quality is discussed and an overview of argument mining is provided. As several of the herein proposed approaches conceptually build upon existing formal models, the corresponding background is discussed in the necessary depth and hence the main focus of this section.

Generally, formal argumentation is concerned with reasoning in situations of incomplete or inconsistent information (Prakken and Vreeswijk, 2001) and can hence be applied to resolve conflicts arising for example from differing beliefs, knowledge or goals. The interdisciplinary nature of the field and the differing requirements for the envisioned applications lead to a variety of formal models that can (in part) be transformed into and/or derived from another. Consequently, divisions and taxonomies were proposed to summarize these models based on their properties and addressed scenarios to provide researchers with a consistent overview that allows for an a priori selection of models that suit their task.

A common division is the one into *structured* and *abstract* argumentation frameworks. Abstract frameworks thereby denote models that make no assumptions regarding the internal structure of arguments, focusing only on their relation. In contrast, structured frameworks also investigate the logical composition of arguments, for example in the form of elements from a logical language connected by inference rules. This disposition is also reflected in the taxonomy proposed by Bentahar et al. (2010), who refer to frameworks concerned with the composition of arguments as *monological models*, that address the micro-structures of arguments and implement them as tentative proofs. Abstract frameworks that are concerned with the macro-structure of arguments, i.e. their relation towards each other are referred to as *dialogical models*. In addition, the authors add the class of *rhetorical models* that focus on a persuasive arrangement of arguments and take into account the perception of an audience. The separation utilized by Prakken (2018) in his historical overview of formal argumentation distinguishes *inference-based* argumentation and *dialogue-based* argumentation, thereby focusing not on the individual properties of the models but on their addressed task: Whereas *inference-based* argumentation is concerned with the investigation of conclusions that can be (legally) drawn from a set of arguments, *dialogue-based* argumentation addresses the formal description of argumentation as a dialogical interaction between interlocutors.

Throughout this work, the separation into *inference-based* argumentation and *dialogue-based* argumentation is adapted, not because it is generally preferred over the others but because it reflects the requirement of the present work best: The envisioned capacity of conversational agents to argue flexible over multiple topics and in an intelligent and challenging manner requires a model to represent knowledge about existing arguments and their relations as well as an approach to structure the interaction. Whereas the representation of arguments can be addressed with models from the class of *inference-based* argumentation, the structuring of the dialogue is an instance of the *dialogue-based* argumentation. However, as both models are to be combined in a single system architecture, they have to be compatible with each other (in addition to the requirements imposed by the task of ensuring flexibility).

The following subsections address the formal background on models of both classes, starting with *inference-based* argumentation. The different levels of abstraction in the reviewed approaches are highlighted to ensure an overview that is also consistent with the other above discussed divisions. *Dialogue-based* argumentation is discussed subsequently with a focus on dialogue games for argumentation as an approach to formally model argumentative interaction. In addition, the taxonomy of argument quality and the essential concepts of argument mining are introduced in separate subsections.

### 2.2.1 Inference-based Argumentation

The discussion of *inference-based* argumentation covers works that are relevant for the approaches investigated in the scope of this thesis as well as influential models that are required to place them in the context of existing work. As the main task of the models for the considered application in conversational agents is the representation of knowledge regarding available arguments, the discussion is mainly focused on how arguments are represented and/or constructed in the respective model. Nevertheless, approaches to infer conclusions within these models are included, whenever they are relevant. As for notation, the present discussion deviates from the ones in the original works to ensure consistency within the remainder of the thesis. An argument of any form is denoted as $\Phi \in Args$ with $Args$ the set of all arguments in the investigated scenario. In addition, argument components used for the construction of arguments are denoted with $\varphi$.

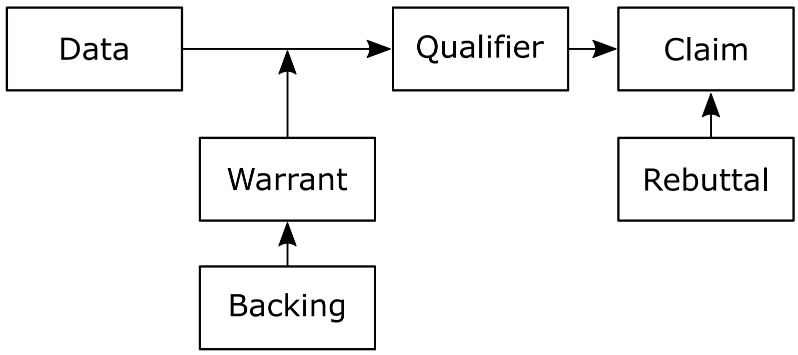


Figure 2.2: Pictorial representation of an argument in Toulmin's model.

#### Toulmin Model

One of the arguably most influential works on the structure of arguments, i.e. the question of how an argument is composed was presented by Toulmin (2003). In contrast to other works like (Pollock, 1987) the model does not rely on formal logic to define components and inference rules. In its original form, the model composes an argument out of six different components:

- *Claim:* The general proposition or statement that is to be proven by the argument.
- *Data:* Information that justifies the claim.
- *Warrant:* Generalization of the data that legitimates the conclusion of the claim.
- *Qualifier:* The argumentative strength of the claim as a result of the data, determined by the warrant.
- *Backing:* Additional grounds for the warrant to support it against a challenge.
- *Rebuttal:* Challenge of the claim based on exceptional circumstances that either rebut the general validity of the warrant in the present case or the inference of the claim from it.

An argument constructed with these components can then be graphically depicted as shown in Figure 2.2. To demonstrate the functionality of the model, an example argument inspired by the discussion in (Verheij, 2009) is investigated. It is concerned with the *claim* that *"Lara's first language is English"*. This claim can be supported by the *data "Lara was born in the US"*. The corresponding *warrant* is *"English is the official language in the US"*, which can be supported by statutes and regulations that serve as *backing*. Based on the data and the warrant, it follows that *"Lara's first language is, presumably (qualifier) English, unless her parents immigrated to the US and raised her with a different first language" (rebuttal)*.

Toulmin's model was extended in multiple ways to meet the requirements of specific application domains and to investigate additional aspects like the chaining of multiple arguments to a dispute or the role of rebuttals in a dialogical context. An example that focuses on the formalization of (legal) dialogues and is hence of particular interest in the context of the present work was introduced in (Bench-Capon, 1998). The therein adapted model omits the qualifier components and allows the claim, data and rebuttal as a link to other arguments. For claim and data, the author explicitly specifies this link by allowing both components to serve as data for the new arguments. In addition, a new component called *presupposition* was introduced representing topic-specific knowledge that is generally accepted (implicitly) but can now be made explicit if desired or necessary.

In addition, the use of Toulmin's model to structure natural language text with the perspective to automatically recognize argument structures (i.e. for argument mining) was investigated in (Habernal et al., 2014). Again, a set of modifications is applied to meet the requirements of the desired application scenario: As in (Bench-Capon, 1998), the qualifier is omitted, based on the observation that it is often not explicitly stated in natural language argumentation. In addition, the warrant is omitted, again with the justification that the reason for concluding the claim from the data is usually not explicitly stated. A further modification is the extension of the backing to support the argument as a whole and not just the warrant as in the original model. Finally, as in other modifications, an attack on the rebuttal in the form of a *refutation* is included. It is pointed out that the strength of the model is its fine-grained structure which allows distinguishing between different roles of the components. On the downside, the authors report that this complexity and the informal definition of the components make it difficult to apply the scheme to natural language argumentation (especially discussions on the web), since the corresponding arguments do not necessarily comply with this structure and often leave several components implicit.

### Claim-Premise Approach

In contrast to the previously discussed model, the term claim-premise approach does not refer to a strictly defined (single) framework but is herein used as a summarizing description of models and approaches sharing a general property. The underlying assumption is that arguments consist of a (possibly empty) set of premises from which a conclusion can be inferred, as for example in (Besnard and Hunter, 2008; Habernal et al., 2014). In contrast to the Toulmin model, claim-premise approaches therefore focus on two component types only from which arguments are constructed. In this general form, the ways by which the conclusion is derived from the premises is not specified and the term is hence applicable to multiple approaches from different fields. An area where these models are of particular interest is the identification of argument patterns or structures in natural language argumentation, i.e in the field of argument mining. As mentioned before and discussed in (Habernal et al., 2014), detailed and fine-grained structures as proposed by Toulmin are not frequently reflected in the corresponding sources and claim-premise models are hence more suitable for this task. As the goal of this thesis includes a combination of argument mining technology with dialogue systems, a claim-premise approach is chosen to represent arguments and their relations.

The specific model utilized herein is discussed in detail in the following. It was introduced by Stab and Gurevych (2014a) and builds on the distinction between claims and premises to represent the argument structure in persuasive essays. To this end, two directed relations (*support* and *attack*) between argument components are considered. However, a formal separation of the

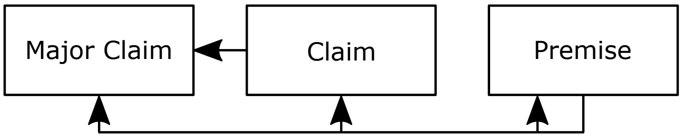


Figure 2.3: Illustration of allowed relations (arrows) between the different component types.

attack relations into undercutting and rebutting as for example in (Besnard and Hunter, 2008) is not included. As for notation, if a component $\varphi_1$ *supports* or *attacks* another component $\varphi_2$, $\varphi_2$ is called the target of $\varphi_1$ (or that $\varphi_1$ targets $\varphi_2$, respectively). The overall annotation scheme is then defined to meet the general structure of persuasive essays by distinguishing the three component types *Major Claim*, *Claim* and *Premise*. The general topic including the author's position (or stance) is encoded in the *Major Claim* component which is the only argument component without a relation towards another one. An example for this component is *"I believe that every student should make the experience of living abroad"*. A *Claim* is then a general statement that refers to the *Major Claim* but needs additional backing or justification. Consequently, each *Claim* is only allowed to target the *Major Claim* with a relation, not other *Claims* or *Premises*. Finally, a *Premise* provides information, evidence or general justification for either the *Major Claim*, a *Claim* or another *Premise*. Therefore, *Premises* can target all other component types with a relation. The general concept of component types and allowed relations is illustrated in Figure 2.3.

An argument in this annotation scheme includes an arbitrary number of *Premises* and either a *Claim* or the *Major Claim*. However, not all premises of an argument have to target the *Claim* directly, as can be seen in the following example from the original work comprised of a *Claim* and three *Premises*:

*"Living and studying overseas is an irreplaceable experience when it comes to standing on your own feet. One who is living overseas will of course struggle with loneliness, living away from families and friends but those difficulties will turn into valuable experiences in the following steps of life. Moreover, the one will learn living without depending on anyone else."* (Stab and Gurevych, 2014a, p. 1504)

The first sentence is a general statement supporting the previously discussed *Major Claim* that every student should make the experience overseas ($Claim_1$). The next sentence includes two premises, one that *attacks* the *Claim* by stating that being apart from families and loved ones leads to loneliness ($Premise_1$) and one that *attacks* its predecessor by stating that this loneliness will turn into a valuable experience later ($Premise_2$). Finally, the third sentence includes an additional premise that directly *supports* the *Claim* by saying that living abroad allows learning to live independent ($Premise_3$).

It can be seen that the annotation scheme is hence quite flexible: Although specific component types are defined to meet the requirements of the investigated persuasive essays, the formal difference between the components is solely determined by their allowed relations.

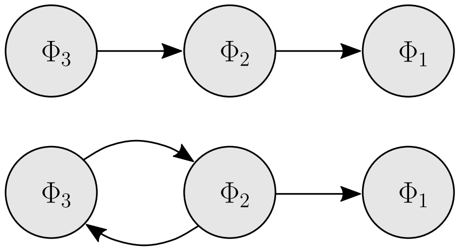


Figure 2.4: Illustration of two argumentation frameworks $AF_1$ and $AF_2$ with arguments (circles) and attack relations (arrows).

### Abstract Argumentation Frameworks

So far, the discussion was focused on approaches that can be described as *structured* argumentation frameworks, i.e. frameworks that model the internal structure of the described arguments. In contrast, the following approaches abstract from this internal structure and make no assumption regarding the specific composition of an argument. Instead, arguments are assumed to be given as atomic units and the topic of interest is their relation and the conclusions that can be drawn from a set of (related) arguments. A seminal approach in this direction was proposed by Dung with the motivation to *"study the fundamental mechanism, humans use in argumentation, and to explore ways to implement this mechanism on computers"* (Dung, 1995, p. 321). To this end, an (abstract) *argumentation framework* (AF) is defined as follows:

**Definition 1** (Abstract Argumentation Framework)**.** *Let Args be a (finite) set of arguments, then an abstract argumentation framework is defined as tuple $(Args, attack)$, with $attack \subseteq Args \times Args$ a set of attack relations between arguments in Args.*

As can be seen from the definition, the framework relies solely on attack relations to model the interplay of arguments. For two arguments $\Phi_1, \Phi_2$, the attack relation $(\Phi_1, \Phi_2) \in attack$ denotes that $\Phi_1$ attacks $\Phi_2$. Consequently, circular attacks ($\Phi_1$ attacks $\Phi_2$ and $\Phi_2$ attacks $\Phi_1$) as well as self-attacks $(\Phi_1, \Phi_1)$ are included in the framework. Another consequence of this abstraction is that an argument framework can be represented as directed graph with arguments as nodes and attack relations as edges. This is illustrated in Figure 2.4, showing two example frameworks $AF_1$ and $AF_2$. For both frameworks, the set of arguments is given as $Args_{1,2} = \{\Phi_1, \Phi_2, \Phi_3\}$. For $AF_1$, the set of attack relations is $attack_1 = \{(\Phi_3, \Phi_2), (\Phi_2, \Phi_1)\}$ whereas for $AF_2$ it also includes a circular attack and hence is given as $attack_2 = \{(\Phi_3, \Phi_2), (\Phi_2, \Phi_1), (\Phi_2, \Phi_3)\}$.

Given an argumentation framework that encodes arguments and conflicts between them in the form of attack relations, the question can be raised which arguments remain valid in the encoded conflict and which conclusions can safely be drawn from them. This can be approached through the definition of *semantics* that formally determine the outcomes of a conflict encoded in an AF. The following discussion is limited on the general concepts as they are sufficient the herein addressed approaches. For a detailed overview, the interested reader is referred to the extensive introduction into argumentation semantics by Baroni et al. (2011) that serves as a foundation for the present discussion. In particular, the examples below are adapted from the referenced work.

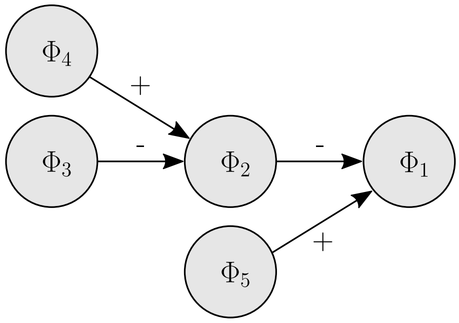


Figure 2.5: A bipolar argumentation framework, including five arguments with support (+) and attack (-) relations.

For the definition of semantics, two different approaches can be found in the literature: The first one also used in the original work is based on *extensions*, which are sets of arguments that remain valid throughout the conflict of interest encoded in the AF. An example extension for the above shown example framework $AF_1$ could be $\{\Phi_3, \Phi_1\}$ since the attack $(\Phi_3, \Phi_2)$ arguably invalidates $\Phi_2$ and, as a consequence also the attack $(\Phi_2, \Phi_1)$. The second approach to semantics is referred to as *labelling-based*. In this approach, arguments are assigned a label that determines their status in the conflict. The set of labels used in the discussion of Baroni et al. (2011) consists of the three labels $\{in, out, undec\}$, where *in* indicates that an argument is valid, *out* the opposite, i.e. that the argument is refuted and *undec* that the status is undecided. For the earlier example of $AF_1$, the status *in* can for example be assigned to $\Phi_3$ (since it is not attacked). Consequently, $\Phi_2$ is refuted and hence *out* which in turn reinstates $\Phi_1$, resulting in the corresponding status *in*. For the herein considered labels, it can be shown that extension and label based approaches can be translated into another (Baroni et al., 2011).

Several adaptations of the original definition of Dung's argumentation framework were proposed, again in the context of specific application domains or scenarios. An extension that is of interest in the context of the present work is the introduction of an additional *support* relation between arguments (Cayrol and Lagasquie-Schiex, 2005; Cayrol and Lagasquie-Schiex, 2009), leading to the definition of a *bipolar argumentation framework* (BAF):

**Definition 2** (Bipolar Argumentation Framework)**.** *Let Args be a set of arguments, then a bipolar argumentation framework is defined as tuple* $(Args, attack, support)$ *with* $attack \subseteq Args \times Args$ *the set of attack relations and* $support \subseteq Args \times Args$ *the set of support relations for which* $attack \cap support = \emptyset$ *holds to ensure consistency.*

It follows directly from the definition that an argument cannot support and attack another argument at the same time. Again, the resulting framework can be represented as a directed graph with arguments as nodes and two kinds of edges that denote support and attack relations between them, as shown in Figure 2.5. Based on this definition, it is again possible to define extensions and semantics similar to the original framework as for example discussed in (Cayrol and Lagasquie-Schiex, 2005). It should be noted that although there is a conceptual similarity between bipolar argumentation frameworks and the previously discussed claim-premise approaches, the notion of arguments in BAFs is (in contrast to the claim-premise approach) still abstract, i.e. no assumption is made regarding the composition of arguments. Nevertheless, the similarity between the two approaches will be utilized in the scope of this thesis to expand concepts developed for one framework into another.

A further extension of BAFs was the introduction of weights that indicate the strength of an argument (Amgoud and Ben-Naim, 2018), leading to the definition of a *bipolar weighted argumentation graph* (BWAG):

**Definition 3** (Bipolar Weighted Argumentation Graph)**.** *Let Args be a set of arguments, then a bipolar weighted argumentation graph is defined as tuple* $(Args, attack, support, w)$*, with* $attack \subseteq Args \times Args$ *attack relations between arguments,* $support \subseteq Args \times Args$ *support relations between arguments and* $w : Args \rightarrow [0,1]$ *a function that assigns each argument a weight.*

The weight of an argument $w(\Phi)$ determines its (intrinsic) validity with a value close to one indicating high validity and a value close to zero the opposite. Consequently, semantics are now defined to model the effect of related arguments on the current one based on their weights:

**Definition 4** (Semantics)**.** *For a bipolar weighted argument graph* $(Args, attack, support, w)$*, a semantics* $\mathbb{S}$ *defines a function* $k : Args \rightarrow [0,1]$ *that assigns each argument an overall strength.*

Based on this definition, a variety of principles for semantics were investigated alongside the transformation of the original (extension-based) semantics from Dung's framework into the new one (Amgoud and Ben-Naim, 2018). In addition, the authors introduce the *Euler-based restricted semantics*, which takes into account supporting and attacking relations in an exponential manner. To this end, a *restricted semantics* is defined analogous to Definition 4 for acyclic non-maximal graphs. The corresponding semantics can then be defined as:

**Definition 5** (Euler-based Restricted Semantics)**.** *For an acyclig non-maximal bipolar weighted argument graph* $(Args, attack, support, w)$*, let* $att(\Phi_i) \subseteq Args$ *and* $supp(\Phi_i) \subseteq Args$ *denote the subset of arguments that attack and support* $\Phi_i$*. The Euler-based restricted semantics then recursively defines the function* $k : Args \rightarrow [0,1]$ $\forall \Phi_i \in Args$ *as*

$$k(\Phi_i) = 1 - \frac{1 - w(\Phi_i)^2}{1 + w(\Phi_i)e^{E_i}}, \tag{2.1}$$

*where* $E_i$ *denotes the energy of* $\Phi_i$ *given as*

$$E_i = \sum_{\Phi_j \in supp(\Phi_i)} k(\Phi_j) - \sum_{\Phi_j \in att(\Phi_i)} k(\Phi_j). \tag{2.2}$$

It should be noted that the present definitions slightly deviates from the referenced work. However, the fundamental concepts are the same in both notations and definitions. An example BWAG is depicted in Figure 2.6. According to the above definition, the leaves of the graph, i.e. arguments that are not addressed by other arguments with a relation have an overall strength that equals their weight. For the second argument with internal strength 0.25, the attacking and supporting relation equal each other, leading to an energy of 0 and hence also to an overall strength of 0.25. For the

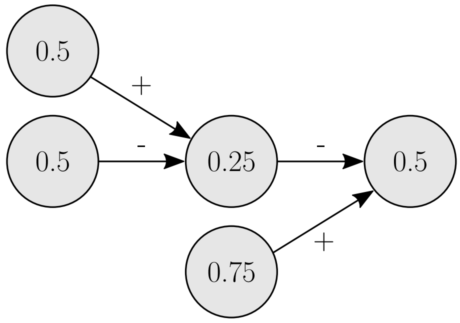


Figure 2.6: A bipolar weighted argument graph with internal strength of arguments denoted in the circles.

root argument, the energy is then 0.5, resulting in an overall strength of 0.59. Consequently, the overall strength of the root argument is higher than its weight due to the influence of the strong support.

The last herein discussed adaptation of abstract argumentation frameworks was introduced in (Rosenfeld and Kraus, 2016) for use in a persuasive agent. The proposed *weighted bipolar argumentation framework* includes, similar to the above-discussed extensions, bipolar relations but also weights assigned to relations (instead of arguments), as well as beliefs that represent an agents confidence in an argument.

**Definition 6** (Bipolar Weighted Argumentation Framework)**.** *Let $v$ be an ordered set with a minimum and a maximum. Then a weighted bipolar argumentation framework is a tuple $(Args, attack, support, w, B, \Phi_0)$ with Args the set of arguments, $support \subseteq Args \times Args$ and $attack \subseteq Args \times Args$ the support and attack relations between them, $w : attack \cup support \to v$ a weight function, $B : Args \to v$ a belief function and $\Phi_0$ a specified argument that indicates the disputed topic.*

It can be seen that, although different in the interpretation, the belief function is formally analogous to the internal weight in BWAGs if $v = [0,1]$. However, it is stressed by the authors that the belief and the weight function are associated with a specific agent and different agents in their considered scenario can hence hold different beliefs and weights. This reflects the dialogical perspective of the desired application scenario, where agents exchange arguments and try to persuade each other, in contrast to the inference-based perspective where the focus is on a (monological) determination of reasonable consequences from the conflicting positions encoded in the arguments. Therefore, the bipolar weighted argumentation framework could also be seen, despite its origins in inference-based argumentation, as an approach to dialogue-based argumentation discussed in the following section.

### 2.2.2 Dialogue-based Argumentation

Approaches that are associated with dialogue-based argumentation focus on the dialogical aspect of argumentation, i.e. the exchange of (and dispute over) arguments. According to Prakken

(2018), formal works addressing this task are concerned with two main aspects: The first is modelling the dialogue itself, i.e. the definition of *protocols* that structure the interaction. The second is the study of dialogue strategies, i.e. the selection of utterances in the dialogue. Although the work in this thesis includes contributions to both aspects, the focus of this section is clearly on dialogue models as the aforementioned contribution in this regard directly builds on the existing formal background. In contrast, the contribution to argument strategies is approached from the machine learning perspective and competing methods are hence summarized in related work instead of a detailed theoretical discussion in this section.

In the context of dialogue models in general and protocols in particular, the following terminology is used: The focus is on competitive dialogues with two interlocutors (or agents) which are consequently referred to as *proponent* and *opponent*. In the typical example of a persuasive dialogue, the goal of the proponent is to defend a topic or position against attacks through the opponent. Although some approaches limit the dialogue to the exchange of arguments, additional utterances are in general possible and, arguably, necessary for complex argumentative dialogues. Examples for such utterances can be the critical questions as a challenging response to argument schemes (Walton, 1996), as well as the general conceding to an argument of the opposition. For the sake of clarity, utterances of this kind are referred to as *moderating utterances* in contrast to *argumentative utterances* that introduce an additional piece of information in support of their position. The difference between the two is illustrated in Table 2.1.

Table 2.1: Example dialogue on the topic animal testing, including argumentative and moderating utterances.

| Speaker | Utterance | Type |
|---|---|---|
| Proponent | *Animal testing should be banned.* | argumentative |
| Opponent | *Why do you think that?* | moderating |
| Proponent | *Animal testing is cruel and violates animal rights, therefore it should not be allowed.* | argumentative |

As for protocols, distinctions can be made (Prakken, 2005) between *immediate-response* protocols, i.e. protocols that enforce a direct response to an utterance once it is introduced, and *delayed-response* protocols that allow the interlocutors to deviate from an unresolved issue in the discussion and to address it later. Furthermore, *single-move* and *multi-move* protocols are distinguished. Whereas single-move protocols restrict speakers on one utterance (argument, speech act, etc.) at a time, multi-move protocols allow a speaker to chain multiple utterances together. Finally, a separation between *single-response* and *multi-response* protocols is made, where single-response protocols allow only one response to an utterance and multi-response protocols, as the name suggests, multiple alternative ones. The theoretical background on the formal frameworks utilized in this thesis is discussed in the following, starting with dialogue games in general. Subsequently, the specific dialogue game instantiation that is utilized in later chapters is introduced in detail.

### Dialogue Games for Argumentation

In the following the theoretical background on dialogue games for argumentation is recalled, following the formal description introduced in (Prakken, 2005). Dialogue games in general are a model of conversation, meaning that they extend the formal approach of speech acts to their effect on the listener (Mann, 1988). A dialogue game for argumentation with two players $proponent, opponent \in \mathscr{P}$ can be described as tuple $(\mathscr{L}, D)$ with $\mathscr{L}$ a logic for defeasible argumentation that encodes the arguments available in the game and $D$ the dialogue system proper that includes the rules of the game.

**Definition 7** (Logic for Defeasible Argumentation)**.** *A logic for defeasible argumentation $\mathscr{L}$ is a tuple $(L_t, R, Args, \rightarrow)$ with $L_t$ a logical language, $R$ a set of inference rules over $L_t$, $Args$ a set of arguments that can be constructed from $L_t$ through the application of rules in $R$, and $\rightarrow$ defeat relations defined over $Args$.*

It can be seen from the definition, that $\mathscr{L}$ includes a structured model to represent and construct arguments which are then, in combination with the defeat relations $\rightarrow$ an instance of Dungs abstract argument framework (see Section 2.2.1). In particular, support relations between arguments are not considered. It should be noted that, in contrast to the definition of the abstract argumentation framework, the term defeat instead of attack is used for the relation between arguments but that both relations are formally analogous. Consequently, each argument is represented as AND-tree with nodes out of $L_t$ connected by links that are instantiations of inference rules in R. The root of an argument $\Phi_i$ is then called *conclusion* and denoted with $conc(\Phi_i)$ whereas the set of leaves is called *premises* and denoted with $prem(\Phi_i)$.

For example, an argument $\Phi_i = (\varphi_1, \varphi_2 \Rightarrow \varphi_3)$ has the premises $prem(\Phi_i) = \{\varphi_1, \varphi_2\}$ and the conclusion $conc(\Phi_i) = \varphi_3$ with $\varphi_1, \varphi_2, \varphi_3 \in L_t$. Furthermore, an argument $\Phi_i$ is called an extension of another argument $\Phi_j$, if $conc(\Phi_i) \in prem(\Phi_j)$. Throughout this thesis, $\mathscr{L}$ is said to encode the argument structure of the dialogue game.

**Definition 8** (Dialogue System Proper)**.** *A dialogue system proper $D$ is a 3-tuple $(L_c, P, C)$, with $L_c$ a communication language, $P$ the game protocol and $C$ the commitment rules of the game.*

The communication language $L_c$ is a set of speech acts (or locutions) of the form $\beta = per(c)$, where *per* is a performative referred to as speech act type and $c$ is either an element of $L_t$, a subset of $L_t$ or an element of $Args$. Over $L_c$, two binary and irreflexive relations $R_a$ and $R_s$ are defined that determine *attacking* and *surrendering* replies and fulfill the following conditions:

- If a speech act $\beta_i$ attacks a speech act $\beta_j$, it does not surrender to another speech act in $L_c$.
- If a speech act $\beta_i$ surrenders to a speech act $\beta_j$, no speech act in $L_c$ attacks $\beta_i$.

It follows from these conditions that replying speech acts can be divided into attacking and surrendering replies and that surrendering replies cannot be attacked. In addition, each $(\beta_i, \beta_j) \in R_s$ is assigned one or more counterparts $(\beta_k, \beta_j) \in R_a$ by a function $cp : R_s \rightarrow 2^{R_a}$. In this case, $\beta_k$ is denoted as an *attacking counterpart* of $\beta_i$ in the following.

For the discussion of the protocol $P$, the concepts of game moves and dialogues in the game are required and consequently introduced next. A dialogue game for argumentation is played in turns, whereas each turn can consist of one or multiple game moves. A game move is of the form

$$m_k = (k, \beta_k, p_k, \tau_k), \tag{2.3}$$

where $k$ is a temporal identifier, $\beta_k \in L_c$ the speech act uttered in the move, $p_k \in \mathscr{P}$ the corresponding player and $\tau_k$ the temporal identifier the current move refers to. As a direct consequence, each move can reply to exactly one earlier move. In case a move does not reply to another specific move, as for example the opening move $m_1$, $\tau_k = 0$ is used to denote an *empty target*. A temporally ordered sequence of moves $d_k = m_1, ..., m_k$ where each move except for $m_1$ refers to an earlier move is then referred to as *dialogue*. In the scope of this work, let $M$ denote the set of all moves and $M^{<\infty}$ the set of all possible, finite length dialogues. The continuation of a dialogue $d$ with a move $m$ is denoted with $d, m$ and the empty dialogue is denoted with $d_0$.

For each dialogue $d \in M^{<\infty}$, the protocol of the game $P$ determines the current player to move (turn-taking), the set of legal moves available to this player and the (current) outcome of the game. The turn-taking is formalized through a turn-taking function $T : M^{<\infty} \rightarrow 2^{\mathscr{P}}$ that maps a dialogue to a set of current players to move (in this general case, multiple speakers are allowed at the same time), whereas the set of legal moves $M_d \subseteq M$ is determined by a protocol function $Pr : M^{<\infty} \rightarrow 2^M$. Given a continuation of a dialogue $d_{k-1}, m_k$, Prakken (2005) defines a set of rules that are assumed to be met by the protocol:

- $P_1 : p_k \in T(d_{k-1})$ (the player of $m_k$ is the current player to move).
- $P_2$ : If $d_{k-1} \neq d_0$, then $(\beta_k, \beta_{\tau_k}) \in R_s$ or $(\beta_k, \beta_{\tau_k}) \in R_a$ (the response to an earlier move is allowed in the communication language, if $m_k$ is not the opening move).
- $P_3 : p_k \neq p_{\tau_k}$ (players do not respond to their own moves).
- $P_4$ : If there is a $m_i$ in $d_{k-1}$ with $\tau_k = \tau_i$, then $\beta_k \neq \beta_i$ (a repeated response has to be different from any earlier response, moves cannot be repeated).
- $P_5$ : If $m_k$ is an attacking reply and there is a $m_i$ in $d_{k-1}$ with $(\beta_i, \beta_{\tau_k}) \in R_s$, then $\beta_k$ is not an attacking counterpart of $\beta_i$ (a surrendering reply cannot be reversed, or, intuitively, accepted argument components cannot be challenged later).

The last component of the dialogue system proper $D$ are the commitment rules. As pointed out in the referenced literature (Prakken, 2005), commitments are statements or, formally, argument components, that a player has publicly accepted - e.g. by introducing them to the dialogue or by conceding to them. In contrast to beliefs, as for example included in the previously discussed bipolar weighted argumentation framework by Rosenfeld and Kraus (2016), commitments do not necessarily imply that the speaker is convinced by them or even (personally) supports them. For example, a debater in an oxford-style debate will (publicly) commit to statements that support his or her assigned position, while it is possible that he or she is at the same time not convinced by them on a personal level. In the dialogue game framework, commitments are determined by a commitment function $C : M^{<\infty} \times \mathscr{P} \rightarrow 2^{L_t}$ that assigns players sets of argument components they

Table 2.2: Communication language of the dialogue game for relevant dialogues with $\Phi_i, \Phi_j \in Args$ and $\varphi_i, \varphi_j \in L_t$.

| Speech Act | Attacks | Surrenders |
|---|---|---|
| *claim*($\varphi_i$) | *why*($\varphi_i$) | *concede*($\varphi_i$) |
| *why*($\varphi_i$) | *argue*($\Phi_j$) ($conc(\Phi_j) = \varphi_i$) | *retract*($\varphi_i$) |
| *concede*($\varphi_i$) | - | - |
| *retract*($\varphi_i$) | - | - |
| *argue*($\Phi_i$) | *why*($\varphi_i$) ($\varphi_i \in prem(\Phi_i)$), *argue*($\Phi_j$) ($\Phi_j \rightarrow \Phi_i$) | *concede*($\varphi_i$) ($\varphi_i \in prem(\Phi_i)$ or $\varphi_i = conc(\Phi_i)$) |

committed to throughout the dialogue. Based on this general formal definition, specific dialogue games can then be instantiated through a specification of the communication language and outcome rules in combination with definite commitment rules and protocols, as also demonstrated in some examples in (Prakken, 2005).

### Prakken's Dialogue Game for Relevant Dialogues

In the following, Prakken's framework for *relevant dialogues* defined in (Prakken, 2000; Prakken, 2005) is discussed thoroughly. As Prakken instantiates his frameworks step-wise through the subsequent introduction of additional specifications to earlier instantiations, all components of the investigated framework are herein summarized without specific reference to the instantiation they were originally introduced for. The corresponding communication language $L_c$ is shown in Table 2.2, ordered by attacking and surrendering replies. Attacking counterparts are shown in the same line as the corresponding surrendering reply with one exception: $argue(\Phi_j)$ is only an attacking counterpart of $concede(\varphi_i)$ if $conc(\Phi_j) = \neg\varphi_i$. The main motivation behind the framework is, as the name suggests, to limit the dialogues in the game to relevant ones. To this end, only moves that influence the logical acceptability of the initial move (i.e. the discussed topic) can be addressed by the player to move. This property is formalized through a binary status assigned to each played move, defining it as either *in* or *out*. Similar to the approach of labelling-based semantics in the context of abstract argumentation frameworks, a move is *in* in a dialogue $d_k$ if no attacking reply that is also *in* is included in $d_k$. Consequently, freshly introduced moves, as well as moves without a reply are automatically *in*. The goal of both players is then to switch the status of the initial move in their favour, i.e. *in* for the proponent and *out* for the opponent. The *relevance* of a move is then defined as follows:

**Definition 9** (Relevant Target)**.** *If an attacking reply to a move $m_i$ in a dialogue $d_k$ switches the status of the initial move $m_1$, $m_i$ is called relevant or (interchangeably) a relevant target.*

Turn-taking is then determined also by the status of the initial move, meaning that all players have the obligation to move until the status of the initial move is switched in their favour. The concept of relevance can be illustrated with the following example. The dialogue is initialized by the proponent with an *argue* move:

Prop.: *Animal testing should be banned since it's cruel and inefficient.* ($m_1$, in, relevant)

This move is *in* and relevant since it has no attacking reply on it. If the opponent then responds with an attacking *argue* reply, the status of the initial move is switched and the new move is *in* and relevant:

Prop.: *Animal testing should be banned since it's cruel and inefficient.* ($m_1$, out, not relevant)

Opp.: *Well, I think animal testing is important for the development of new medicine.* ($m_2$, in, relevant)

As the status of the initial move is switched in favour of the opponent after this move, it is now the proponent's turn and the only relevant target is $m_2$. The following *why* response attacks this move and consequently again switches the status of the initial move in favour of the proponent:

Prop.: *Animal testing should be banned since it's cruel and inefficient.* ($m_1$, in, relevant)

Opp.: *Well, I think animal testing is important for the development of new medicine.* ($m_2$, out, not relevant)

Prop.: *Could you please elaborate?* ($m_3$, in, relevant)

Consequently, it is again the turn of the opponent, who can now decide between two relevant moves ($m_1, m_3$) that can be addressed. The difference between the status and the relevance of a move becomes evident in the following case where the opponent decides to backtrack and challenge the initial move instead of the last one:

Prop.: *Animal testing should be banned since it's cruel and inefficient.* ($m_1$, out, not relevant)

Opp.: *Well, I think animal testing is important for the development of new medicine.* ($m_2$, out, not relevant)

Prop.: *Could you please elaborate?* ($m_3$, in, not relevant)

Opp.: *Well I think we can discuss this point in more detail later. But first I would like to know why you think that animal testing is cruel and inefficient.* ($m_4$, in, relevant)

Since there is no attacking reply to $m_3$, it remains in. However, as $m_4$ challenges the initial move directly, an attack on $m_3$ would only reinstate $m_2$ and not switch the status of the initial move. Consequently, the relevance criterion implicitly enforces $P_3$ (players do not respond to their own moves). The example dialogue, including status and relevance, is illustrated in its abstract form in Figure 2.7. The *current winner* of a dialogue is defined as the player who is not to move, as the initial move and hence the topic under discussion has a (logical) validity matching his or her position. The dialogue game is terminated if a player runs out of legal moves without changing the status of the initial move in his or her favour and this player, consequently, loses the game. Formally, these specifications are defined through adding the following rules to the game protocol, assuming again a continuation of a dialogue $d_{k-1}, m_k$:

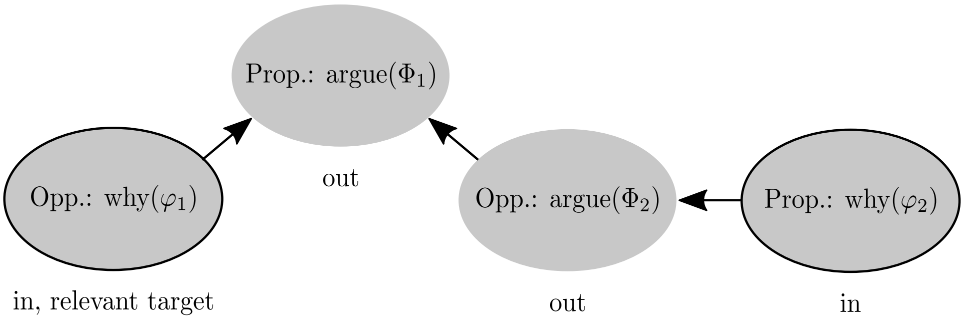


Figure 2.7: Illustration of the relevance criterion. Both *why* moves are not attacked and therefore *in* (indicated by black margins of the circles). Consequently, their targets are *out*. Only the *why(*$\varphi_1$*)* move is a relevant target since an attack on it would change the status of the opening move *argue(*$\Phi_1$*)*. Therefore, it is the turn of the proponent.

- $P_6 : T(d_0) = proponent$ and $T(d_{k-1}) = p, p \in \mathscr{P}$ if $p$ is not the current winner of $d_{k-1}$ and $d_{k-1} \neq d_0$ (the player to move is the one who has to switch the status of the initial move).
- $P_7$ : If $d_{k-1} = d_0$, $\beta_k$ has the type *argue* or *claim* (the game is initialized with the introduction of the topic).
- $P_8$ : If $m_k$ replies to $m_i$ in $d_{k-1}$, then $m_i$ is relevant in $d_{k-1}$ (only relevant targets can be addressed by the player to move).

Although $P_7$ is not directly related to the relevance criterion, it ensures that the game is initialised with the topic and the remaining rules can be applied. Moreover, it follows from the turn taking-rule $P_6$ that players can play an unspecified number of surrendering moves, followed by a single attack move. Consequently, the protocol is a multi-move, multi-response and delayed-response protocol. Detailed examples of dialogues generated with this framework are included in later chapters when the discussed model is used to structure the interaction between two virtual agents.

### 2.2.3 Argument Quality

In contrast to the models in the previous sections, the study of argument quality is not concerned with modelling arguments (or argumentation) but with their assessment. Whereas the term argument evaluation is also used to refer to the study of conclusions that can logically be drawn from a set of arguments, e.g. in the case of semantics, the assessment of arguments is concerned with (mostly) natural language argumentation and its overall quality (for example, its strength). In comparison to the previously discussed models, it is evident that this assessment is to a certain degree subjective, as arguments are perceived differently by individuals. As a consequence of this subjectivity, some approaches assess argument quality through human ratings (Habernal and Gurevych, 2016b) and are hence referred to as *rating-based approaches*. On the other hand, a variety of of works was dedicated to argument quality aspects that can be objectified, in the following denoted with *theory-based approaches*. The majority of these works were recently summarized in a single taxonomy by Wachsmuth et al. (2017b) which is discussed in this section. In the context of the present thesis, no approach (neither theory-based nor rating-based) is preferred over the other since it was shown that the theoretical insights are reflected in the human ratings (Wachsmuth et al., 2017a). Instead, the assumed position is that the theoretical insights help in identifying the right questions for human raters.

In their taxonomy, Wachsmuth et al. (2017b) investigate argument quality on different levels of granularity, namely *argument components* (or units), *arguments*, *monological argumentation* (i.e. argumentative speech) and *dialogical argumentation*. In addition, they identify three broad aspects of argument quality:

- *Logical quality* denotes the logical consistency of an argument or argumentation, also including its (formal) strength.
- *Rhetorical quality* is the effectiveness of arguments or argumentation in their ability to persuade the target audience.
- *Dialectical quality* refers to the contribution of an argument or argumentation to the resolution of the topic argued over.

The authors then discuss existing work on all three aspects and also different dimensions that can influence each of them, resulting in the proposed taxonomy with a total of 15 fine-grained sub-dimensions. For the definition of these dimensions, the scenario of a single speaker addressing an audience is considered. The logical quality is included in the quality dimension *cogency* and defined as follows:

**Definition 10** (Cogency)**.** *An argument is cogent if its premises are relevant for the conclusion, sufficient to draw the conclusion and rationally worthy of acceptance.*

It can be seen that the cogency dimension is hence defined for individual arguments as properties of their argument components. The three sub-dimensions of cogency are, as included in the definition, the *local relevance*, *local sufficiency* and *local acceptability* of an argument's premises. The rhetorical quality is reflected in the quality dimension *effectiveness* and defined as:

**Definition 11** (Effectiveness)**.** *Argumentation is effective if it is able to persuade the audience of the author's position.*

This definition is rather general as several different influences can contribute to the persuasion of a target audience. They are included in the taxonomy in five sub-dimensions of effectiveness. The first one is referred to as *credibility* and argumentation is said to be credible if it is presented in a way that supports the author's authority on the issue. The second dimension is denoted as *emotional appeal* and refers to argumentation that is emotionally appealing to the audience. The third dimension is *clarity* and indicates how structured the argumentation is and if it deviates from the topic under discussion. The *appropriateness* of argumentation indicates if it uses language that supports its credibility, its emotional appeal and if it is appropriate for the topic under discussion. Finally, the structure of the argumentation, i.e. if it presents the issue and related arguments in a comprehensive manner, is addressed by the dimension *arrangement*. The effectiveness and its sub-dimensions are all defined on the level of argumentation, i.e. on the level of the complete argumentative speech, and include (in contrast to the formal models above) also effects that are related to the speaker instead of the argument content. Especially the influence of emotions and subliminal effects that contribute to persuasion were also investigated in psychology. Respective persuasion models that describe the process of persuasion from the human perspective can be found for example in (Van Kleef et al., 2015).

The dialectical quality is represented in the dimension *reasonableness* and defined as:

**Definition 12** (Reasonableness)**.** *Argumentation is reasonable if it sufficiently supports the resolution of the issue in a way that is accepted by the audience.*

Again, this dimension is defined on the level of argumentation and includes (as indicated by the definition) three sub-dimensions: *Global acceptability* indicates if the stated arguments and the way they are proposed are accepted by the audience. The dimension of *global relevance* assesses the contribution of the argumentation to an overall resolution of the issue, whereas *global sufficiency* indicates if the argumentation adequately addresses known counter-arguments.

Despite their definitions on different levels of granularity, all 15 dimensions were used by the authors to annotate an argument quality corpus. In the discussion of the results, it is emphasised that the dimensions are not independent and influence each other. Moreover, the subjectivity of the annotation task is highlighted as one of the main issues during annotation, also indicating that the assessment of arguments and argumentation is, as mentioned at the beginning of the section, subjective to a certain degree.

### 2.2.4 Argument Mining

The field of argument mining is concerned with the automatic extraction of arguments and argument structures from natural language sources and has, similar to other natural language processing areas, seen a lot of progress recently. Due to its increasing popularity, a variety of approaches that utilize target structures with different levels of granularity were provided. Extensive overviews of the overall field were provided for example by Lippi and Torroni (2016) and Lawrence and Reed (2020). Following the discussion by Lawrence and Reed (2020), argument mining includes the following three sub-tasks:

- Automatic identification of argument components or units, i.e. text parts that are considered to be *argumentative*.
- Automatic recognition and analysis of properties of argumentative text spans.
- Automatic identification of relations between argument components.

In addition, argument mining was divided into the two broad groups of *discourse-level argument mining* which recognizes argument structures within a single document, such as essays (Stab and Gurevych, 2014b), and *information-seeking argument mining* which considers a multitude of heterogeneous sources (Trautmann et al., 2020). Approaches for *discourse-level argument mining* typically extract the complete argument structure of their source, for example utilizing the claim-premise model discussed above. In doing so, they cover at least two of the three previously mentioned argument mining tasks, i.e. the recognition of argument components and the identification of relations. In contrast, *information-seeking argument mining* is focused on recognizing argumentative text spans that are concerned with a given topic in various sources and only estimate their binary polarity regarding this topic. A recently evolved application for *information-seeking argument mining* are argument search engines (Stab et al., 2018a; Wachsmuth et al., 2017c). Similar to web search engines, the respective system retrieves a ranked list of pro and con arguments concerned with a specific search query from a large corpus of documents. In doing so, it includes argument mining but also document indexing as well as argument evaluation and different paradigms are pursued by the individual systems (Ajjour et al., 2019). In the scope of this thesis, argument search engines are considered as an approach to argument acquisition for dialogue systems.

## 2.3 Machine Learning

Machine learning has evolved as a sub-field of artificial intelligence and gained a lot of attention in recent research. In contrast to traditional computer programs, where pre-defined rules are coded by the programmer, machine learning systems infer programs from data (Domingos, 2012). Mathematically, a machine learning model approximates a function that maps data points (for example an image or the positions of pawns in a board game) to correct targets (for example a class or a game move). The function approximation is then a numerical optimization problem on a given data set or in a pre-defined setup with respect to a cost function that encodes the goal of the machine learning task. The optimization is usually referred to as training or learning of the model. Machine learning has shown remarkable success in different domains and scenarios like for example natural language processing (Devlin et al., 2019) and games like go (Silver et al., 2016). Generally, machine learning techniques are divided into three groups:

- *Supervised learning* is concerned with problems where labelled data is available that represents the patterns that should be inferred. A supervised learning model is then trained to recognize the labels from the data in a way that allows it to infer the same patterns in new data (generalization). The two most popular tasks addressed by supervised learning models are *classification* and *regression*.

- *Reinforcement learning* describes computer programs (or agents) that learn from experience, i.e. from a reward that is received for taking a certain action in a pre-defined environment. The training of a reinforcement learning agent is then the optimization of its strategy (or policy) with the aim to maximize the overall reward.

- *Unsupervised learning* models infer rules from data without specification through a label, including for example clusters or repeating patterns.

In the scope of this thesis, supervised learning and reinforcement learning techniques are used and the relevant theoretical background, as well as the specific approaches, are described in detail in the following sections.

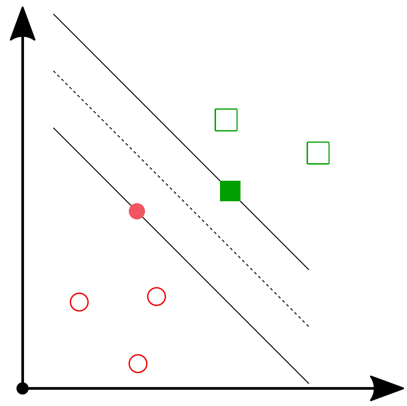

Figure 2.8: Illustration of the maximum margin classification into two classes through a SVM.

### 2.3.1 Supervised Learning

Supervised learning models are trained to generalize patterns indicated in a set of training data through labels. As supervised learning is only applied to classification tasks in the scope of this thesis, the discussion is limited to the background of this sub-task. The goal of a machine learning model in a classification task is to assign data points to pre-defined classes. An intuitive example is a division of emails into the classes *spam* or *no spam* based on their content. A perfect classifier is then a function $f : X \rightarrow Y$ that assigns each data point in $X$ the correct class in $Y$. In practice, data points are represented by a set of features that are used in the classification and $X$ is hence called the *feature space*. A machine learning-based approximation of this perfect classifier is then comprised of the three elements *model*, *evaluation* and *optimization* (Domingos, 2012). The *model* describes the representation of the function that is used to approximate the classifier. Depending on the chosen model, different classifiers (i.e. functions) can or cannot be learned and the choice of the model directly influences the model performance. Examples of frequently used models are support vector machines (Vapnik, 2013) and artificial neural networks (Zhang, 2000). *Evaluation* refers to the definition of a suitable evaluation metric that indicates the quality of a trained model. It should be noted that the metric used to measure the quality of a model does not need to be the same as the cost function used for training. Typical evaluation metrics that are used in the scope of this thesis are discussed in the following subsection. Finally, *optimization* refers to the optimization procedure used to train the classifier, i.e. to optimize the approximated function with respect to their cost-function. In the following, the models that are used in the scope of this thesis are introduced alongside the underlying optimization procedure. The applied evaluation metrics are discussed subsequently.

#### Support Vector Machines

Support vector machines (SVM) separate the data points through the hyperplane in the feature space with the maximum margin to samples on both sides as depicted in Figure 2.8. It is therefore also called a *maximum margin classifier* and the samples on the margin are denoted as *support* *vectors*. The basic concepts of the SVM classification are discussed in the following for a two-class problem with labels $y \in \{1, -1\}$. The decision function can then be described with the normal vector $W$ that is perpendicular to the hyperplane and some offset $b$ as

$$\mathbb{D}(x) = W^\top x - b\,, \tag{2.4}$$

where $x \in X$ denotes the classified data point. The assigned class is determined from the decision function as $y = sign(\mathbb{D}(x))$. The learning objective is hence to find the parameters $W$ that maximizes the width between the margins on both sides. This width can be derived geometrically as $\frac{2}{\|W\|}$ and the maximum margin is hence obtained for a minimized norm $\|W\|$.

Finding the minimum norm can then be formulated as quadratic optimization problem $\min\|W\|^2$ under the constraint

$$y_i\mathbb{D}(x) \geq 1, \tag{2.5}$$

which reflects the condition that all data points are on the side of the hyperplane corresponding to their label and not closer to the hyperplane than indicated by the margin. The case of equal one in the above formula corresponds to the support vectors on the margin. Mathematically, this is approached through the use of Lagrange multipliers as discussed in detail in (Boser et al., 1992). The resulting Lagrange function then reads as

$$L(\alpha, b) = \sum_i \alpha_i(1 - y_i b) - \frac{1}{2}\sum_i\sum_j \alpha_i\alpha_j y_i, y_j\langle x_i, x_j\rangle. \tag{2.6}$$

Here, $\alpha_i, \alpha_j$ denote Lagrange multipliers, $\alpha$ the vector of all multipliers and $y_i, y_j \in Y$ labels assigned to the corresponding data points $x_i, x_j \in X$. Moreover, the decision function can also be expressed in terms of Lagrange multipliers, labels and samples in the data set as

$$\mathbb{D}(x) = \sum_i \alpha_i y_i\langle x_i, x\rangle + b\,, \tag{2.7}$$

where $x \in X$ is the data point that is to be classified. Although both functions depend (formally) on all samples in the data set, the Lagrange multipliers for samples that are not on the margin turn out to be zero, which limits the dependency to the data points on the margin. The drawback of this approach to classification is that it can in this original form only deal with problems that are linearly separable, i.e. that can be separated by a hyperplane in the corresponding feature space. To enable classifications of problems that are only separable by a non-linear function, a transformation $\phi(x)$ into a higher dimensional space is required in which the data points are again linearly separable. The concept is illustrated in Figure 2.9. However, solving a quadratic optimization problem in a high (possibly infinite) dimensional space is prone to be computationally expensive.

This computational effort can be avoided in SVMs through a method usually called *kernel trick*: As can be seen from the above equations, the function to optimize and the decision function both do not depend directly on the data points $x_i, x_j$ but on an inner product of data points. The inner product of two vectors in the transformed space however can be expressed as a function of the original vectors through a kernel

$$K(x_i, x_j) = \langle\phi(x_i), \phi(x_j)\rangle. \tag{2.8}$$

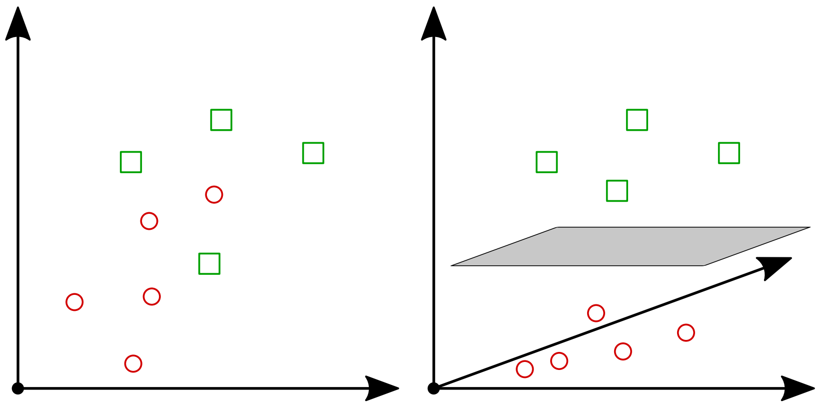

Figure 2.9: Illustration of the maximum margin classification into two classes using the transformation of features into a higher dimensional space.

Given that the kernel function for the corresponding transformation is known, the inner product can then be directly computed from the data points in the original feature space. It can be seen from the above derivation that a perfectly separable data set is assumed (each data point is on the correct side of the hyperplane). This case is referred to as *hard margin SVM*. However, this assumption is unrealistic in most cases due to factors like labelling errors or noise. The *soft-margin SVM* addresses this issue through the addition of the *hinge loss* to the cost function, yielding the optimization problem

$$\min \left[ \frac{1}{N} \sum_i \max(0, 1 - y_i \mathbb{D}(x_i)) + \Lambda \|W\|^2 \right], \tag{2.9}$$

where $\Lambda$ is the soft margin parameter that regulates the trade-off between a large margin and correct predictions. Similar to the hard margin case, the optimization problem can be reformulated through the introduction of constraints and the use of Lagrange multipliers and the above-discussed kernel trick can hence also be applied in the case of a soft margin SVM.

### Random Forest

Random forest classification is an ensemble method based on decision trees, meaning that an ensemble of classifiers is trained and the prediction about a data point $x$ is then determined through the majority vote in the ensemble. A decision tree is comprised of leave nodes that correspond to a decision of the classifier and decision nodes that split the data according to a decision rule. The decision rules at the corresponding nodes can be learned from data by comparing different splits with respect to the information gain obtained from the split. To this end, the Shannon entropy (Shannon, 1948)

$$H(\tilde{Y}) = -\sum_i \mathbb{P}(y_i) \log(y_i) \tag{2.10}$$

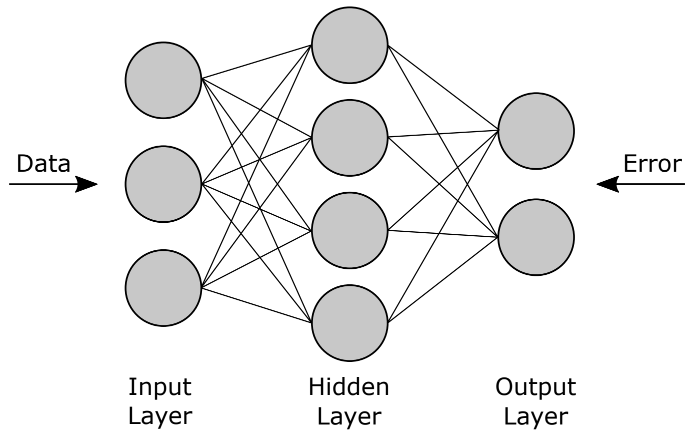


Figure 2.10: Illustration of a multi-layer perceptron with three layers. Data is processed from left to right, whereas the error is propagated backwards through the network.

can be utilized as a typical measure of the level of uncertainty. Here, $\mathbb{P}(y_i)$ is the probability with which the class $y_i$ occurs in the set $\tilde{Y}$. The *information gain* resulting from a split is then computed as the difference between the entropy at the decision node (parent) and the (combined) entropy at the child nodes. The advantage of decision trees in comparison to other classifiers is that they can learn non-linear patterns and do not require a re-scaling of the data points $x$. On the downside, they are prone to learn overly complex models that include artificial patterns in the data. This is usually referred to as *overfitting*.

Random forest classification addresses this issue through training multiple decision trees simultaneously on different random subsets of the data. In addition, randomized sub-spaces of the feature space can be used to train the different classifiers on a limited amount of features and reduce the chance of overfitting further. The final decision is then obtained, as mentioned above, by taking the majority vote of the forest. In comparison to other classifiers like SVMs with non-linear kernels and artificial neural networks, the results of a decision tree (and hence of the random forest) are interpretable, i.e. an explanation for the classification based on the features can be provided.

### Artificial Neural Networks

Artificial neural networks denote machine learning models with an architecture inspired by neural models of the human brain (Rosenblatt, 1958). They are comprised of a collection of nodes (or neurons) that are connected through directed, weighted links (weights). Each node outputs a real-valued signal that is processed (weighted) through the links to connected neurons. Consequently, each node can receive inputs from multiple others and transmit its signal to multiple others. The output of a node is then derived from the sum of its inputs. It can be adjusted by a bias and is frequently subject to a non-linear *activation function* like the sigmoid or step-function. A group of nodes on the same level in the network is referred to as a layer and if multiple (more than two) layers are stacked in a model, it is called a *deep neural network*.

In the following, the example of a multilayer perceptron (MLP) as depicted in Figure 2.10 is used to illustrate the concepts. In this case, a (single) layer is formally described as

$$\mathrm{act}(Wx+b) = y\,, \tag{2.11}$$

with $x$ the input of the layer, $y$ the output of the layer, $W$ the weight matrix encoding the weighted links, *act* the activation function and $b$ a bias vector. It can be seen that in this case, information is processed in one direction only through the network and architectures of this kind are referred to as *feed-forward networks*. During training, the weights of a network are adjusted to minimize the error computed from the difference between the network output and the expected output. A frequently-used approach for feed-forward architectures is propagating the error back through the network (backpropagation), as also illustrated in Figure 2.10. The strength of artificial neural networks lies in their flexibility, as the corresponding architecture can be adapted to the task at hand. In addition, the capability of learning different representations of the input data in hidden layers can replace (to a certain extent) manual feature engineering.

Throughout this work, the Bidirectional Encoder Representations from Transformers (BERT) model (Devlin et al., 2019) is used which provides a pre-trained language model that can be fine-tuned on different NLP tasks. The architecture is comprised of stacked Transformer encoder layers introduced in (Vaswani et al., 2017). The encoder blocks are build of two sub-layers with the general form

$$y = \mathrm{norm}(x + f_{sublayer}(x)), \tag{2.12}$$

where $x, y$ denote input and output of the layer, *norm* the normalization applied on the sub-layer and $f_{sublayer}$ the function that defines the operations in the layer. The first sub-layer of the encoder block is a multi-head self-attention layer whereas the second is a feed-forward layer. Both are discussed in detail in the following. A self-attention layer utilizes the input $x$ in three different roles: As queries $x^q = W_q x$, keys $x^k = W_k x$ and values $x^v = W_v x$ with dimension $dim_k$ (queries and values) and $dim_v$ (values). Here, $W_q, W_k, W_v$ denote (trainable) weight matrices. In practice query, key and value vectors are combined in matrices $X_q, X_k$ and $X_v$ which allows computing the self-attention as

$$f_{att}(X_q, X_k, X_v) = \mathrm{softmax}\left(\frac{X_q X_k^\top}{\sqrt{dim_k}}\right) X_v\,, \tag{2.13}$$

where the scaling of the dot product $\sqrt{dim_k}$ is added to facilitate learning and $softmax$ denotes the softmax normalization. In the case of multi-head attention, several attentions (called attention heads) are computed with different matrices $W_q^i, W_k^i, W_v^i$ in parallel. The resulting output is then concatenated (denoted with *conc*) and projected back to the model dimension with a matrix $W_o \in \mathbb{R}^{h \cdot dim_k \times dim_v}$. Consequently, it is expressed as

$$f_{multihead}(X) = \mathrm{conc}(f_{att}^1(X_q^1, X_k^1, X_v^1), \ldots, f_{att}^h(X_q^h, X_k^h, X_v^h))W_o\,. \tag{2.14}$$

In the case of BERT, this enables the model to take into account the different roles a word can play in different contexts. The fully connected layer on the other hand is given as

$$f_{connected} = \max(0, xW_1 + b_1)W_2 + b_2\,, \tag{2.15}$$

where $W_1, W_2$ and $b_1, b_2$ denote weight matrices and bias vectors, respectively. In the case of BERT, the input vector can be either a sentence (i.e. a text span) or a sentence pair which is encoded using WordPiece embeddings (Wu et al., 2016). At the beginning of each input sequence, a special $[CLS]$ token is added. The representation of this token in the final hidden layer is later used for classification tasks. In addition, sentence pairs are separated by a token $[SEP]$. The pre-training of BERT is then done in an unsupervised manner on two different tasks, namely predicting masked words in a sentence and predicting whether, given a pair of sentences, the second one can follow the first one. For fine-tuning a pre-trained model, an additional output layer that suits the corresponding task can be added and the complete model is adjusted end-to-end through supervised learning to the task. In addition, the sentence embeddings of the pre-trained model can be used as input to a new classifier. Within this work, fine-tuning on classification tasks and the use of BERT sentence embeddings as features are considered. For classification tasks, the representation of the $[CLS]$ token is utilized by the output layer to predict the class.

### Evaluation Metrics

As the goal in classification tasks is generalization, the available labelled data is frequently split into different sets to evaluate the model on data that is not utilized during training. The error of a supervised learning model is typically discussed in terms of *variance* and *bias*. A high bias indicates that the corresponding model oversimplifies in the sense that it does not capture dependencies relevant to the problem. It is indicated by a low performance on both training and testing data and the effect is also referred to as *underfitting*. In contrast, a high variance indicates a model that is too complicated and captures artificial dependencies which is referred to as *overfitting*. An indicator for high variance is a good performance on the training data and a comparatively low performance on the testing data. For most problems, there is a trade-off between variance and bias as minimizing one usually increases the other.

In the following, evaluation metrics that are used to assess the quality of a trained classifier are discussed. As only binary classification problems are addressed in the scope of this thesis, the discussion is focused on this case but the metrics can be applied to multi-class problems in a similar manner. In terms of notation, the classes are generalized as a positive class and a negative class without specifying the actual labels. Given a set of predictions, the different metrics are derived from the *confusion matrix*

$$ConfMatrix = \begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix}, \tag{2.16}$$

where $TP$ (true positives) denotes the number of samples that are correctly assigned to the positive class, $FP$ (false positive) the number of samples incorrectly assigned to the positive class, $FN$ the number of samples that are incorrectly assigned to the negative class and $TN$ (true negative) the number of samples predicted correctly as members of the negative class.

The accuracy of a classifier measures the percentage of correct predictions and is hence given as

$$ACC = \frac{TP + TN}{TP + FP + FN + TN}. \tag{2.17}$$

Although frequently used, it can be misleading for problems with highly imbalanced data. If for example 80% of the data samples in the test set correspond to the positive class, a classifier that always predicts this majority class reaches an ACC of 80%, although no distinction between the samples on the basis of features is made. Moreover, depending on the respective task, the correct prediction of members of a specific class can be considered more important than overall high accuracy. For example in the case of medical tests, it is more important to correctly identify subjects that carry disease, even if they represent only a small fraction of the tested subjects. On the other hand, the rate of incorrect positive predictions should also be as small as possible to avoid for example unnecessary treatment or quarantine. Typical measures that address these issues and indicate the classifier performance with respect to a relevant class are *recall* and *precision*. In the following, it is assumed that the positive class is the relevant one but both metrics can be analogously computed for the negative class as well. The *recall* measures the percentage of correctly predicted positive samples and is defined as

$$recall = \frac{TP}{TP+FN}. \tag{2.18}$$

In the above example, the recall is the ratio of subjects that are correctly predicted to carry the disease and the overall number of subjects that carry the disease. However, it does not include information about the amount of false-positive predictions and a classifier that assigns all samples the positive class would reach a *recall* of 100%. In contrast, the precision measures the percentage of correct positive predictions and is defined as

$$precision = \frac{TP}{TP+FP}. \tag{2.19}$$

In the above example, the precision corresponds to the ratio of subjects that were correctly predicted to carry the disease and the overall number of subjects that were predicted to carry the disease. Hence, it also provides information regarding the number of subjects that are incorrectly predicted as positive. However, it does not take into account the overall amount of relevant samples meaning that a classifier that correctly predicts a single subject as positive and all others as negative also reaches a *precision* of 100%.

A measure that combines both aspects covered by *recall* and *precision* is the *F1-score*

$$F1 = \frac{2TP}{2TP+FN+FP}. \tag{2.20}$$

It considers both, the amount of overall positive samples and the amount of overall positive estimates. In cases without a specific relevant class but imbalanced data, a measure of the overall performance can be derived by averaging the class-wise recalls. This is referred to as unweighted average recall (UAR) and defined as

$$UAR = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right). \tag{2.21}$$

It can be seen that for the above discussed example with the majority classifier, the corresponding UAR would be only 50% which shows that the classification is not better than random if the classes are equally distributed.

In addition to the performance metrics discussed above, measures for *agreement* and *correlation* between sets of ratings are also frequently used in the context of supervised learning. They can be applied to measure the agreement/correlation between the ratings of a human annotator and a trained machine learning model or as a reliability measure for the generation of training data from human annotations. In the scope of this thesis, Fleiss' kappa (Fleiss, 1971) and Krippendorf's alpha (Hayes and Krippendorff, 2007) are used to measure agreement whereas linear correlations are investigated by means of the Pearson correlation coefficient (Rodgers and Nicewander, 1988). Fleiss' kappa is an extension to Cohen's kappa (Cohen, 1960) for the case of three or more raters. In addition, not all instances have to be rated by the same annotators but all instances need to have the same amount of ratings. Consequently, Fleiss' kappa is not capable of dealing with missing data. The agreement is computed as

$$\kappa_{Fleiss} = \frac{p_a - p_c}{1 - p_c}, \tag{2.22}$$

where $p_a$ is the agreement between the annotators and $p_c$ is the agreement by chance. Here, the denominator indicates the agreement obtainable above chance whereas the numerator is the actual agreement above chance. Consequently, $\kappa_{Fleiss}$ ranges from -1 to 1, where 1 indicates perfect agreement, 0 an agreement by chance and -1 complete (systematic) disagreement. The formula is analogous to Cohen's kappa and only the derivation of the agreement values $p_a$ and $p_c$ is adapted for the more general case.

Krippendorf's alpha was introduced as a general reliability measure and can be applied to ordinal, nominal, interval and ratio data (Hayes and Krippendorff, 2007). In addition, it imposes no restriction regarding the number of raters and is able to handle missing data, meaning that not all instances need to have the same amount of ratings. In contrast to the above-discussed kappa, it relies on a measure of *disagreement* and is computed as

$$\alpha_{Kripp} = 1 - \frac{D_o}{D_e}, \tag{2.23}$$

where $D_o$ indicates the observed disagreement in the data and $D_e$ the disagreement expected by chance. Again, a value of 1 indicates perfect agreement, a value equals 0 corresponds to an agreement by chance and a value smaller than 0 indicates systematic disagreement.

The Pearson correlation coefficient (PCC) measures the linear correlation between two sets of variables $x_1$ and $x_2$ with dimension $n$ as

$$PCC = \frac{\sum_{i=1}^{n}(x_{i,1} - \bar{x}_1)(x_{i,2} - \bar{x}_2)}{\sqrt{\sum_{i=1}^{n}(x_{i,1} - \bar{x}_1)^2}\sqrt{\sum_{i=1}^{n}(x_{i,2} - \bar{x}_2)^2}}, \tag{2.24}$$

where $\bar{x}_{1,2}$ denotes the average of the corresponding variables. The PCC values range from -1 to 1, where -1 indicates a perfect negative correlation, 1 a perfect positive correlation and 0 no linear correlation.

### 2.3.2 Reinforcement Learning

In reinforcement learning (RL), the objective of the learning agent is to optimize the selection of actions in an environment with respect to feedback, referred to as reward. In contrast to supervised and unsupervised approaches, the agent does not directly learn from input/output data but from the feedback signal and reinforcement learning can hence be (informally) described as *learning from experience*. It is usually applied as an approximative method for solving markov decision processes (MDPs) or markov games in scenarios where the solution cannot be computed analytically, for example in the case of incomplete information (transition and/or reward function unknown) or increasing dimensionality. Most work on RL is focused on the single-agent case and single-agent learners are thus well studied (Sutton and Barto, 1998). In the following, the formal background on MDPs as well as markov games is introduced, followed by a discussion of learning methods to solve them.

### Formalization

In scenarios with a single agent, the learning can be formally described as MDP. It is defined as $(S, A, r, T, \gamma)$, where $S$ denotes a set of environment states and $A$ the set of actions the agents can choose from. The reward $r : S \times A \rightarrow \mathbb{R}$ determines the return the agent receives for taking a certain action in a given state, whereas the transition function $T : S \times A \times S \rightarrow [0, 1]$ determines the (possibly probabilistic) state transitions given a state-action pair. Finally, the discount factor $\gamma$ weights future rewards against the current reward. The objective of the agent is to find the policy $\pi$ that maximizes the sum of expected future rewards

$$V_\pi(s) = \mathbb{E}(\sum_{t=k}^{\infty} \gamma^t r_t \mid s_k = s), \tag{2.25}$$

where $V_\pi(s)$ is the so called value function and $r_t = r(s_t, a_t)$ denotes the reward received at time $t$. The policy is either a mapping $\pi : S \rightarrow A$ (deterministic) or a probability distribution $\pi : A \times S \rightarrow [0, 1]$ (probabilistic).

A discounted markov game is an extension of MDPs to multi-agent scenarios. Following the discussion of Barlier et al. (2015), it is described as tuple $(I, S, \mathbf{A}, \mathbf{r}, T, \gamma)$ with $I = \{1, ..., n\}$ the agents, $S$ again the set of states or state space, $\mathbf{A} = A_1 \times ... \times A_n$ the joint action space that includes the (sub)set of actions available to player $p \in I$ in the respective state $A_p(s)$ and $\gamma$ the above discussed discount factor. The reward function $r^p : S \times \mathbf{A} \rightarrow \mathbb{R}$ determines the real valued reward, given the current state $s$ and the current actions of all players. It is included in the joint reward function $\mathbf{r} = r^1 \times ... \times r^n$ of all players. The transition probability function $T : S \times \mathbf{A} \times S \rightarrow [0, 1]$ determines the probability for reaching a state $s'$, given the current state $s$ and the actions of all players. It can be seen from this definition that the special case of $n = 1$ corresponds to the formal description of a MDP.

Again, each agent selects actions according to a (possibly stochastic) policy function $\pi^p : A_p \times S \rightarrow [0, 1]$. The goal of each agent is to find the policy that maximizes its expected discounted sum of future rewards

$$V^p_{\boldsymbol{\pi}}(s) = \mathbb{E}(\sum_{t=k}^{\infty} \gamma^t r^p_t \mid s_k = s), \tag{2.26}$$

with $\boldsymbol{\pi} = \pi^1 \times ... \times \pi^n$ the joint strategy of all agent and $V_p$ the value function of agent $p$. A joint policy $\boldsymbol{\pi}$ is defined to be optimal if each agents policy $\pi_p$ is optimal with respect to the policies of all other agents. This is called a Nash equilibrium and discounted markov games were shown to posses at least one in stationary policies (Filar and Vrieze, 2012). A special case of markov games is the two-player zero-sum game, where two agents have opposite goals, i.e. where the gain of one agent is equivalent to the loss of the other. In this case, the reward for both agents, any state $s$ and any actions $a^1, a^2$ can be written as $r^1(s,a^1,a^2) = -r^2(s,a^1,a^2)$. For two-player zero-sum games with altering turns, the existence of an optimal deterministic policy is ensured (Littman, 1994).

Reinforcement learning methods can be divided into *model-based* approaches that utilize a model of the environment (transition and reward function) in combination with planning algorithms and *model-free* approaches that optimize either the policy function directly (policy-based methods) or a value function (value-based methods) from which the policy can be derived. In the scope of this thesis, model-free approaches are used and the specific algorithms are discussed in detail in the following subsections.

Although algorithms for simultaneous training of all agents were introduced and explored (Bowling and Veloso, 2001), single-agent learning can be utilized in markov games as well. Since the herein pursued approach utilizes single-agent learning in the markov game setup, the following discussion of methods is also focused on the single-agent case.

### Value-based Methods

Value-based methods aim at approximating the optimal state-action function $Q : S \times A \to \mathbb{R}$. It encodes the expected future reward given action $a$ in state $s$ and following policy $\pi$ afterwards. For the single-agent case it can be described in terms of the state value function in Equation 2.26 as

$$Q_\pi(s,a) = r(s,a) + \gamma \sum_{s'} T(s,a,s') V_\pi(s'). \tag{2.27}$$

If the optimal Q-function is known, the optimal policy can be derived by always taking action $a$ with the maximal Q value:

$$\pi^*(s) = \operatorname*{argmax}_a Q(s,a). \tag{2.28}$$

The Q-learning algorithm (Watkins and Dayan, 1992) approximates the Q-function during training by updating it according to

$$Q(s_t,a_t) = (1-\alpha)Q(s_t,a_t) + \alpha(r_t + \gamma \max_a Q(s_{t+1},a)), \tag{2.29}$$

where $s'$ denotes the state the agent is in after executing action $a$ and $\alpha$ is called the learning rate. Due to the max operator over actions, this is an *off-policy* algorithm. The update rule for the respective *on-policy* variation reads as

$$Q(s_t,a_t) = (1-\alpha)Q(s_t,a_t) + \alpha(r_t + \gamma Q(s_{t+1},a_{t+1})), \tag{2.30}$$

with $a_{t+1}$ the action chosen according to the current policy in state $s_{t+1}$. It is usually referred to as SARSA - state action reward state action - algorithm (Sutton and Barto, 1998). It can be seen that both, the Q-learning and the SARSA algorithm consider only the current state and the immediate predecessor in their update, which is referred to as *one-step* Q-learning/SARSA. This can be extended through the combination of the respective algorithms with *eligibility traces* to enable a more efficient learning. The *eligibility trace* $\psi_t(s,a)$ encodes a (decaying) memory of previously visited state-action pairs and is updated after each time step. Throughout this thesis, the *replacing traces* update is used although different approaches are possible (Sutton and Barto, 1998). The corresponding updates of the eligibility trace and the Q-function are discussed in the following for both Q-learning and SARSA and the respective algorithms are denoted as $Q(\lambda)$ and SARSA($\lambda$).

In both cases, the eligibility trace is initialized as $\psi(s,a) = 0 \; \forall s,a$, as no state-action pair has been visited yet. For SARSA, the trace update at time step $t$ is then given as

$$\psi_t(s,a) = \gamma\lambda\,\psi_{t-1}(s,a) \qquad \forall s,a \neq s_t,a_t \tag{2.31}$$

for all state-action pairs that are not the currently visited one. Here, $\gamma$ is again the discount factor and $\lambda \in [0,1]$ the trace-decay parameter (Sutton and Barto, 1998). For the current state-action pair, the *replacing traces* approach yields the update

$$\psi_t(s_t,a_t) = 1\,, \tag{2.32}$$

meaning that even if a state-action pair is visited repeatedly, the corresponding trace value has an upper bound of 1. The respective update for the Q-function is then

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha\delta_t\,\psi_t(s,a) \qquad \forall s,a\,, \tag{2.33}$$

with

$$\delta_t = r_t + \gamma Q(s_{t+1},a_{t+1}) - Q(s_t,a_t). \tag{2.34}$$

It can be seen that for the case of $\lambda = 0$, this corresponds to the *one-step* SARSA algorithm discussed above.

For $Q(\lambda)$, the trace update is similar as in SARSA, with the difference that the trace values are set to zero for all state-action pairs that are not the currently visited one whenever a non-greedy action is taken:

$$\psi_t(s,a) = \begin{cases} \gamma\lambda\,\psi_{t-1}(s,a) & if\ Q_{t-1}(s_t,a_t) = \max_a Q_{t-1}(s_t,a) \\ 0 & else \end{cases} \qquad \forall s,a \neq s_t,a_t\,. \tag{2.35}$$

Again, the trace for the current state-action pair is set to 1 in any case. The update of the Q-function is then analogous to Equation 2.33 with

$$\delta_t = r_t + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t). \tag{2.36}$$

Like in the case of SARSA($\lambda$), the case of $\lambda = 0$ corresponds to the one-step Q-learning discussed earlier.

In a straightforward approach to Q-learning and SARSA, each state-action value is stored separately in a lookup table. However, for large state spaces, this requires an impractical amount of training episodes and computational power which is usually referred to as Bellmans' curse of dimensionality (Keogh and Mueen, 2011). To bypass this issue, different function approximations for the Q-function were employed, ranging from linear approximations (Sutton and Barto, 1998) to deep neural networks (Mnih et al., 2013) and stochastic processes (Engel et al., 2005;

**Algorithm 1:** SARSA($\lambda$) with linear function approximation (Sutton and Barto, 1998)

Init: $\omega = (0,...,0)^\top$, $\alpha =$ learning rate
**foreach** *episode* **do**
  $\psi \leftarrow (0,...,0)^\top$
  $s_0, a_0 \leftarrow$ initial state and action
  $F_\phi \leftarrow$ non-zero feature indices for $s_0, a_0$
  **foreach** *step t in episode* **do**
    **foreach** $i \in F_\phi$ **do**
      $\psi_i \leftarrow 1$
    take action $a_t$, get reward $r_t$ and transition to state $s_{t+1}$
    $\delta \leftarrow r_t + \sum_{i \in F_\phi} \omega_i$
    **if** $s_{t+1}$ *is final* **then**
      $\omega \leftarrow \omega + \alpha\delta\psi$
      terminate episode
    **foreach** $a \in A(s_{t+1})$ **do**
      $F_\phi \leftarrow$ non-zero feature indices for $s_{t+1}, a$
      $Q_a \leftarrow \sum_{i \in F_\phi} \omega_i$
    get $a_{t+1}$ from policy $\pi$ ($\varepsilon$-greedy)
    $\delta \leftarrow \delta + \gamma Q_{a_{t+1}}$
    $\omega \leftarrow \omega + \alpha\delta\psi$
    $\psi \leftarrow \gamma\lambda\psi$

Rasmussen and Williams, 2006). The herein utilized value-based methods use a linear function approximation of the form

$$Q(s,a) = \sum_i \omega_i \phi_i(s,a), \tag{2.37}$$

where $\omega_i$ denote weights that are optimized during learning and $\phi_i$ features characterizing the present state-action pairs. As similar state-action pairs have similar features, this approximation offers an educated guess for the Q-function value of states that were not visited during training based on the values for features this state shares with others. Learning is then done by updating the weight parameters. In the scope of this thesis, linear function approximation with binary features $\phi$ as well as eligibility traces is utilized. Consequently, the dimension of the weight vector $\omega$ equals the number of features in $\phi$. As the eligibility trace is also a function of state-action pairs, it is as well expressed as a vector $\psi$ with the same dimension as $\omega$. The corresponding learning for SARSA($\lambda$) is shown in Algorithm 1. For Q($\lambda$), the combination of eligibility traces with linear function approximation can be derived in an analogous manner by taking Equation 2.35 into account (Sutton and Barto, 1998).

### Experience Replay

Before discussing more complex algorithms, some notations are introduced and a concept to facilitate optimization in the case of more complex learning problems is discussed, namely importance-weighted experience replay. For the remainder of the section, the general case of a probabilistic policy $\pi : A \times S \to [0,1]$ is considered. For a given policy $\pi$, the probability of reaching state $s$ after starting in state $s_0$ and following policy $\pi$ is denoted as

$$d^{\pi}(s) = \lim_{t\to\infty} \mathbb{P}(s_t = s \mid s_0, \pi). \tag{2.38}$$

In addition the notation $\mathbb{E}_{s\sim d^{\pi}, a\sim\pi}[\cdot] = \mathbb{E}_{\pi}[\cdot]$ for the expected value is used.

An approach to speed up the learning process in cases of complex function approximations (like neural networks) is to train the agent off-policy with samples collected from past iterations. This is referred to as experience replay (Lin, 1992) and a common approach in this regard is *importance sampling*. Given a current policy $\pi(a_t|s_t)$ and a behaviour policy $\mu(a_t|s_t)$ under which the replay sample was generated, the importance weight of the sample is

$$\rho_t = \frac{\pi(a_t|s_t)}{\mu(a_t|s_t)}. \tag{2.39}$$

For a (partial) trajectory from $l$ to $t+1$ $\{s_l, a_l^{\mu}, r_l, ..., s_t, a_t^{\mu}, r_t, s_{t+1}\}$ generated under the behaviour policy $\mu$ , the importance weight is computed as (Mahmood and Sutton, 2015)

$$\rho_l^{t+1} = \prod_{i=l}^{t} \frac{\pi(a_i|s_i)}{\mu(a_i|s_i)} = \prod_{i=l}^{t} \rho_i. \tag{2.40}$$

It can be seen from the above equation, that the importance weight is sensitive towards high ratios due to the product. To address this issue, a truncation of the importance weight were proposed, e.g. $\tilde{\rho}_t = \min(1, \rho_t)$.

### Actor-Critic Methods

Actor-critic methods (Konda and Tsitsiklis, 2000) combine aspects of policy-based methods and value-based methods to get the best out of both approaches. To this end, a policy function $\pi(a|s)$ (actor) is updated based on a value-based feedback (critic). The policy function and the value or Q-function can be approximated by different techniques, in the present case through neural networks with parameters $\theta$ (policy) and $\omega$ (critic). The critic network parameters are updated to minimize the loss function

$$L(\omega) = \mathbb{E}_{\pi}[(Q_{\pi}(s,a) - Q_{\omega}(s,a))^2], \tag{2.41}$$

where $Q_{\pi}$ denotes the Q-function under the current policy $\pi$ that has to be estimated with observed rewards. As for the actor, the policy network is updated with respect to an objective function $J(\theta)$ for which the gradient can be computed using the policy gradient theorem (Sutton et al., 1999)

$$\nabla_{\theta} J(\theta) = \sum_s d^{\pi_\theta}(s) \sum_a \nabla_{\theta} \pi_{\theta}(a|s) Q_{\omega}(s,a) \tag{2.42}$$

$$= \mathbb{E}_{\pi_\theta}[\nabla_{\theta} \log \pi_{\theta}(a|s) Q_{\omega}(s,a)]. \tag{2.43}$$

The herein employed Actor-Critic with Experience Replay (ACER) algorithm for discrete action spaces (Wang et al., 2016) utilizes off-policy training with importance-weighted experience replay (see previous section). Using the policy gradient theorem and following the approach discussed in (Degris et al., 2012), the following approximation of the gradient can be derived

$$g_{marginal} = \mathbb{E}_{\mu}[\rho_t \nabla_\theta \log \pi_\theta(a_t|s_t) Q_\pi(s_t, a_t)], \tag{2.44}$$

with a behaviour policy $\mu$ and importance weight $\rho_t$ as discussed above. For estimating the action value function $Q_\pi$ in Equation 2.44, ACER utilizes the Retrace algorithm (Munos et al., 2016). It is a return based method that can be used to estimate $Q_\pi$ for a given policy from trajectories generated under a behaviour policy $\mu$. The corresponding estimator $Q^{retrace}$ is given (Wang et al., 2016) as

$$Q^{retrace}(s_t, a_t) = r_t + \gamma\tilde{\rho}_{t+1}(Q^{retrace}(s_{t+1}, a_{t+1}) - Q(s_{t+1}, a_{t+1})) + \gamma V(s_{t+1}), \tag{2.45}$$

where $Q(s,a)$ denotes a value estimate of $Q_\pi(s,a)$ and $V(s) = \mathbb{E}_{a\sim\pi}[Q(s,a)]$. In addition, $\tilde{\rho}_t = \min\{c, \rho_t\}$ denotes the importance weight truncated with a constant $c$ as discussed in the previous section. In the present case, the estimate $Q(a,s)$ is taken from the critic network with parameters $\omega$, yielding the recursive update for the Retrace estimator

$$Q^{retrace}(s_t, a_t) = r_t + \gamma\tilde{\rho}_{t+1}(Q^{retrace}(s_{t+1}, a_{t+1}) - Q_\omega(s_{t+1}, a_{t+1})) + \gamma V_\omega(s_{t+1}), \tag{2.46}$$

with $V_\omega(s) = \mathbb{E}_{a\sim\pi_\theta}[Q_\omega(a,s)]$. In addition, ACER utilizes the Retrace estimator for training the critic network. The respective gradient of the loss function utilized to update the network weights is then given as

$$\nabla_\omega L(\omega) = (Q^{retrace}(s_t, a_t) - Q_\omega(s_t, a_t))\nabla_\omega Q_\omega(s_t, a_t). \tag{2.47}$$

Although the truncated importance weight is used in the retrace update, the importance weight in the marginal gradient (Equation 2.44) is still unbound and can hence lead to instability in the learning process. As the truncation can in turn introduce bias to the learning, the authors in (Wang et al., 2016) propose an importance weight truncation with bias correction. To this end, a correction term is added to the gradient to ensure an unbiased estimate. The final ACER gradient approximated with a trajectory sampled from behaviour policy $\mu$ is then given as

$$\begin{aligned} g_{acer} = {} & \tilde{\rho}_t \nabla_\theta \log \pi_\theta(a_t|s_t)(Q^{retrace}(s_t, a_t) - V_\omega(s_t)) + \\ & \mathbb{E}_{a\sim\pi}\left[\max\left(0, \frac{\rho_t(a) - c}{\rho_t(a)}\right) \nabla_\theta \log \pi_\theta(a|s_t)(Q_\omega(s_t, a) - V_\omega(s_t))\right], \end{aligned} \tag{2.48}$$

with $\rho_t(a) = \frac{\pi(a|s_t)}{\mu(a|s_t)}$. The subtraction of $V_\omega(s_t)$ is included to further reduce variance and it can be seen that the bias correction term is only active in the case $\rho_t(a) > c$.

To prevent variance in the policy update while at the same time maintaining an efficient learning speed, ACER utilizes trust region optimization (Schulman et al., 2015) with additional modifications: As limiting the updated policy to be close to the current one is computationally inefficient for complex problems, the updated policy is restricted to be close to an average policy $\tilde{\pi}$ that encodes an average over previously learned policy parameters (Wang et al., 2016).

## 2.4 Summary

Throughout this chapter, the relevant background for the present thesis was introduced. As it covers techniques and formal frameworks from multiple fields, a concise summary with respect to the relevance for the thesis contributions is provided in the following:

For dialogue system technologies, an overview of different system types and architectures was discussed. It provides the necessary context for the herein introduced modules and system implementations. In addition, the motivation for the preliminary choices of the utilized models in Chapter 4 builds upon this discussion. Subsequently, different evaluation approaches for dialogue systems were summarized. They are primarily used in Chapter 6 for the development of meaningful evaluation setups and categories.

The second part of the chapter addressed methods and models from the field of computational argumentation, including inference-based argumentation, dialogue-based argumentation, argument quality and argument mining. The discussed inference-based models are again relevant for the preliminary choices in Chapter 4 but also for the mapping of argument search results into argument structures in Section 6.2, the development of the preference model in Section 7.1 and the discussion of related work in the following chapter. The background on dialogue-based argumentation is also relevant for the preliminary choices in Chapter 4 and provides the theoretical foundation for Chapter 5. Argument quality is, similar to the evaluation of dialogue systems, required for the evaluation setups in Chapter 6 and the definition of a reasonable winning criterion in the context of the investigated dialogue game in Section 5.1.2. The overview of argument mining provides the background for the techniques and technologies utilized in Chapter 6 to enable topic flexibility for argumentative dialogue systems.

The last part of the chapter introduced different machine learning techniques from the subfields of supervised learning and reinforcement learning. The selection is focused on methods that are applied in the context of the present work and all following contributing chapters except Chapter 4 utilize one or more of the introduced methods.

# 3 Related Work

This chapter summarizes related work from the different fields considered within the thesis and complements the previous chapter with alternative or preceding approaches to the herein addressed tasks. In the first section, relevant contributions to the development of argumentative dialogue systems will be reviewed. The focus of the discussion will be on the underlying dialogue model, the utilized argument representation as well as the acquisition of arguments. In addition, the development of the system strategy and the corresponding evaluation setup will be discussed. Approaches that investigate isolated aspects or modules of an argumentative system, like strategy optimization or user adaptation will also be included if they bear relevance and/or similarity to the present work. Subsequently, recent work on argument quality assessment is recalled with an emphasis on approaches that either build on the taxonomy discussed in Section 2.2.3 or pursue an approach that is not considered therein. The discussion covers the methods to evaluate and/or annotate argument quality as well as approaches to automatize this assessment. The third section covers works from the field of affective computing with relevance for the present thesis. The focus of the discussion is on approaches that estimate human *agreement* and *interest* from multimodal data. At the end of the chapter, a brief summary with respect to the herein introduced contributions is provided. As the discussion of argument quality and affective computing are mainly relevant for the proposed estimation of subjective argument quality aspects from non-verbal cues shown by the user (Chapter 7), the respective discussion of related work is an extension of the one published in (Rach et al., 2021a).

## 3.1 Argumentative Dialogue Systems

Computational argumentation as a whole and argumentative dialogue systems in particular have gained a lot of attention recently. One early implementation of a system for human-computer argumentation was introduced in (Bench-Capon, 1998). It models the interaction through a dialogue game based on a specified version of the Toulmin model (Toulmin, 2003) discussed in Section 2.2.1. Based on the formal definition of the game moves, the strategy of the machine player is defined in a set of rules. The implementation also includes an interface with a list of available game moves from which users can select their next move. However, an evaluation of the system with human users was not explicitly discussed in the referenced work. In addition, the automatic acquisition of arguments, strategy optimization and user adaptation as considered herein were not addressed.

The system introduced in (Yuan et al., 2008) utilizes a specification of the dialogue game DC (Mackenzie, 1979) called DE to model the dialogue in the addressed domain of educational debates. The knowledge base utilizes a structured representation of propositions and counter-propositions. Each proposition can have multiple supporting propositions and a single counter-proposition. In addition, it distinguishes between evidence and opinion. However, the acquisition of arguments for the knowledge base is not explicitly addressed. The system strategy is based on extensively discussed heuristics and the interaction with human users is approached through an interface that enables the corresponding user to select the type of move and the content of the move separately. The system was assessed in an expert evaluation (Yuan, 2004) and a user study (Ævarsson, 2006), both confirming the functionality of the system. In addition, the expert evaluation indicated two drawbacks of the system strategy, namely that the system repeatedly hands over its turn and that the system does not voluntarily concede. The user study was conducted with ten students, of which eight ended the debate in a draw and two conceded to the system stance. However, the automatic acquisition of arguments, a machine learning-based optimization of the system strategy and adaptation to human users as included in the present thesis were not included.

In (Rosenfeld and Kraus, 2016), the authors proposed a persuasive agent that models the argumentation using the *bipolar weighted argumentation framework* discussed in Section 2.2.1. In contrast to the previously reviewed works, where different types of moves in the utilized argument games were considered, the interaction is in this instance limited on the exchange of arguments, i.e. each utterance includes an argument that responds to a previous one. The corresponding arguments for the discussed topics were acquired in a data collection from human-human discussions and subsequently assigned the weights required in the framework based on human ratings. The strategy of the system is optimized through a formalization as a partially observable markov decision process and the use of Monte-Carlo planning with samples from the collected argument corpus. To enable users to interact with the proposed agent via natural language, a human expert assigned the user utterances to arguments in the corpus during the dialogue. The system was evaluated through user studies with an attitude change topic and a behaviour change topic. For both scenarios, the system was compared to the human performance observed in the data collection and a baseline. The results indicate a performance that is similar to the human one and significantly better than the baseline in both cases. Although the system utilizes an optimized strategy, it depends directly on the collected topic data. In contrast, the herein proposed approach to policy optimization does not require a training corpus. In addition, the proposed automatic acquisition of arguments and an adaptation to individual users were not included.

Higashinaka et al. (2017) proposed a complete argumentative system that is capable of handling natural language input by mapping the user input into the four utilized speech acts *assertion*, *question*, *concession* and *retraction*. Subsequently, the cosine distance between the user utterance and arguments encoded in the knowledge base is utilized to identify the corresponding node in the argument structure and to identify out-of-domain utterances. The corresponding argument representation is based on a claim-premise model with the two relations support/non-support and five different domains/topics were included (in separate structures) in the introduced version. As for the dialogue strategy, no specific approach is discussed apart from the fact that the system retrieves premises that are related to the current user argument, hence indicating an immediate response protocol. In addition, the system utilizes a chat module to address out-of-domain utterances. However, an evaluation of the complete system and/or the individual modules was not included in the referenced work. As for the utilized argument structures, a detailed discussion of their generation was presented in a separate work (Sakai et al., 2018b). They were generated by humans and assessed through the generation of artificial dialogues that were evaluated in a user study. The results indicate that the artificial dialogues are reasonable and understandable and the suitability of the underlying structures for the use in dialogue systems is concluded by the authors. In addition, the structures were utilized as training data for a machine learning-based estimate of the relation (support/non-support) between two statements. In follow-up work, the collected argument structures were used to generate argumentative dialogues that consider the user opinion (Sakai et al., 2020) expressed in the form of agree/disagree statements during the interaction. The system architecture is similar to the one discussed above, extended by a rule-based dialogue strategy that reacts to the opinion expressed by the user. For evaluation, dialogues generated with the new strategy were compared to dialogues generated without taking the user agreement into account. The results indicate that the adaptive model was perceived as more persuasive than the compared baseline. Again, policy optimization and an automatic acquisition of the required argument structures as proposed in this thesis were not included.

Increasing the persuasiveness of a system through strategy adaptation to user opinion was also addressed in (Chalaguine and Hunter, 2020) through the use of user concerns. Building on the concepts introduced in (Chalaguine et al., 2019; Hadoux and Hunter, 2019), the system utilizes an argument graph based on the abstract argumentation framework (Dung, 1995) discussed in Section 2.2.1. Consequently, only attack relations between the arguments are considered. The corresponding argument graph was collected in a crowd-sourcing setup (Chalaguine and Hunter, 2019). The dialogue is focused on the exchange of arguments and immediate responses, i.e. the system responds directly to the previous user argument if a suitable response argument is present in the utilized graph. The user on the other hand is not forced to reply with an argument that is directly related to the previous system utterance. The system allows for (written) natural language input which is mapped through a semantic similarity measure to an argument in the graph. To recognize and adapt to the user concerns, a topic-specific list of five concerns was defined and a prediction model was trained to estimate the user concern from the utterance text. In terms of strategy, the proposed system then utilizes the recognized concern to select the next argument with the same concern (if present). For evaluation, the system was compared in a user study to a baseline that did not take user concerns into account. The results show that the consideration of concerns leads to a significantly higher change in the user opinion. Although the system includes an approach to user adaptation, an optimization of the system strategy through machine learning as well as an automatic acquisition of arguments was again not included.

The use of semantic similarity measures to retrieve responses was also investigated in the case of argumentative chatbots. The system introduced in (Rakshit et al., 2019) retrieves its responses from an argument mining corpus for three different topics. In contrast to the previously discussed works, relations between the arguments are not considered as the corpus only includes annotations of the stance, topic and response characterization. In addition, a regression-based argument quality score is provided. The authors then cluster the arguments in the corpus thematically. For a given user utterance, the system either compares the utterance to all arguments in the corpus (baseline) or the arguments in the corresponding cluster, also using a semantic similarity measure. The most similar argument with an opposing stance is then chosen as the system response. In doing so, an implicit immediate-response protocol is used as the system is not able to backtrack to previous utterances. In particular, an optimization of the strategy and adaptation to human users as proposed in the scope of this thesis are not included in the system. The authors evaluated their approach through a comparison of the time needed to retrieve a response and found that (as expected), the clustering-based retrieval performs better. A similar approach was compared by Le et al. (2018) to a generative model that was trained on a corpus of debate posts on various topics (Abbott et al., 2016). In contrast to the retrieval-based approach, the generative one is not strictly limited to specific topics or domains. In addition, the generative model is not strictly immediate-response, as it encodes the dialogue history and is hence, in principle, able to address previous utterances. In turn, however, the response is solely based on the utilized data and aspects like additional strategy optimization or user adaptation as proposed in this thesis are not considered, yet. The authors verified the functionality of their approaches through the use of established metrics for the underlying technological tasks. For the retrieval-based approach, the assessment of similarity in a question pair corpus was used whereas the generative approach was evaluated in terms of model perplexity and distinctness on an argumentative test set.

The arguably most prominent argumentative system is the IBM project debater (Slonim et al., 2021) which engages in a full-scale debate with a human interlocutor. Due to the specific setup that enforces strict rules regarding turn-taking and speaking time, strategic aspects in this regard are not considered. Instead, the focus is on generating argumentative speech that meets the requirements of the setting. To this end, the system includes four different modules, namely an argument mining module, a knowledge base of arguments, a rebuttal module and a debate construction module. The argument mining module performs argument search on a large newspaper corpus, including offline processing of documents and indexing of sentences as well as and online retrieval of claims and evidence. The knowledge base of the system complements the arguments retrieved by the argument mining module with human-generated, general arguments applicable in multiple domains. The rebuttal module predicts possible arguments that might be used by the opposition, processes the opponent's speech using IBM Watson and aligns the predictions with the actual responses. In addition, it identifies rebutting arguments which are then used in combination with clustering techniques and a speech template by the debate construction module to derive the final system response. Amongst other evaluations, the system was compared against several baselines on the task of generating an opening speech. It was shown to outperform all baseline systems and yielded a performance close to that of human experts. Although the system utilizes state-of-the-art argument mining approaches to ensure topic flexibility, the corresponding modules are tailored to the domain of debates and not publicly available for use in other systems, yet. In addition, the clear structure of a debate imposes strict rules on the interaction. As a consequence, the system strategy is focused solely on the selection of suitable arguments and their compositions in a debate speech. In particular, certain aspects of argumentative dialogue like the length of an utterance as part of the strategy as well as an adaptation to individual users as considered herein are not addressed.

So far, the discussion was focused on fully implemented argumentative systems. The remainder of this section will discuss works that address isolated aspects of human-computer argumentation and are relevant for this thesis. For the development of argument strategies, Hadoux et al. (2015) proposed a formalization of *argumentation problems with probabilistic strategies* (Hunter, 2014) as a mixed observable markov decision process to optimize the argumentation strategy of an agent against a probabilistic opponent strategy. The results indicate that the proposed approach performs well in scenarios with small amounts of arguments but is not yet feasible for larger problems. The approach that is closest to the ones investigated in the scope of this thesis was introduced in (Alahmari et al., 2019). The authors propose the use of reinforcement learning to optimize an agent policy in the dialogue game DE. To this end, the agent is trained against pre-defined opponent strategies to maximize its overall reward. Although the approach is, as mentioned before, similar to the one pursued herein, it relies on pre-defined strategies and hence, on knowledge about the behaviour of the opponent.

## 3.2 Argument Quality

In this section, related work that is concerned with the assessment of argument quality is reviewed. As discussed in Section 2.2.3, two different approaches to argument quality assessment are distinguished in recent literature. They are referred to as theory-based approaches and comparative or rating-based approaches. Theory-based argument quality assessment builds upon theoretical definitions of quality dimensions to derive annotation guidelines. In contrast, comparative approaches are motivated by the subjective nature of the task and assess the quality of arguments in pairwise comparison to others based on (intuitive) human judgements.

The unified taxonomy introduced by Wachsmuth et al. (2017b) discussed in Section 2.2.3 has become a frequent reference work for theory-based approaches. The authors in (Lauscher et al., 2020) utilized the therein defined argument quality dimensions *cogency*, *reasonableness* and *effectiveness* to build an annotated argument quality corpus from forum posts in three different domains, namely, question-answering, debate portals and business reviews. The fine-grained subdimensions in the original work were not explicitly annotated but utilized as guiding questions for the annotators. The authors observed varying inter-annotator agreements for the different domains and highlight the subjective nature of the task. In addition, methods to automatically estimate the annotated scores from different feature sets and with different models were explored and the synergies between the utilized theoretical approach and a practical approach to argument quality assessment were investigated. The authors conclude from the results that both approaches can benefit from each other in the sense that the theory-based models can assist in interpreting the practical approach and in turn, that the practical approach can be utilized to improve the performance of the theoretical ones.

Also along the line of theory-based assessment, Potthast et al. (2019) utilized expert ratings in the categories *logical*, *rhetorical* and *dialectical* quality to compare different retrieval approaches for argument search in combination with the information retrieval notion of relevance. Four retrieval models were compared over 20 topics for each of which a neutral topic formulation and a biased topic formulation was generated. Each evaluation category was rated on a four-point Likert scale and a different annotator was recruited for each of the 40 investigated instances. In addition, the effect of personal bias as a result of the opinion of the annotator was investigated. The authors found no statistically significant effect of the personal opinion on the ratings although the subjective nature of the overall task is emphasized.

A rating-based approach to argument quality assessment utilizing crowd-sourcing annotation was introduced in (Habernal and Gurevych, 2016b). It was motivated by the subjective nature of the task and the resulting personal bias in direct ratings of individual arguments. In the proposed approach, the crowd workers were instead asked to rate a pair of arguments with respect to their convincingness, i.e. which of the two arguments they perceived as more convincing. In addition, the raters were asked to provide a written explanation for their decision. Properties of the annotated convincingness relations were then investigated and the annotated corpus was published. In addition, the authors trained machine learning models to predict the relation between pairs of arguments and to perform argument ranking. In a subsequent work (Habernal and Gurevych, 2016a), an additional annotation scheme was introduced to infer argument quality labels from the written explanations provided by the annotators in the original crowd-sourcing experiment.

Also along the line of comparative argument quality assessment, a corpus for the pairwise comparison of the convincingness of evidence was introduced in (Gleize et al., 2019). Like in the previously discussed work, the annotation was done in a crowd-sourcing setup. Participants were asked to rate which of the two presented evidence sentences they would use in a discussion about the given topic if they had to choose one.

The correlations between the theoretical and the comparative crowd-sourcing approach were also investigated (Wachsmuth et al., 2017a). The authors found that a lot of reasons provided for the annotations in the crowd-sourcing setup could be matched to theoretical quality dimensions and that the comparative annotations correlate with theory-based absolute ratings. In addition, the authors discuss the strengths and weaknesses of both approaches, concluding that the reasons provided for a rating in the comparative approach are in some instances less informative than theory-based ratings whereas theory-based assessment is, in turn, difficult in practice for some of the utilized categories. The overall quality of single arguments, as well as argument pairs, was discussed by Toledo et al. (2019) together with automatized approaches for argument ranking and argument-pair classification. A comparison of the single argument and the argument pair annotation showed that the results of both approaches are mostly consistent and the authors suggested using pair-wise approaches mainly for argument pairs with a low difference in the individual rating. In addition, a large corpus with binary annotations regarding the quality of arguments and their stance was introduced and methods to derive argument quality scores from the labels were investigated alongside a prediction model for argument quality (Gretz et al., 2020).

Despite the remarkable results of these works, they all focus on properties of the arguments to determine their quality. In contrast, the herein considered approach investigates the subjective user perception of arguments and hence their subjective quality.

## 3.3 Affective Computing

To enable an automatic assessment of the subjective user opinion regarding arguments, indicating features have to be extracted from the reaction of the individual user (instead of the arguments, as in the previously discussed argument quality approaches). Therefore, methods from the area of affective computing are required. This field has yielded several approaches that bear similarity to the herein considered tasks of estimating whether an argument is *interesting* and *convincing* for a user. In particular, the automatic recognition of (dis)agreement is closely related to the herein addressed *convincing* task. Following the notion of Poggi et al. (2011), for two persons A and B, *"[...] B agrees with A when B assumes that his/her opinion is the same, similar or in any case congruent (in the same line, not conflicting)"*. Consequently, it can be said that if B is convinced by an argument of A, he/she also agrees with A (on the content of the argument). On the other hand, agreement does not necessarily implicate that the agreeing person is convinced. For example, A and B can agree on a premise but draw different conclusions from it, therefore establishing contradicting opinions and arguments. In addition, social signals associated with agreement can also be shown out of politeness or as a sign of attention. This is referred to as apparent agreement (Poggi et al., 2011). Consequently, it can be said that if a person A is convinced, he or she is likely to show signs of agreement but signs of agreement do not necessarily implicate a convincing argument.

From a technical perspective, a variety of different approaches to recognizing (dis)agreement were introduced and overviews are provided for example in (Bousmalis et al., 2009; Bousmalis et al., 2013). The authors distinguish three types of (dis)agreement – direct speakers (dis)agreement, indirect speakers (dis)agreement and nonverbal listeners (dis)agreement – and discuss the corresponding indicating cues. Besides verbal indicators gestures, postures, facial expressions and eye movements are listed as frequent cues. Regarding the herein considered task, the non-verbal cues are of more interest as the participants assumed a passive and quiet role. An approach to (dis)agreement recognition that utilizes only non-verbal cues was presented in (Bousmalis et al., 2011) and serves as a reference point for the comparison of convincingness recognition and (dis)agreement recognition.

For the second affective computing task considered throughout this thesis, namely the recognition of the participants' interest in the presented arguments, multiple related approaches were introduced. An approach based on emotion recognition that estimates the level of interest from facial expressions was presented in (Yeasin et al., 2006). To this end, the authors map the recognized facial expressions into a three-dimensional affect space (valence, arousal, stance) and subsequently, infer the level of interest from it. Moreover, Schuller et al. (2007) recorded a multimodal corpus of human-human conversation and annotated five levels of interest. Afterwards, different features including facial expressions, eye movements as well as linguistic and acoustic features were merged in an (early) fusion and utilized to estimate the annotated labels. The authors found that from the individual modalities, audio-related features perform best but also that the result can be improved by combining audio-related features with ones related to activity derived from eye movement. The collected data was also used in (Jeon et al., 2014), where an estimation of utterance-wise interest levels in a dialogue system setup was investigated. The proposed approach utilizes acoustic and lexical features extracted from the user utterance to build two separate regression models and subsequently combines their predictions in a late fusion. The results show that this fusion leads to better results than predictions with one of the two models only. In (Sasaka et al., 2016), the user interest in movie trailers was assessed on a binary scale. The interest label was inferred from the rating provided by the participants of the conducted study after watching the corresponding trailer. The authors then proposed a method to estimate the interest labels from facial expressions as well as biological signals. The results indicate a reliable performance of the proposed method in comparison to several competitive approaches, whereas the role of the different features was not explicitly discussed. Hirayama et al. (2010) estimated user interest in displayed content from eye movements. To this end, indicating eye-movement cues were defined and the interest of participants was assessed in an interview subsequently to the conducted experiment on a five-point scale. The authors conclude from their results that the proposed approach is feasible but also discuss the use of personality types to improve performance. For the case of multimodal human-machine dialogue, facial expressions alongside prosodic features related to the user's voice were used to determine the turn-wise interest of the human user in the discussed topic (Tomimasu and Araki, 2016). To this end, the level of interest was annotated on the exchange level in three categories (interested, neutral, not interested) which were then merged into binary class labels. The results indicate again that the performance is increased by combining both feature types. Due to the similarity of the setup to the herein discussed one, this work is chosen as a reference point for the comparison of the recognition of interest in arguments with related literature. Finally, a cognitive model of user interest in different objects for the application in smart environments was introduced in (Ahmed et al., 2020). To this end, a formal model to map activity to interest was proposed. The approach was validated in an experiment where the user activity was estimated from statistical user data on an internet forum.

Despite the similarities between these setups and the herein considered tasks, the individual interest and the perceived convincingness of an argument were not considered in the referenced work.

## 3.4 Summary

The overview of related work indicates several challenges in the development of argumentative systems. In the following, a summary of the pending issues addressed in the scope of this thesis is provided. It is divided into the three sub-tasks addressed in the scope of the thesis, namely argument strategies, topic flexibility and user adaptation.

As for the development of argument strategies, most of the reviewed argumentative systems rely on robust, yet inflexible rules. In cases where strategy optimization is considered, the corresponding approaches either rely on a corpus with training data, pre-defined opponent strategies and/or indicate scaling problems regarding the number of arguments that can be discussed. In addition, most of the proposed approaches are investigated in the specific setup of the corresponding system. The herein pursued approach based on multi-agent reinforcement learning addresses all these issues through scalable training in self-play. In addition, the proposed general reformulation of dialogue games for argumentation as markov games allows for an application of this method in multiple scenarios and setups.

The topics that can be discussed by the individual systems are in most of the discussed cases limited by the reliance on a previously collected database or argument structure. The only exceptions in this regard are the two argumentative chatbots (Le et al., 2018; Rakshit et al., 2019) and the IBM project debater (Slonim et al., 2021). In the case of the argumentative chatbots, this flexibility comes at the price of a more restricted dialogue setup that does not allow for strategy optimization and user adaptation whereas the technology of the IBM project debater is tailored to the specific application domain of debates. The herein proposed use of argument search technology in combination with relation classification bridges the gap between data-based and model-based systems as it allows for the automatic acquisition of argument structures required in systems that rely on formal models in their functionality. It is flexible in the sense that it enables the retrieval of argument structures for any topic on which the search engine can find suitable arguments.

Regarding user adaptation, the reviewed approaches in argumentative dialogue systems focus mostly on pre-defined concerns that are derived either from the utterance of the user, assessed in a pre-study or derived from personality traits. In contrast, the herein proposed approach adapts to the individual user perspective of arguments, i.e. the subjective quality of an argument. To this end, argument quality is not assessed from the argument itself but from non-verbal reactions shown by the user during its presentation through a virtual avatar. It can be seen from the discussion of argument quality assessment and affective computing approaches that the two fields are, up to now, separated. In contrast, the herein proposed automatic assessment of subjective argument quality aspects from user reactions combines approaches from both areas.

Moreover, it can be seen that the majority of the investigated approaches consider argumentation as a purely dialogical task and that multimodal systems that combine their argumentative utterances with synthetic emotions are mostly unexplored. In addition, the majority of the systems focus on direct human-machine interaction and the use of multi-agent systems as a way to investigate and evaluate models as well as isolated aspects of argumentation is also still in its infancy.

# 4 Architectural Design and Pilot Study

This chapter covers the preliminaries required for the contributions made in later chapters to solve the three tasks of challenging agent strategies, topic flexibility and user adaptation. In particular, existing frameworks that are utilized as a starting point, reference and for evaluation are introduced. A reasonable choice in this regard depends on the specific role the corresponding approaches assume in an actual application. Consequently, a general architectural design including the positioning of the contributions made in this thesis is proposed in the first part of the chapter. Based on this general setup, a selection of suitable formal frameworks is discussed and motivated. This includes a specific dialogue model to structure the argumentative interaction as well as a formal representation of arguments and their relations. For both cases, specific instantiations are implemented and discussed.

In the second part of the chapter, an agent-agent system employing the selected models is introduced. It is used to validate the previous choices and identify pending issues. Moreover, it provides the basis for testing, evaluation and system implementations in the remainder of this thesis. Following the classification of Reed and Norman (2003), the focus is on *persuasion*, also referred to *pure* argumentation (Prakken, 2018), between a proponent and an opponent. Consequently, each agent has the goal to establish a convincing line of argumentation and to weaken the one of the opposing side. The proposed system is used to generate artificial dialogues which are then assessed in a user study regarding their logical consistency and naturalness. Although the results indicate the compatibility and general functionality of the selected models, they also highlight several pending issues that are addressed one by one in the remainder of the thesis. The present chapter is based on the work published in (Rach et al., 2019a) and builds in parts on the bachelor's thesis of Langhammer (2017).

## 4.1 Architectural Design

As a first step, the preliminaries for the methods proposed in the following chapters are introduced. To this end, the use of a general dialogue system architecture is motivated and the positioning of the thesis contributions in this architecture is discussed. Based on this setup, the formal representation of arguments and their relations is addressed next. The selected framework is utilized to generate an argument structure based on human annotations. It is used for the evaluation in the second part of the chapter and also serves as a reference and/or default setting in the following chapters. Subsequently, the dialogue model to structure the interaction is discussed, including the specific framework utilized for evaluation. As the main focus of this thesis is the development of approaches that are independent of a specific application or setup, the flexibility of the chosen frameworks is the main selection criteria throughout this section.

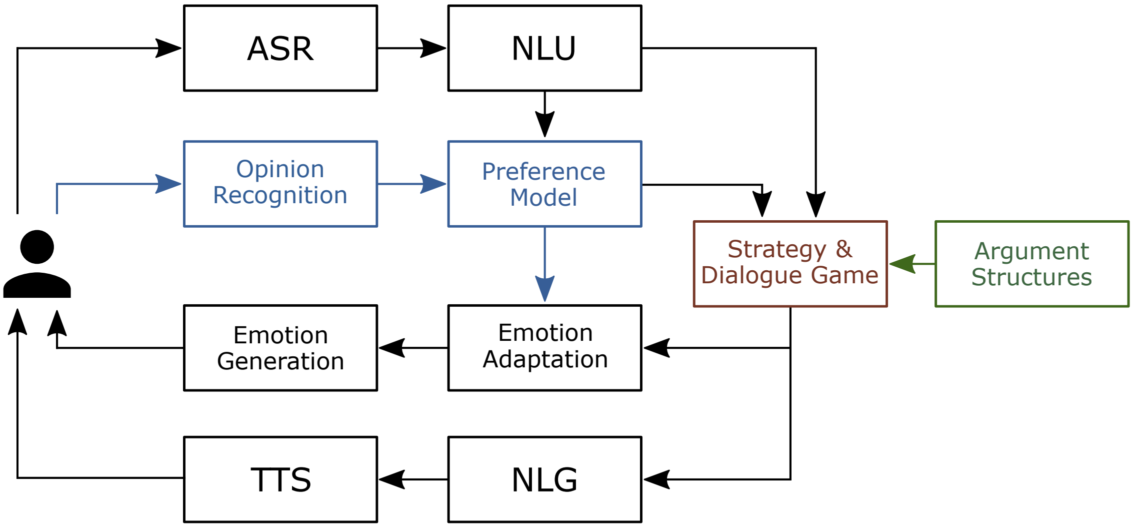


Figure 4.1: Positioning of the approaches proposed in Chapter 5 (brown), Chapter 6 (green) and Chapter 7 (blue) in a multimodal pipeline architecture.

### 4.1.1 Dialogue System Architecture

As discussed in Section 2.1, two conceptually different dialogue system architectures can be found in literature, namely modular pipeline architectures and data-driven end-to-end architectures. However, the complexity of argumentation renders flexible end-to-end solutions unfeasible (Slonim et al., 2021). Hence, most of the current argumentative dialogue systems utilize some sort of modular approach to ensure functionality, as discussed in Chapter 3. The present thesis follows this conceptual approach by dividing the overall task into three sub-tasks and consequently also utilizes a pipeline architecture. The methods proposed in the following chapters to solve the individual sub-task can then be implemented as different modules. It is thereby possible to flexibly combine the investigated approaches with each other and additional modules in a way that suits the task of the corresponding system best.

Figure 4.1 shows the positioning of modules corresponding to the three herein considered sub-tasks of challenging agent strategies (brown), topic flexibility (green) and user-adaptation (blue) in a multimodal pipeline architecture. The development of agent strategies and the modelling of the interaction as a dialogue game are part of the dialogue management and discussed in Chapter 5. They determine the general capacity of the system to handle a specific type of interaction. The argument structures encode the knowledge of the system regarding the topic under discussion and replace the application in the traditional dialogue system setup. They hence serve as a system database that directly influences possible replies and strategies. The methods proposed in Chapter 6 to acquire argument structures for dialogue systems build on this positioning and, in particular, on the compatibility of the DM module with the respective argument representation. Finally, the methods proposed for user-adaptive argumentation in Chapter 7 are encoded in the multimodal extension of the traditional architecture and consequently as separate modules. This separation is required as the approach to recognize user opinions relies on non-verbal cues which are not included in the speech signal. In addition, the adaptation approach discussed in Chapter 7 utilizes the proposed preference model to adapt the conveyed system emotion and hence employs an additional system modality.

As previously mentioned and evident from the module dependencies in Figure 4.1, the DM (i.e. dialogue game and strategy) and the formal representation of arguments utilized in the argument structure need to be compatible to ensure functionality. Framework choices considering these dependencies are introduced and motivated in the following sub-sections alongside specific instantiations that are used in the evaluation study.

### 4.1.2 Argument Structure and Data Annotation

In the following, the selection of a suitable framework to represent arguments is discussed. In contrast to the dialogue model which formalizes the interaction, this representation formalizes the knowledge about existing arguments and their relations towards each other. Suitable candidates are hence inference based (or monological) models for argumentation as discussed in Section 2.2.1. A formal representation of this kind is called an *argument structure* for the remainder of the thesis. Due to the extensive research in the field of inference-based argumentation, a variety of argument representations is available which can be (to some extent) transferred into another. For a definite selection of an appropriate model, the following requirements are derived based on the tasks addressed in this thesis: For the goal of providing conversational agents with the capability to discuss various topics, automatic approaches to retrieve arguments, i.e. argument mining techniques are of particular interest. Consequently, the respective representation of arguments has to be compatible with these techniques to enable their use in conversational agents. This is hence a requirement from the acquisition side. In contrast, the second requirement results from the desired application of the representation, i.e. the artificial argumentation with conversational agents. In particular, the representation should enable the agents to construct arguments from it during the discussion. However, the order of the arguments in the discussion presumably differs from the one in the original source the arguments are retrieved from. Therefore, the utilized model has to enable an abstraction from the specific argument structure in the original source while at the same time providing building blocks for arguments that are compatible with the utilized dialogue model.

Both requirements render a fine-grained representation like the Toulmin model impractical for the present case. Even though it was used in some instances for argument mining experiments, the corresponding insights provided in (Habernal et al., 2014) indicate that the explicit form is frequently not represented in natural language sources. This makes an application of the model difficult for a lot of sources (especially web documents). Consequently, the focus herein is on claim-premise representations which are less specific and hence provide more flexibility. The utilized framework is based on the argument annotation scheme in (Stab and Gurevych, 2014a), also discussed in Section 2.2.1. It distinguishes the three argument component types *Major Claim*, *Claim* and *Premise* as well as the two relations *support* and *attack*. While it was introduced for a specific type of source, namely persuasive essays, and includes a corresponding composition of arguments it still provides the desired flexibility: The formal difference between the three component types is only their allowed relations to other component types and hence only determines

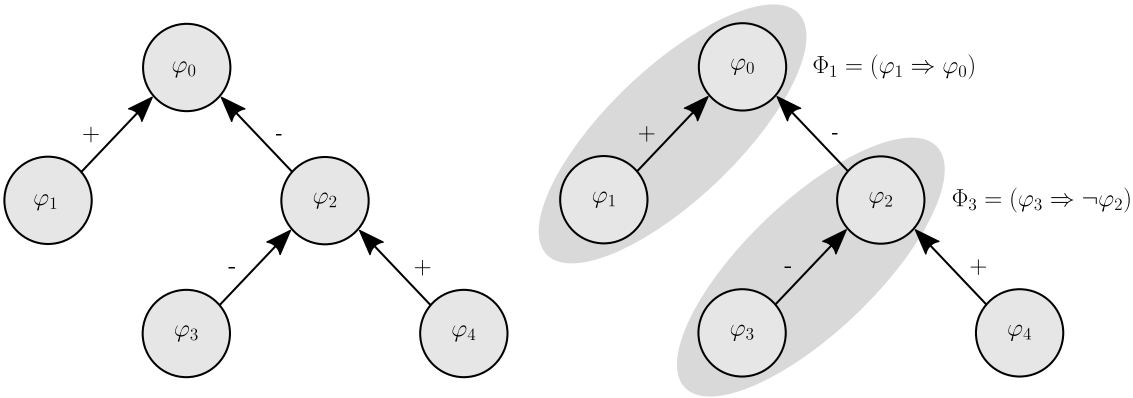


Figure 4.2: Left side: Abstract tree structure derived from the annotated argument components and their relations. Component types are abstracted into components $\varphi_i$. Right side: Generation of arguments from the abstract tree structure utilizing only components and relations.

their position in the resulting argument structure. In addition, the condition that each component targets exactly one other component with a relation is adapted from (Stab and Gurevych, 2017) to prevents circular relations. Consequently, the resulting argument structure can be abstracted as a directed tree with the *Major Claim* (or topic) as root, the remaining components as nodes and the relations as edges. As the definition of an argument as a set of *Premises* that are (directly or indirectly) linked to a *Claim* also imposes no restrictions regarding the structure, it can be abstracted from as well to generate new arguments in the context of the discussion or apply the annotation scheme to new types of sources. For the remainder of this thesis, the different component types are considered only for acquisition purposes (if beneficial) and abstracted from otherwise. In the abstract representation, components of all three kinds are formally denoted as $\varphi$ from which arguments of the form $\Phi_i = (\varphi_i \Rightarrow \varphi_j)$ ($\varphi_i$ supports $\varphi_j$) and $\Phi_i = (\varphi_i \Rightarrow \neg\varphi_j)$ ($\varphi_i$ attacks $\varphi_j$) are derived for use in the discussion. The abstract tree representation and the construction of arguments are shown in Figure 4.2. It can be seen that the abstract tree structure bears similarity to bipolar argumentation frameworks (Cayrol and Lagasquie-Schiex, 2005), with the difference that therein the nodes are complete arguments and not components. The formal distinction between arguments and argument components is necessary for the application of the dialogue model discussed in the next subsection but the similarities are exploited in Chapter 7 to apply concepts developed for BAFs to the utilized structures.

In (Rach et al., 2019a), we introduce a sample argument structure based on human annotation for the desired evaluation and as a reference for subsequently acquired structures. The source of choice is a sample debate from the *Debatabase* of the idebate.org[1] website. The reasons for this choice are as follows: Firstly, idebate.org is operated by the International Debate Education Association (IDEA), a global network of organizations devoted to debating education. Hence, the

[1] https://idebate.org/debatabase (last accessed 29 August 2021)

debates offered here can be expected to meet certain quality standards regarding both form and content. Secondly, all debates explore both sides of their respective topic. Lastly, all *Debatabase* debates adhere to a specific structure which both facilitates the quick screening for suitable candidates and potentially aids the argument annotation process later on. The sample debate employed in the scope of this work is concerned with the topic *Marriage is an outdated institution*. This choice is mostly due to the high amount of arguments provided by the *Debatabase* for this topic. The annotation was done by an expert based on the guidelines of Stab and Gurevych (2014a) by first identifying argument components in the debate and secondly annotating the relations between them. The identification of the overall topic and stance of the author included in the original work as a separate step are not required here, as both aspects are brought out by the structure of the debates. The annotation resulted in a total of 72 argument components (one *Major Claim*, ten *Claims* and 61 *Premises*) and their corresponding relations and is encoded in an OWL ontology (Bechhofer, 2009) for further use. In order to facilitate the natural language generation (NLG) in the envisioned argumentative systems, the original annotated sentences are modified slightly to form a complete and reasonable utterance. Thus, implications are made explicit, references and citations are reformulated and expressions exclusively used in the debate format are adapted to the dialogue context (for example *"[...] as the opposition claims"* is changed to *"[...] as you claim"*).

### 4.1.3 Dialogue Model

The dialogue management in the previously discussed architecture requires a model of the interaction to determine aspects like turn-taking (and hence speaking time), the outcome of the discussion as well as possible strategies. As the main focus of this work is the flexibility of the proposed approaches, a dependency on additional resources like conversational training data has to be minimized. Consequently, a formal model of the dialogue is considered (instead of a data-driven one). In the scope of this thesis, the interaction between interlocutors is formalized as a dialogue game for argumentation. This choice is due to several reasons: First of all, dialogue games are a well-studied approach to model dialogical argumentation and hence include theoretical insights in the interaction model. Moreover, they are not limited to the exchange of arguments, like in approaches that rely directly on abstract or structured argumentation frameworks (Rosenfeld and Kraus, 2016). Instead, they also include moderating utterances like requesting additional information regarding a specific aspect (*"Could you please elaborate on $\varphi$?"*), conceding to an aspect (*"I think we can agree on $\varphi$."*) or asking critical questions (*"Is it the case that $\varphi$?"*). Finally, it was argued that any form of argumentative dialogue discussed in (Reed and Norman, 2003) can be formalized as a dialogue game (Atkinson et al., 2017) which makes this a flexible and reasonable choice. The system of Prakken (2005) is used as a starting point for evaluation and testing since it was motivated by providing a flexible and moderate framework that ensures logically coherent dialogues. As discussed in Section 2.2.2, it is based on a multi-move and multi-reply protocol, meaning that players in the game can play several moves in a row and can also introduce multiple responses to one utterance. For the desired testing scenario, the framework is implemented *as is* without additional modifications.

As introduced in Section 2.2.2, the game is formally defined as $(\mathscr{L}, D)$, where $\mathscr{L} = (L_t, R, Args, \rightarrow)$ denotes a logic for defeasible argumentation and $D = (L_c, P, C)$ the dialogue sys-

Table 4.1: Arguments in *Args* generated from the argument structure (Arg. Structure) and implications in the logic for defeasible argumentation $\mathscr{L}$.

| Arg Structure | *Args* | $\mathscr{L}$ |
|---|---|---|
| $\varphi_j$ supports $\varphi_i$ | $\Phi_j = (\varphi_j \Rightarrow \varphi_i)$ | - |
| $\varphi_l$ attacks $\varphi_j$ | $\Phi_l = (\varphi_l \Rightarrow \neg\varphi_j)$ | $\Phi_l$ defeats $\Phi_j$ |
| $\varphi_h$ supports $\varphi_j$ | $\Phi_h = (\varphi_h \Rightarrow \varphi_j)$ | $\Phi_h$ extends $\Phi_j$ |

Table 4.2: Communication language $L_c$ of the dialogue game for argumentation (Prakken, 2005) with arguments of the herein considered form.

| Speech Act | Attacks | Surrenders |
|---|---|---|
| *claim*($\varphi_i$) | *why*($\varphi_i$) | *concede*($\varphi_i$) |
| *why*($\varphi_i$) | *argue*($\varphi_j \Rightarrow \varphi_i$) | *retract*($\varphi_i$) |
| *concede*($\varphi_i$) | - | - |
| *retract*($\varphi_i$) | - | - |
| *argue*($\varphi_j \Rightarrow \varphi_i$) | *why*($\varphi_j$), *argue*($\varphi_l \Rightarrow \neg\varphi_j$) | *concede*($\varphi_j$) |

tem proper. In the present case, $\mathscr{L}$ is derived from the argument (tree) structure discussed previously, and hence specified as follows: The set of arguments *Args* includes all arguments that can be created from the tree structure and have the form $\Phi_j = (\varphi_j \Rightarrow \varphi_i)$ if $\varphi_j$ supports $\varphi_i$ and $\Phi_j = (\varphi_j \Rightarrow \neg\varphi_i)$ if $\varphi_j$ attacks $\varphi_i$. For an argument $\Phi_j = (\varphi_j \Rightarrow \varphi_i)$ it follows directly that $conc(\Phi_j) = \varphi_i$ and $prem(\Phi_j) = \{\varphi_j\}$. It should be noted that in the context of the dialogue game, only the abstract representation of the argument structure is utilized, meaning that *claim* moves are not related to the component type *Claim* in the annotation scheme. The same holds for the component type *Premise* which is not to be confused with the premise $prem(\Phi_i)$ of an argument $\Phi_i \in Args$. An argument $\Phi_i$ then defeats another argument $\Phi_j$, if the conclusion of $\Phi_i$ contradicts the premise of $\Phi_j$. Conversely, an argument $\Phi_i$ extends another argument $\Phi_j$, if the conclusion of $\Phi_i$ equals the premise of $\Phi_j$. The generation of arguments and the implications in $\mathscr{L}$ regarding defeat and extensions are summarized in Table 4.1.

Based on the specific arguments, the communication language $L_c$ is also specified as shown in Table 4.2. A direct consequence of this specification is that the relation of each speech act is unique as well, meaning that each speech act attacks or surrenders to one specific other speech act. For speech acts of the type *argue*, this is a direct result of the unique target relation in the tree structure. It ensures that each argument $\Phi \in Args$ either extends or defeats one unique other argument. Moreover, since each argument has a single premise, speech acts that include individual argument components $\varphi \in L_t$ also have a unique relation since the corresponding component is the premise of one (and only one) argument and can consequently only be introduced to the dialogue in the context of this argument.

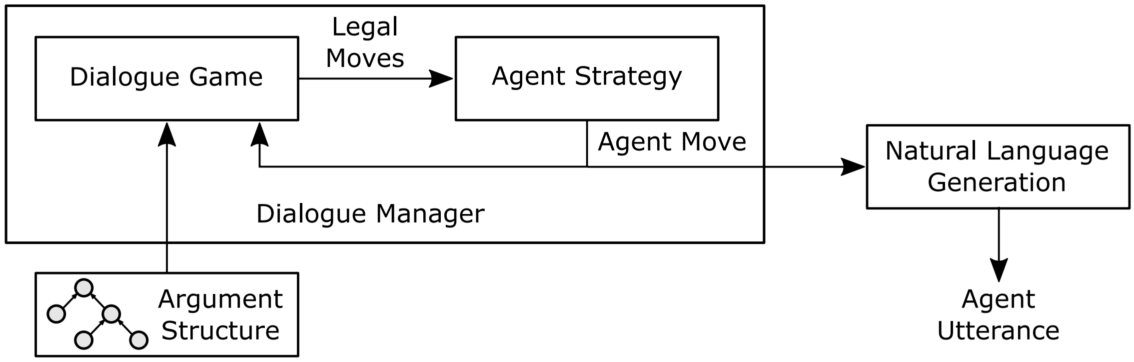


Figure 4.3: Sketch of the multi-agent system proposed as experimental setup.

## 4.2 Pilot Study

To validate the previous model choices and identify pending issues, a multi-agent setup including the introduced argument structure, a dialogue manager and a natural language generation (NLG) is introduced next. It is used to generate artificial dialogues between the two agents Alice and Bob. The transcripts of these dialogues are then discussed and evaluated subsequently. To this end, one of the artificial dialogues as well as an excerpt of a human-generated discussion are examined first to highlight structural differences. The findings of a user study comparing both types of dialogues are discussed. The focus of the survey is on the logical consistency of the lines of argumentation established by the agents as this consistency depends directly on the interplay of the model choices, i.e. the argument structure and the dialogue game. In addition, pending issues in the current setup are identified. The work in this section is also published in (Rach et al., 2019a).

### 4.2.1 Agent-Agent System

The core of the proposed evaluation system consists of two agents (Alice and Bob) that argue about a certain topic. Their interaction is formalized as a dialogue game which determines the player to move and the available moves in each state of the dialogue. Both agents choose their next move from this list based on an agent strategy. The set of available arguments within the game is provided by the previously discussed argument structure. For the desired evaluation, each dialogue has a fixed length, i.e. a maximal number of turns after which the interaction ends and thus, no termination criterion is employed. Each played move is subject to a template-based NLG transforming the game moves into a natural language utterance. In the following, the agent strategy and the NLG utilized in the pilot study are described in more detail. The complete architecture of the system is depicted in Figure 4.3.

### Agent strategy

As the utilized dialogue game for argumentation only imposes a set of formal rules, a strategy to select the next move from the list of allowed ones is required. For the desired evaluation, probabilistic rules that allow for the generation of multiple different dialogues with the same structure are employed. This strategy also serves as a strong reference for the optimization of agent strategies addressed in Chapter 5. The corresponding rules are based on the condition, that both agents attack whenever possible. Consequently, they only surrender to an opponent move if there is no other option left. This choice is reasonable for the discussed competitive case, as each participant has the goal to convince the opponent and thus to strengthen the own and weaken the opponent's line of argumentation whenever possible. To keep the dialogue focused on the current topic, agents prefer moves that respond to the latest opponent move over topic switches. In addition, a preference of *argue* over *why* moves is added to prevent extensive use of the latter one. The rules thus read as:

- Attack if possible. If you do so,
  - If possible, attack the previous utterance of the opponent.
  - Prefer *argue* moves over *why* moves.
- If no attack is possible, *surrender*. If possible, *surrender* to the latest opponent move.

By using these rules, each agent identifies the next move from the set of allowed ones. If there is more than one move fulfilling the same conditions, the next move is chosen from this subset randomly.

### Natural Language Generation

The NLG of the system relies on the original textual representation of the argument components, i.e. the sentences in the original source. As discussed in Section 4.1.2, the annotated sentences are slightly modified to form a stand-alone utterance which serves as a template for the respective *argue* (and *claim*) move. In case an *argue* move refers to the latest move of the opposing player, the utterance only includes the premise of the argument and the conclusion is left implicit. In addition, a list of natural language representations for each additional type of move is defined. The explicit formulation is chosen from this list randomly during the generation of the utterances. Examples of these formulations are *"Could you be more specific?"* for *why* moves and *"It's hard to disagree with that. I see your point there."* for *concede* moves.

As the game protocol allows both agents to backtrack, i.e. to address utterances that are not the immediate predecessor, natural language indicators for these *topic switches* are also included. They are comprised of a referencing formulation, followed by the textual representation of the referenced argument component as for example *"I'd like to go back to something you mentioned earlier. I think you said: [...]"*. Again, the explicit formulation is chosen randomly for each utterance. In the case of an *argue* move, this reference (formally) is the conclusion of the utilized argument. Finally, connecting phrases like *"That's not a valid point."* are added to some *argue* moves in order to increase the naturalness of the dialogue.

Table 4.3: Excerpt of an example dialogue generated with the dialogue game for argumentation and the annotated tree structure. Sentences that correspond to an annotated argument component in the argument structure are shown in italic font.

| Speaker | Utterance | Speech Act |
|---|---|---|
| Alice | *Marriage is an outdated institution.* | claim($\varphi_0$) |
| Bob | Why do you think that? | why($\varphi_0$) |
| Alice | *The frequency and accessibility of divorce undermines the entire purpose of marriage.* | argue($\varphi_1 \Rightarrow \varphi_0$) |
| Bob | Could you please elaborate? | why($\varphi_1$) |
| Alice | *Marriage no longer leads to a stable or permanent relationship.* | argue($\varphi_2 \Rightarrow \varphi_1$) |
| Bob | That's not a valid point. *The purpose of marriage is not an eternal, unrelenting union, whether it is wanted or not.* | argue($\varphi_3 \Rightarrow \neg\varphi_2$) |
| Alice | Could you be more specific? | why($\varphi_3$) |
| Bob | *The purpose of marriage is to foster a more stable relationship than would be possible without marital vows.* | argue($\varphi_4 \Rightarrow \varphi_3$) |
| Alice | Could you be more specific? | why($\varphi_4$) |
| Bob | Well, maybe we can come back to this point later. You said earlier: *Marriage no longer leads to a stable or permanent relationship.* I'm not sure I understand what you're getting at. | why($\varphi_2$) |
| Alice | *Our society no longer respects marriage as a permanent institution.* | argue($\varphi_5 \Rightarrow \varphi_2$) |
| Bob | Could you be more specific? | why($\varphi_5$) |
| Alice | *Serial monogamy is becoming ever more common.* | argue($\varphi_6 \Rightarrow \varphi_5$) |
| Bob | Could you please elaborate? | why($\varphi_6$) |
| Alice | *50% of all divorces in the UK go on to remarry according to the Office for National Statistics.* | argue($\varphi_7 \Rightarrow \varphi_6$) |

### 4.2.2 Artificial Dialogue

In the following, an artificial dialogue generated with the system by means of the annotated argument structure is compared to a human-generated dialogue[2]. The artificial dialogue, including speaker, NLG output (utterance) and the corresponding speech act are shown in Table 4.3.

It can be seen that this dialogue does not include a surrendering move, as both agents prefer attack moves and the dialogue has fixed length, so no agent has run out of attack options yet. In addition, there is one topic switch (*why*($\varphi_2$)) by Bob, which makes a new branch of the tree structure available for both agents and allows Alice to support one of her points ($\varphi_2$) with additional arguments. Overall, a frequent use of *why* moves can be observed due to the fact that the explicit reply structure does not allow for chained arguments, i.e. an extended utterance covering multiple arguments. This is particularly evident in comparison with the excerpt of a human-generated dialogue shown in Table 4.4, where the opposite can be observed. In this case, a single utterance

[2]Material reproduced from www.idebate.org with the permission of the International Debating Education Association. 

Table 4.4: Excerpt of a human-generated dialogue using the same textual source. Content taken from the original debate is shown in italic font.

| Speaker | Utterance |
|---|---|
| Alice | *Marriage is an outdated institution.* |
| Bob | Why do you think that? |
| Alice | *According to the Office for National Statistics the average age, in the UK, to get married is approximately 30 years old. Life expectancy in the UK is approximately 80 years. This means the average marriage expects people to commit to maintain a certain way of life for a period that is longer than they have actually been alive. This goes hand in hand with the rise of social acceptability of people having more than one life partner in their life to show that either marriage is an unreasonable expectation of someone or a meaningless charade that is not actually expected to be maintained.* |
| Bob: | *This argument only works under the assumption that we live in a society where divorce does not exist. If a person enters into a marriage without full awareness of what they have committed to and later need to get out of that marriage, they are free to. Being able to leave a marriage, though, does not make marriage a meaningless charade, as you claim. It is still more difficult to leave a marriage than it is to leave a non-marital committed relationship and so it makes a big difference.* |

encapsulates multiple argument components and the dialogue as a whole consists of fewer turns. Moreover, only one *why* move is present in this case. The last difference regarding the line of argumentation is that in the human-generated discussion some premises occur before the corresponding conclusion is present. Up to now this is not possible for virtual agents in our system as it could lead to inconsistent lines of argumentation and is only reasonable in combination with utterances containing multiple arguments. In order to enable the system to consistently argue this way, a modification of the employed dialogue game is required. This modification has to extend the rules of the game while at the same time preserving the protocol properties that ensure coherent dialogues.

### 4.2.3 Survey

To validate the selection of models and evaluate their interplay, the two kinds of dialogues discussed above were compared in a pilot study realized by clickworker[3]. Each participant was given one random instance out of the two possible categories (agent-agent or human dialogue). To include all aspects of the original debate, five human-generated dialogues and 20 agent-agent dialogues were utilized as the argument density is higher in the human case. The 122 participants were from the UK and assigned randomly to one instance, resulting in a splitting of 54 participants rating the agent-agent case and 68 participants rating the human-generated case. At the beginning of the survey, information about the purpose of the study and instructions regarding the rating process were provided. In particular, participants were asked to assess the logical consistency of

[3]https://marketplace.clickworker.com (last accessed 29 August 2021)

the argumentation instead of providing a personal opinion regarding the discussed content. The rating was done on a five-point Likert scale from *completely disagree* (1) to *completely agree* (5) for ten statements about the persuasiveness of the involved parties, logical consistency of the argumentation and an overall impression of the dialogue. The statements distinguish between the two agents Alice and Bob and assess both of them separately:

- I was not convinced by Bob/Alice and how he/she presented his/her case. (Strat. Bob/Alice)
- It was always clear which previous utterance Bob/Alice addressed in his/her turn. (Prev. Bob/Alice)
- The arguments presented by Alice/Bob are logically consistent responses to the utterances they refer to. (Arg. Bob/Alice)
- Alice's/Bob's line of argumentation is not logically consistent. (Arg. line Bob/Alice)
- It was difficult to follow the line of argumentation throughout the debate. (Arg. line diff.)
- The whole debate is natural and intuitive. (Nat. and int.)

It should be noted that due to the different formulations the best ranking is not always the highest. Table 4.5 shows the corresponding statistical results for all statements. Each line includes the median for the artificial dialogues (Agent), the human dialogues (Human) and the corresponding $p$ value achieved with a Mann-Whitney-U test (Mann and Whitney, 1947) and all 122 ratings. As mentioned in the beginning, the focus of this section is on the statements assessing the logical consistency of the argumentation. These are in particular the ones asking for appropriateness of the arguments (Arg. Bob/Alice) and the statements that assess the complete line of argumentation (Arg. line Bob/Alice). The statements related to the agent strategy (Strat. Bob/Alice) were posed in order to decouple the rating of the dialogical behaviour from the rating of the lines of argumentation. Thus, the corresponding results are not discussed further, here.

It can be seen that in the case of the overall consistency (Arg. line Bob/Alice) the ratings for the different scenarios are close to each other as the median is equal for both agents. Moreover, the case of Bob yields no significant difference whereas Alice is on the threshold of $p = 0.05$. For the single step rating (Arg. Bob/Alice), both cases show a significant difference between the human-generated and the agent-agent dialogue. Nevertheless, the median is the same for the case of Bob and still above the neutral value of 3.0 for Alice. As mentioned earlier, the agent-agent dialogues show a frequent use of *why* moves as well as some unintuitive topic changes that may have lead to distraction and irritation of the participant. This is mostly reflected in the different ratings for the two last statements which assess the overall impression of the dialogue. As the argumentation was nevertheless in each case rated as consistent, it is concluded that the utilized argument structure is an adequate approach to represent arguments in systems of the discussed kind.

Table 4.5: Median and $p$ value for both agent-agent (Agent) and human-generated (Human) dialogues. Bold lines indicate statements related to the logical consistency of the argumentation.

| Statement | Agent | Human | $p$ |
|---|---|---|---|
| I was not convinced by Bob and how he presented his case. | 3.0 | 2.5 | 0.11 |
| I was not convinced by Alice and how she presented her case. | 3.0 | 3.0 | 0.26 |
| It was always clear which previous utterance Bob addressed in his turn. | 4.0 | 4.0 | 0.06 |
| It was always clear which previous utterance Alice addressed in her turn. | 4.0 | 4.0 | 0.29 |
| **The arguments presented by Bob are logically consistent responses to the utterances they refer to.** | **4.0** | **4.0** | **0.02** |
| **The arguments presented by Alice are logically consistent responses to the utterances they refer to.** | **3.5** | **4.0** | $\leq$ **0.01** |
| **Bob's line of argumentation is not logically consistent.** | **2.0** | **2.0** | **0.72** |
| **Alice's line of argumentation is not logically consistent.** | **2.0** | **2.0** | **0.05** |
| It was difficult to follow the line of argumentation throughout the debate. | 3.0 | 2.0 | $\leq$ 0.01 |
| The whole debate is natural and intuitive. | 2.0 | 4.0 | $\leq$ 0.01 |

## 4.3 Discussion and Conclusion

Throughout this chapter, preliminary choices of the general frameworks and models utilized in the following chapters were discussed. This included the considered dialogue system architecture as well as formal frameworks to represent arguments and structure the interaction. To validate the choices, a multi-agent setup capable of generating artificial persuasive dialogues between two agents was proposed. Transcripts of these dialogues were then analyzed and assessed in a pilot study.

The comparison of the artificial dialogues with the human-generated ones indicate different strengths of the proposed setup but also reveal room for improvement in several instances. Since the investigated dialogues were perceived as logically consistent by the study participants, the selected combination of argument structure and dialogue game fulfils the requirement of ensuring coherent interaction. Consequently, the task of providing flexibility with respect to the discussed topics can now be reduced to the mapping of arguments in available sources into the investigated tree structure. Since neither the dialogue model nor the argument structure include topic-specific dependencies, this can be approached as a general argument mining task and is thoroughly discussed in Chapter 6.

The most obvious difference between the artificial dialogues and the human-generated ones in the perception of the users is their naturalness and the difficulty in following the discussion.

A possible explanation for this perception is revealed by the direct comparison of example dialogues from both groups: Whereas the artificial dialogue consists of short utterances and includes multiple explicit requests for additional information, the utterances in the human-generated dialogue are more extensive and combine multiple aspects into a single utterance. This is a direct consequence of the explicit reply structure in the dialogue game since it only allows for extending the line of argumentation after an explicit request for further information. In addition, the utilized rule-based strategy does not explore the multi-move nature of the game protocol due to its focus on attack moves. Although related, the two issues are shortcomings on different levels of abstraction: The repeated *why* moves are directly related to the specific dialogue game instantiation, whereas the rule-based strategy is a general issue that is independent of the specific formal model. Hence, both points are approached in separate steps. The inflexibility of the agent strategy is addressed through the application of reinforcement learning in the general context of dialogue games for argumentation. In contrast, the repeated *why* moves are addressed through an extension of the herein utilized dialogue game. Both steps are discussed extensively throughout Chapter 5.

In general, the proposed architecture yields promising results with respect to full-scale application systems and also in its capacity to reveal pending issues in the utilized models. However, user-related issues are not considered in the present scenario, since the dialogues are generated between virtual agents. As interaction with human users is arguably more complex and requires additional functionalities of the system, extensions of the setup with respect to user adaptation are discussed in Chapter 7.

# 5 Development of Challenging Agent Strategies

In the following, the first sub-task of the thesis, namely the development of flexible agent strategies is addressed. Since formal modelling of the interaction is considered herein, this task can be divided into the development of models that formalize the desired scenario and the development of suitable strategies within these frameworks (Prakken, 2018).

The first part of the chapter is focused on the appropriate selection of utterances in a given dialogue model, thereby building on existing work on formal frameworks. To ensure flexibility, the proposed approach abstracts from specific topics, does not depend on any pre-defined strategy or training data and enables the system to explore multiple strategies within a given dialogue framework. The general idea of this part was first published in (Rach et al., 2017b) and is thoroughly introduced in (Rach et al., 2018b). Although the strategy can include additional aspects like the emotional tone with which an utterance is conveyed or the wording of this utterance (i.e. how something is said), the decision about the next utterance (i.e. what is said) is at the core of each strategy and hence considered first. The additional aspects are addressed in the context of user-adaptive argumentation in Chapter 7 and Chapter 8 based on the results of the present one.

The second part of the chapter addresses the specific issues raised by the preliminary evaluation in Chapter 4, namely the naturalness of dialogues generated with the investigated dialogue game. To this end, the utilized formal model is extended to increase the freedom of choices for the players. The extension is focused on lifting some constraints of the original model while at the same time preserving the properties that ensure logically coherent dialogues. The approach is introduced and evaluated in (Rach et al., 2020b).

## 5.1 Markov Games for Persuasive Dialogue

The problem of identifying the next utterance in an argumentative dialogue is a decision making problem and herein formalized as markov game (Littman, 1994). The main advantage of this formalism is that it enables the use of multi-agent reinforcement learning and hence to optimize the agent strategy in self-play with respect to the received reward. Consequently, no pre-defined opponent strategy or training data of the conversation is required. As for the interaction model, the focus is on dialogue games for argumentation as they set reasonable boundaries for the strategy based on theoretical insights while abstracting from the specific topic under discussion. Moreover, this choice offers the desired flexibility as it was argued that argumentative dialogue between agents can always be formalized as a dialogue game (Atkinson et al., 2017).

First, a general reformulation of dialogue games for argumentation as markov games is proposed. This formal step is independent of a specific dialogue game instantiation. Subsequently, the approach is applied to an existing dialogue game with a known optimal strategy for a proof-of-principle evaluation. In addition, a more complex winning criterion in the dialogue game is investigated to allow for the learning of more complex strategies.

### 5.1.1 Formalization

This section addresses the formulation of dialogue games for argumentation as a markov game as introduced in (Rach et al., 2018b). As discussed above, this formulation is applicable to any game that has the formal structure discussed in Section 2.2.2 and is not limited to a specific instantiation. The required definitions of a general dialogue game for argumentation and a markov game are recalled before discussing the reformulation of one into the other. A dialogue game for argumentation is given as tuple $(\mathscr{L}, D)$ with $\mathscr{L}$ a logic for defeasible argumentation and $D = (L_c, P, C)$ the dialogue system proper that encodes the communication language $L_c$, the protocol $P$ and the commitment rules $C$. A markov game on the other hand is defined as $(I, S, \boldsymbol{A}, \boldsymbol{r}, T, \gamma)$, with the set of agents $I$, the state space $S$, the joint action space $\boldsymbol{A}$, the joint reward function $\boldsymbol{r}$, the transition function $T$ and the discount factor $\gamma$. The dialogue game is played by two players (proponent and opponent), which are now associated with the agents in the markov game $p \in I = \{1,2\}$. As in Section 2.2.2, a dialogue $d_t$ that is continued with the move $m_{t+1}$ is denoted as $d_t, m_{t+1}$. The reformulation is then approached by defining each component of the markov game in terms of dialogue game components, starting with the state space $S$.

**Definition 13** (State Space)**.** *Let $d_t = m_1, ..., m_t$ be a dialogue in the dialogue game for argumentation at time t and $C_t^p$ the commitments of player p. The state at time t is then defined as*

$$s_t := (d_t, C_t^1, C_t^2). \tag{5.1}$$

Consequently, each state encodes the complete dialogue history up to time $t$ and the state space $S$ is the set of all possible states $s$. The action space of each player $A_p$ is defined based on these states as follows.

**Definition 14** (Action Space)**.** *Let $s_t$ be the state at time t as defined above and let $d(s_t)$ denote the dialogue encoded in that state. Further, let $2^M$ be the power set of the set of all moves M and $Pr : M^{<\infty} \to 2^M$ be the protocol function that maps each dialogue to a set of legal moves in the dialogue game. Then the set of actions available to the current player to move p at time t is defined as*

$$A_p(s_t) := Pr(d(s_t)). \tag{5.2}$$

The available actions are thus equivalent to the legal moves in the dialogue game for argumentation (see Equation 2.3). To include multi-move protocols, a *wait* action $a_w$ that does nothing is assigned to the player who is currently not to move, i.e. to the current winner of the dialogue game. The complete action space for every player $p \in I$ is hence given as $A_p = M \cup \{a_w\}$.

The transition function $T : S \times \boldsymbol{A} \times S \to [0,1]$ determines the state transitions given a state and the actions of both players. As a deterministic environment is considered in the present case, it can be expressed as $T : S \times \boldsymbol{A} \to S$ and is defined as follows.

**Definition 15** (Transition Function)**.** *Let $s_t$ be the state at time t, $a_t^p$ the action of player p at time t and $m_{t+1}$ the move that corresponds to the action of the player to move. Then the transition function is given as*

$$T(s_t, a_t^1, a_t^2) := (d_{t+1}, C_{t+1}^1, C_{t+1}^2). \tag{5.3}$$

*with $d_{t+1} = d_t, m_{t+1}$ and $C_{t+1}^p$ the commitments of player p updated with $m_{t+1}$ according to the game protocol P.*

Finally, the reward function of each agent is derived from the winning criterion of the dialogue game. According to Prakken (2006), a dialogue game for argumentation of the herein discussed kind is a zero-sum game, meaning that whenever one player wins, the opposing side loses the game. Consequently, the reward function for each agent can be defined as follows.

**Definition 16** (Reward Function)**.** *Let $d(s_t)$ be the dialogue encoded in state $s_t$ at time t and let a dialogue be called final if it is terminated by the game protocol P. Then, the reward function of player p is defined as*

$$r_p(s_t, a_t^1, a_t^2) := \begin{cases} 0, & \textit{if } d(s_t), m_{t+1} \textit{ is not final} \\ 20, & \textit{if } d(s_t), m_{t+1} \textit{ is final and p wins according to } P \\ -20, & \textit{if } d(s_t), m_{t+1} \textit{ is final and p loses according } P \end{cases}, \tag{5.4}$$

*with $m_{t+1}$ the move that corresponds to the action of the player to move.*

It should be noted that the herein introduced formalization does not rely on the zero-sum assumption and can be applied to general-sum games as well by modifying the reward function appropriately. Finally, the discount factor is set to $\gamma = 0.9$ in the herein considered case.

The above-introduced formulation is as general as possible but also results in a large state space that is impractical for implementations. Consequently, specifications of this general formalization depending on the specific utilized formalism are of particular interest. In the following, two modifications are introduced to decrease the dimensionality of the state space. The first one is rather general and relies on a single assumption regarding the game protocol, whereas the second one is tailored to the dialogue game (Prakken, 2005) utilized for experiments. For the first modification, it is assumed that the protocol of the corresponding game meets the following condition:

*If the legality of a move depends on the temporal order of previous moves, this dependency is restricted to the latest move.*

Although it is not fulfilled by definition, there is (to the best of our knowledge) no actual formal system violating this assumption. The most common systems in which the temporal order of arguments is relevant are the ones with immediate-response protocols and systems of this kind do not violate the above posed-condition. Thus it is a rather weak restriction regarding implementations. If it is fulfilled, the definition of the state space can be adapted as follows.

**Definition 17** (Reduced State Space)**.** *Let $d_t = m_1, ..., m_t$ be a dialogue in the dialogue game. Further, let $G_t = (g_t, e_t)$ be a graph with $g_t \subseteq L_c$ the set of all speech acts included in moves in $d_t$ and $e_t \subseteq R_a \cup R_s$ the set of corresponding speech act relations used in $d_t$. The reduced state of player p at time t is then defined as*

$$s_t := (G_t, C_t^1, C_t^2, m_t). \tag{5.5}$$

Table 5.1: Overview of the reformulations (complete and reduced) of a dialogue game for argumentation as a markov game.

| Markov Game | | Dialogue Game | Dialogue Game (Mod.) |
|---|---|---|---|
| $I$ | $=$ | $\{1,2\}$ | $\{1,2\}$ |
| $s_t$ | $=$ | $(d_t, C_t^1, C_t^2)$ | $(G_t, C_t^1, C_t^2, m_t)$ |
| $A_p(s_t)$ | $=$ | $Pr(d(s_t))$ | $Pr(\Delta(s_t))$ |
| $r_p(s_t, a_t^1, a_t^2)$ | $\in$ | $\{0, 20, -20\}$ | $\{0, 20, -20\}$ |
| $T(s_t, a_t^1, a_t^2)$ | $=$ | $(d_{t+1}, C_{t+1}^1, C_{t+1}^2)$ | $(G_{t+1}, C_{t+1}^1, C_{t+1}^2, m_{t+1})$ |

The latest move $m_t$ is included in the state to enable immediate response protocols. This reduced state space summarizes similar states in the original formalization in one new state. In practice, the graph can be encoded in the respective adjacency matrix to be computationally efficient. As the set of possible actions is currently defined based on the dialogue in the state, it needs to be modified as well.

**Definition 18** (Reduced Action Space)**.** *Let $S$ be the state space as defined above, $M^{<\infty}$ the set of all finite-length dialogues in the dialogue game, $\Delta : S \to M^{<\infty}$ a function that assigns each state a corresponding legal dialogue and $Pr$ the protocol function of the dialogue game. The set of actions for a player $p$ in state $s_t$ is then defined as*

$$A_p(s_t) := Pr(\Delta(s_t)). \tag{5.6}$$

In general, several mappings $\Delta$ are possible, meaning that a single state can be assigned multiple dialogues. However, these dialogues only differ in the temporal order of the included moves and therefore (according to the above-stated condition) share the same set of legal next moves. Finally, the transition function for the reduced state space is defined as follows:

**Definition 19** (Modified Transition Function)**.** *Let $s_t$ be the state at time $t$, $a_t^p$ the action of player $p$ at time $t$ and $m_{t+1}$ the move that corresponds to the action of the player to move. Further, let $\Delta : S \to M^{<\infty}$ be the previously introduced mapping function and $G_{t+1}$ the graph that encodes the dialogue $\Delta(s_t), m_{t+1}$. Then, the modified transition function is given as*

$$T(s_t, a_t^1, a_t^2) := (G_{t+1}, C_{t+1}^1, C_{t+1}^2, m_{t+1}), \tag{5.7}$$

*where $C_{t+1}^p$ denotes the commitment of player $p$ updated with $m_{t+1}$ according to the protocol $P$.*

Although the introduced modifications reduce the dimension of the original state space, additional modifications can be applied in specific dialogue game instantiations based on their individual properties and formal requirements. The complete reformulation is summarized in Table 5.1.

### 5.1.2 Implementation

Next, the introduced reformulation is applied to the dialogue game instantiation described in Section 2.2.2 to evaluate the approach in an actual framework. Again, the setup considered in this thesis for applications end evaluation is utilized. In particular, the logic for defeasible argumentation $\mathscr{L}$ in the dialogue game is based on the general tree structure discussed in Section 4.1.2.

#### Reformulation as Markov Game

Based on the specific setup, additional modifications of the previously introduced definitions are possible: Due to the considered argument (tree) structure, more precisely the unique target of each therein encoded component, arguments have again the form $\varphi_j \Rightarrow \varphi_i$ ($\varphi_j$ supports $\varphi_i$) or $\varphi_j \Rightarrow \neg\varphi_i$ ($\varphi_j$ attacks $\varphi_i$), respectively. As discussed in Section 4.1.3 this means that every speech act $\beta_i \in L_c$ also has one unique relation towards another. In addition, the utilized game protocol does not enforce commitments, meaning that they are not considered in the selection of legal moves and can consequently be omitted from the dialogue state. This allows for an additional reduction of the state space and enables a more efficient optimization.

**Definition 20** (Minimal State Space)**.** *For all speech acts $\beta_j \in L_c$ let $\sigma^t_{\beta_j} \in \{0,1\}$ be a binary integer that is 1 if $d_t$ includes a move with corresponding speech act $\beta_j$ and 0 otherwise. Then the minimal state at time t is given as $s_t = (\sigma^t_{\beta_1}, ..., \sigma^t_{\beta_N})$.*

As a consequence of this definition, the transition function only updates the integer of the speech act utilized in $m_{t+1}$. All other components of the corresponding markov game are defined as discussed in the previous section. Since the protocol of the considered game is not an immediate response protocol, the last game move is not included in this definition of the state space. However, if the corresponding information is required, the state space can be extended to include this information as well.

#### Winning Criteria

The utilized dialogue game for argumentation introduces a winning criterion regulating the termination and the outcome of the game. In there herein considered case, the player who first runs out of moves loses the game. Although this is reasonable from a purely logical point of view, it encourages strategies that are not considered to be optimal by humans. This is best understood in view of the strategy of the opponent, i.e. the player attacking the initial move. If this player challenges in each turn the previous move of the proponent with a *why* reply, the latter one will always run out of moves at a certain point. Nevertheless, as in this case the optimal strategy is known, it is utilized in the proof of principle scenario to test the RL approach.

In order to learn a more natural strategy, a modified winning criterion is proposed. It was first introduced in (Rach et al., 2021c) and is formally derived in the following based on the argument quality taxonomy discussed in Section 2.2.3 (Wachsmuth et al., 2017b). To this end, argument quality dimensions that can be influenced within the boundaries of the dialogue game and thus provide relevant feedback about the underlying argument strategy are identified. As the original definitions assume a single speaker aiming at convincing a (passive) audience, the utilized dimensions are adapted to the scenario at hand. Based on these reformulations, a hierarchy of moves is derived indicating which kind of move is generally to be preferred over alternatives. This preference is afterwards expressed in a numerical order from which a scoring is derived. Winning is then defined based on the difference between the overall scores of both players.

The dimensions that are considered to be of importance for the winning criterion are *effectiveness*, *clarity*, *global relevance* and *global sufficiency*. In the following, the original definitions of the dimensions and their application to the dialogue game are discussed. Subsequently, the scoring is formally defined and an argument for the completeness of the selection is made.

The dimension of *effectiveness* indicates (in its original form) if the speaker is capable of convincing their audience. In the context of the dialogue game, this corresponds to the status of the initial move at the end of the game. If a player has to accept the stance of the opponent, this is an indicator for the effectiveness of the opposing side. *Clarity* is in its original form related to both the language and the focus of the argumentation. Since the language is not addressed within the dialogue game, this dimension can be reduced for this particular context to the focus on the current topic, i.e. branch of the tree structure. *Global relevance* is in parts related to the content of the arguments that are not influenced by the strategy. However, in a dialogue context, it also indicates that players should contribute to the overall resolution, either through arguments or through an agreement on certain aspects. This agreement is represented in the dialogue game through the surrendering moves *concede* and *retract*. The dimension of *global sufficiency* indicates if the speaker adequately addresses counter-arguments, in the context of argumentative speech those that can be anticipated. In the case of an argumentative dialogue, opposing arguments are stated by the opponent, and *global sufficiency* hence demands the interlocutors to adequately rebut the argumentation of the opposing side. For the dialogue game, this includes counter-arguments as well as *why* attacks on their argumentation. Consequently, this dimension is summarized in the requirement that the argumentation strategy is capable of defending attacks with arguments. The following list contains the herein utilized modification of each dimension:

- *Effectiveness*: The argumentation strategy is effective if the opponent concedes to the player's stance of the issue.
- *Clarity*: The player's argumentation strategy has a clear style if it avoids unnecessary complexity and deviation from the issue.
- *Global relevance*: The argumentation strategy is relevant if it contributes to the issue's resolution. This includes arguments as well as a reasonable agreement on otherwise unresolved issues.
- *Global sufficiency*: The argumentation strategy is sufficient if it adequately rebuts attacks of the opposing side.

From this list, a preference of moves is defined next. As *effectiveness* is related to the final stance of the initial move, it is the only one that is not considered in this list. It is instead included in the final score with a constant value $Eff$. According to the dimensions *global relevance* and *global sufficiency*, arguments are to preferred over surrendering moves, which are on the other hand preferred over unresolved issues: $m_{argue} > m_{surrender} > m_{unresolved} \geq 0$. Based on the *clarity* dimension, topic switches are generally negative but no comparison to the other components of the winning criterion can be made: $m_{switch} < 0$. Without loss of generality it is now defined that $m_{unresolved} = 0$ and that each score is to be a natural number, resulting in the following scores: $m_{argue} = 2$, $m_{surrender} = 1$, $m_{switch} = -1$. From this list, and with the assumption that the status of the initial move at the end of the game is in favour of player $p$, the final game score for this player reads as

$$score_p = 2\xi_1 \#m^p_{argue} + \xi_2 \#m^p_{concede} + \xi_3 \#m^p_{retract} - \xi_4 \#m^p_{switch} + Eff\,, \tag{5.8}$$

where $\#m^p_{type}$ denotes the number of *type* moves played by $p$ in the dialogue. The coefficients $\{\xi_j, Eff\}$ can be chosen appropriately to the setting and corresponding argument structure. It can be seen from this scoring that the strategy of constantly challenging is no longer optimal and that it is now possible to win the game without forcing the opposing player to concede to the preferred status of the initial move.

So far, dimensions that influence the winning criterion were discussed. In a final step, it is argued why the presented list is complete in the context of the employed taxonomy: First of all, all allowed moves except for the *why* moves are explicitly addressed by the dimensions. Since challenging an opponent move (i.e. asking for further arguments) is only reasonable in a dialogue context and has no direct counterpart in single side argumentation, this type of move is not represented in a separate dimension. However, it follows from the above scoring that a *why* move is generally preferred over a deviation from the issue. Hence, all move types in the dialogue game are addressed by the proposed list. Also, the multi-reply nature of the game is included by giving a general preference between the options (avoid topic switches). Thus, all aspects of the utilized dialogue game represented in the taxonomy were included in the list. The second argument of completeness is that the remaining dimensions can either be tied to the performance of other modules in an argumentative dialogue system or are already covered in the above list: *Cogency* and all three sub-dimensions are properties of the argument structure, i.e. the arguments and their relations to each other. The same holds for the dimension of *global acceptability*.

The remaining sub-dimensions of *effectiveness* are related to the wording and/or the general presentation, i.e. the question of *how* to introduce an utterance. Therefore, they are directly related to the NLG module and not influenced by the choice of moves in the dialogue game. The only exception here is the dimension of *arrangement* which can be related to the strategy. However, since the rules of the herein utilized dialogue game ensure a certain structure, all dialogues that are allowed in the game can be seen as equivalent in terms of this dimension. Nevertheless, this sub-dimension can be of importance for protocols that are more liberal.

Lastly, the dimension of *reasonableness* is in parts included through the dimensions *global relevance* and *global sufficiency*. The aspects that are not considered herein address again the way utterances are introduced, i.e. the NLG. It should be noted that the dimensions are generally not independent from each other (Wachsmuth et al., 2017a) but the context of artificial argumentation allows to separate them more clearly than in the case of human argumentation. Also, it is stressed that the winning criterion derived above is directly related to the specific game protocol. Nevertheless, the basic idea can be adapted to other dialogue games as well as it is independent of specific protocol rules.

### 5.1.3 Numerical Experiments

In the following, the numerical experiments conducted with the above-discussed implementation are reported. The focus throughout this section is on two major aspects: First, the general markov game approach is evaluated to address the question of whether or not an optimal policy can be found in the considered scenario. Second, the modified winning criterion that was introduced to achieve a more natural and human-like argumentation is discussed by investigating an artificial dialogue between two agents trained on this particular winning criterion.

For both winning criteria, the training was done via Q($\lambda$) and SARSA($\lambda$) with linear function approximation and an $\varepsilon$-greedy strategy, as introduced in Section 2.3.2. It is important to note that the linear approximation generally enables the encoding of knowledge about the optimal strategy in the features. However, this advantage was not explored herein as the purpose of the approach was to show that this knowledge is not required to optimize the strategy.

To enable the efficient use of single-agent learning methods, a setup similar to Silver et al. (2016) was considered: One agent was assigned a fixed policy (*reference policy*) against which the other agent was trained (*training policy*). Both agents started with a random policy and after every $n$ episodes, the *reference policy* was replaced by the current *training policy*. Therefore, the training agent optimized its strategy in each super-iteration against the previously learned one. In the following, a series of $n$ training episodes is denoted as a super-iteration. The stance of each agent was determined randomly at the beginning of each episode to train a policy that is applicable to both roles in the dialogue game.

#### Proof of Concepts

For the evaluation experiment with the original winning criterion, an overall of 40,000 episodes was considered for both algorithms and the reference policy was updated after every $n = 4{,}000$ episodes. To ensure that the outcomes do not depend on one specific argument structure, training was repeated for ten randomly generated argument structures with ten argument components each. The learning rate was set to $\alpha = 0.02$ and the exploration rate to 5%. After each episode, the current policy was evaluated with 0% exploration.

Figure 5.1 shows the thereby achieved average reward of the training agent when assigned the role of the opponent (red) and overall (green) as a function of the number of super-iterations. We see an overall high reward of the training agent against the random policy during the first 4,000 episodes and that this coincides with a comparatively high reward for the role of the opponent in the same range. Afterwards, the overall reward drops to an average around zero whereas the reward for the role of the opponent drops to -20, indicating that both agents win all games in which they are assigned the role of the proponent. Moreover, an improvement to an average of 20 can be observed between 15,000 and 16,000 steps for both curves. Afterwards, the average reward drops again and stays at around zero.

As discussed earlier, a Nash equilibrium means that both strategies are optimal against each other, i.e. changing the policy does not yield an advantage for any side. As both agents start with the same policy every $n$ episodes, an average reward around zero means that no improvement could be achieved by changing this policy further and both sides perform equally well. Nevertheless, this is not a convergence guarantee since the training episodes or the considered amount

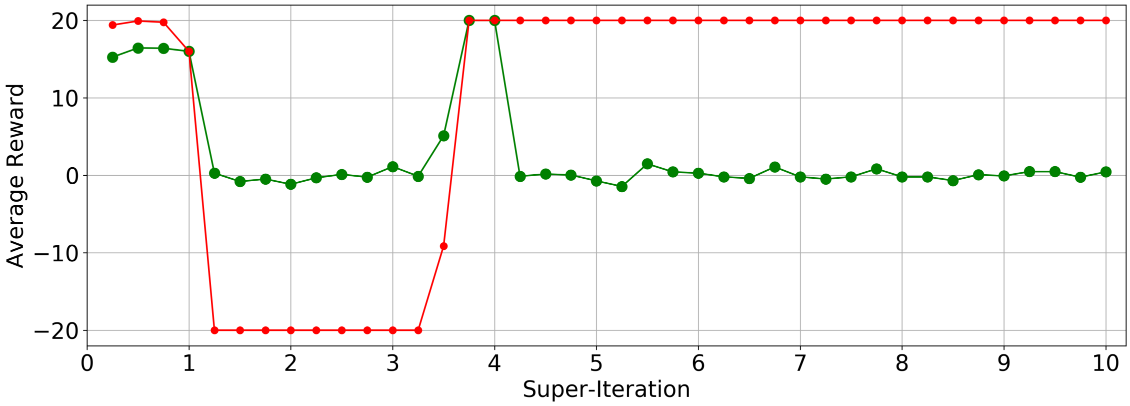


Figure 5.1: Average reward for the training agent as a function of super-iterations when assigned the role of the opponent (average over 500 episodes, red) and overall (average over 1,000 episodes, green). Exploration was set to 0% to derive the plot.

of super-iterations might not be enough to achieve an improvement (as can be seen in Figure 5.1 between 4,000 and 8,000 episodes). However, if the training agent is not able to achieve an improvement over several super-iterations (like in the right half of Figure 5.1), convergence is assumed. This assumption can be tested in the herein considered scenario as it is known that the optimal policy for the opponent always wins independently of the policy of the proponent. Thus, if this (optimal) policy is learned, both agents should be able to win each game in which they are assigned the role of the opponent. This is clearly the case for the right half of the above figure and similar reward curves could be observed for all ten argument structures and both algorithms.

As an additional test, the trained agents played ten games as the opponent against the agent based on probabilistic rules introduced in Chapter 4. The rule-based strategy is semi-optimal in the sense that it does not explicitly take advantage of the argument structure but follows a generally optimal strategy of attacking whenever possible. Since for each investigated argument structure, the trained agent won ten out of ten games once the above-discussed convergence was observed, it is concluded that for this proof-of-principle case, the optimal policy could be found.

### Modified Winning Criterion

For a more application-oriented perspective, training was repeated with the modified winning criterion and the larger argument structure on the topic *Marriage is an outdated institution* (see Section 4.1.2). As the optimal strategy is not known in this case, the results are discussed in terms of algorithm performance and on the basis of natural language dialogues that can be generated with the trained policy, again using the agent-agent setup discussed in Chapter 4. The learning parameters were the same as in the previous scenario, except for the exploration rate which was set to 10%. As the modified winning criterion penalizes topic switches, the state space in Definition 20 was extended to include the information about the latest game move, as discussed in the previous section. The parameters of the winning criterion were set to $\xi_4 = 2$, $\xi_1 = \xi_2 = \xi_3 = 1$ and $Eff = 4$ in the present scenario based on the insights provided by preliminary experiments.

Table 5.2: Example dialogue generated with a policy optimized to the new winning criterion. Sentences that correspond to an annotated argument component are shown in italic font.

| Speaker | Utterance | Speech Act |
|---|---|---|
| Prop. | *Marriage is an outdated institution.* | claim($\varphi_0$) |
| Opp. | *Marriage removes the transient and casual aspects of a monogamous relationship, thus giving a child a far more stable environment.* | argue($\varphi_1 \Rightarrow \neg\varphi_0$) |
| Prop. | *Marriage is redundant in terms of raising children.* | argue($\varphi_2 \Rightarrow \neg\varphi_1$) |
| Opp. | I'm not sure I understand what you're getting at. | why($\varphi_2$) |
| Prop. | *Marriage offers no more stability than a stable relationship.* | argue($\varphi_3 \Rightarrow \varphi_2$) |
| Opp. | Could you be more specific? | why($\varphi_3$) |
| Prop. | I retract that statement. Let's assume I never said that.<br>I'm still thinking about something you mentioned earlier. You said: *Marriage removes the transient and casual aspects of a monogamous relationship, thus giving a child a far more stable environment. Marriage clearly does not offer the stability you claim it does.* | retract($\varphi_3$)<br>argue($\varphi_4 \Rightarrow \neg\varphi_1$) |
| Opp. | Could you please elaborate? | why($\varphi_3$) |
| Prop. | *40% of marriages end in divorce and this is on the rise according to the National Office for Statistics in 1999.* | argue($\varphi_4 \Rightarrow \varphi_3$) |
| Opp. | Could you be more specific? | why($\varphi_4$) |
| Prop. | I think I was wrong there.<br>You said earlier that *marriage removes the transient and casual aspects of a monogamous relationship, thus giving a child a far more stable environment.* I agree with you. | retract($\varphi_4$)<br>concede($\varphi_1$) |

These preliminary experiments also indicated a high complexity of the learning problem and hence a computationally expensive training in the discussed setup. Consequently, several modifications were made to facilitate learning: First, the properties of the new winning criterion were exploited to provide agents with an immediate reward during the interaction based on the score of the selected moves. In addition, the training in each super-iteration was stopped once the training agent was able to win 20 games in a row (without exploration) against its counterpart, thereby assuming that a winning strategy against the current reference policy was found. If the training agent was not able to fulfil this termination criterion for two super-iterations (10,000 episodes per super-iteration) in a row, convergence was assumed.

However, this is no guarantee for the convergence to a Nash equilibrium, especially in view of the much larger state space in this case. An example dialogue[1] achieved within twelve super-iterations and the SARSA($\lambda$) algorithm is shown in Table 5.2. It can be seen that both sides try to stick to the current line of argumentation (i.e. try to avoid a topic switch) as long as possible.

[1] Material reproduced from www.idebate.org with the permission of the International Debating Education Association. 

In addition, the proponent clearly prefers *argue* moves over *why*, whereas the opponent strategy shows (as in Chapter 4) a frequent use of *why* replies. However, it is important to note that the argument structure does not provide a counter-argument for every included argument and the proponent in fact only selects arguments for which the opponent has no argumentative response available. This indicates that the trained strategy not just learned preferences between move types but also explores the problem-specific argument structure to get an advantage. Nevertheless, the explicit reply structure that demands a request for further information before support arguments can be introduced is – as already pointed out in Chapter 4 – an eminent limitation of the utilized game and is addressed in the next section.

In addition, the topic switch in the fourth turn of the proponent is accompanied by a *retract* to the previous move. This makes the topic switch appear more natural but on the other side limits the possibilities of the proponent in later stages of the dialogue. Moreover, the multi-move aspect of the employed dialogue game for argumentation can be observed, as the fourth and the sixth proponent turns includes more than one move. Regarding the winning criterion, the proponent finishes the game with a score of nine, whereas the opponent only provided one argument and hence has a score of two. Consequently, the proponent wins this game, even though the final agreement on the initial move is in favour of the opponent.

From a technical perspective convergence to different final scores in different learning runs as well as a remaining high computational complexity is reported. In addition, some of the strategies observed during training are clearly not optimal. This is a known problem of the utilized learning methods when applied to multi-agent scenarios (Littman, 1994). It is thus fair to say that the employed algorithms gradually reach their limits in these large scale cases and should be replaced by more advanced approaches in the future.

Nevertheless, this appears to be more of a technical than a conceptual issue and it can be seen that the above-shown dialogue is quite complex. The underlying strategies show the characteristics considered to be important by the winning criterion and it is concluded that the agents are indeed able to learn the respective concepts. As for the winning criterion, it can be seen that the goal of providing a more fine-grained assessment of the strategies could be accomplished as it enables agents to agree with the stance of the opposing side while still winning the overall game. In addition, it provides the possibility of immediate rewards which can be used to facilitate learning.

### 5.1.4 Discussion

In this section, it was shown how dialogue games for argumentation adhering to a general structure can be reformulated as markov games and be addressed as a multi-agent RL task. The approach was tested on an implementation based on a specific dialogue game for argumentation and random argument structures. The results indicate that the optimal policy was found in each considered case. Although this strategy is only optimal with respect to the game framework and possibly perceived as unnatural, its successful learning demonstrates the feasibility of the approach. In addition, it indicates that the proposed approach can also be utilized to assess game frameworks as it reveals (possibly unintended) strategies that result from its formal definition.

With the aim of applying the proposed methods in application systems, a modified winning criterion based on argument quality dimensions was introduced and utilized for training a policy in the case of a larger argument structure. The outcome was discussed with respect to algorithm performance and on the basis of an artificial dialogue generated from trained policies, showing the increased complexity of the strategies that reflect the considerations on which the winning criterion is built. However, the limitations of the linear function approximation utilized in the employed learning algorithms could also be observed in this case, indicating that more complex methods will be required in practice.

To address this issue, the use of deep reinforcement learning methods was investigated in the scope of a master's thesis (Yang, 2019) for the above discussed proof-of-principle setup. The corresponding results are discussed in detail in Chapter 8 and indicate that the utilized method is capable of dealing with the increasing state space that results from larger argument structures although the problem of convergence to local optima was observed in this case as well. However, the multi-agent setup allows for comparing policies against each other and a practical approach is hence to train multiple strategies simultaneously and select the best one from a subsequent comparison. An additional limitation for applications is the dependency of the strategy optimization on a purely formal winning criterion, which does not consider the perception of an individual user. Moreover, the strict limitations of the utilized dialogue game regarding supporting arguments also hinders the learning of more complex strategies.

The identified issues are addressed in the remainder of the thesis through different approaches: The strict limitations of the utilized dialogue game is discussed in the following section, where modifications of the game formalism to enable more natural dialogues are introduced. The application of the proposed learning approach is discussed in Chapter 8 on different levels of complexity. In particular, a summarizing system is introduced that utilizes the modified dialogue game, a policy trained through deep reinforcement learning and an approach to further adapt the utterance selection to individual user feedback.

## 5.2 Liberal Game Protocol

Within this section, the previously identified issue of the perceived naturalness of dialogues generated with the herein utilized dialogue game is addressed. Despite the formal advantage of dialogue games, the resulting interactions are restricted by the formalism. In the example of the herein investigated case, this limitation leads to a frequent use of *why* replies. The task at hand is thus to find the balance between reasonable restrictions and a freedom of choice that enables a natural and intuitive interaction. The difficulty lies in the implications that come with this freedom, as the possibility of a more natural response may also include the possibility of responses that are neither natural nor consistent and violate the basic principles of the desired interaction. Especially modifications to an established formalism have thus to preserve the general properties of the model and extend the respective regulations rather than simply omit them.

In the following, this task is addressed for the herein utilized dialogue game instantiation. As argued in Chapter 4, the utilized model allows for multiple as well as postponed responses to an utterance, meaning that it already provides a certain freedom of choices for the players by design. The proposed approach increases this freedom by modifying the underlying game protocol to allow for the chaining of multiple arguments. This is done in compliance with the remaining regulations and does thus not violate the logical consistency of the resulting dialogues. The approach is tested by generating artificial dialogues between two virtual agents. The corresponding

Table 5.3: Communication language $L_c$ of the modified dialogue game with $\Phi_i, \Phi_j \in Args$, $\varphi_i, \varphi_j \in L_t$ and $\Phi_j \rightarrow \Phi_i$ the notion for $\Phi_j$ defeats $\Phi_i$.

| Speech Act | Attacks | Surrenders |
|---|---|---|
| *claim*($\varphi_i$) | *why*($\varphi_i$) | *concede*($\varphi_i$) |
| *why*($\varphi_i$) | *argue*($\Phi_j$), *argue_extend*($\Phi_j$) ($conc(\Phi_j) = \varphi_i$) | *retract*($\varphi_i$) |
| *concede*($\varphi_i$) | - | - |
| *retract*($\varphi_i$) | - | - |
| *argue*($\Phi_i$) | *why*($\varphi_i$) ($\varphi_i \in prem(\Phi_i)$), *argue*($\Phi_j$), *argue_extend*($\Phi_j$) ($\Phi_j \rightarrow \Phi_i$) | *concede*($\varphi_i$) ($\varphi_i = conc(\Phi_i)$ or $\varphi_i \in prem(\Phi_i)$) |
| *argue_extend*($\Phi_i$) | *why*($\varphi_i$) ($\varphi_i \in prem(\Phi_i)$), *argue*($\Phi_j$), *argue_extend*($\Phi_j$) ($\Phi_j \rightarrow \Phi_i$) | *concede*($\varphi_i$) ($\varphi_i = conc(\Phi_i)$ or $\varphi_i \in prem(\Phi_i)$) |

dialogue transcripts are then rated in a user study with respect to their logical consistency as well as their naturalness. The survey setup is adapted from Chapter 4, where the unmodified version of the dialogue game was applied. The results show that the perceived consistency remains whereas the naturalness is significantly increased. The work in this section is published in (Rach et al., 2020b).

### 5.2.1 Extension to Chained Arguments

The main restriction of the utilized formalism with respect to naturalness lies in the inability of introducing more than one argument per turn. More precisely, a player is only allowed to extend an argument, if the corresponding move was challenged by a *why* move. This section introduces an extension that allows players to chain multiple arguments in a single turn without violating the logical consistency of the dialogue. To this end, an additional speech act and modifications to the protocol are introduced.

In the original formalism, a player has to move until he or she switched the status of the initial move in his or her favour, meaning that he or she plays an unspecified number of surrendering moves, followed by a single attack. After the status of the initial move is switched, the turn ends immediately. This rule is now modified by allowing both players to *extend* their attack under the condition, that the attack includes a new argument. An extended attack generally allows the player to introduce additional arguments to defend his or her current move in advance. This extension does thus not reply to an actual attack but to an anticipated one.

The first formal modification is the introduction of an additional speech act *argue_extend*($\Phi$) to $L_c$, as shown in Table 5.3. It can be seen that the new speech act has the same properties as the *argue*($\Phi$) act concerning attacking and surrendering replies. Next, the modifications and extensions of the game formalism to enable the chaining of arguments are defined, again for the continuation of a dialogue $d_{k-1}, m_k$. The original definition of a dialogue demands that $m_k$ replies to a move in $d_{k-1}$ unless $d_{k-1} = d_0$. This definition is adapted so that $m_k$ replies to a move in $d_{k-1}$ unless $d_{k-1} = d_0$ or $m_{k-1}$ has the type *argue_extend*. Moreover, the following adaptations of the protocol rules are introduced:

- $P_6$ (original): $T(d_0) = proponent$ and $T(d_{k-1}) = p, p \in \mathscr{P}$ if $p$ is not the current winner of $d_{k-1}$ and $d_{k-1} \neq d_0$.
- $P_6'$ (modified): $T(d_0) = proponent$ and $T(d_{k-1}) = p, p \in \mathscr{P}$ if $p$ is not the current winner of $d_{k-1}$ and $d_{k-1} \neq d_0$ unless $m_{k-1}$ has the type *argue_extend*. If $m_{k-1}$ has the type $argue\_extend$, $T(d_{k-1}) = p_{k-1}$.

The modification of $P_6$ defines that the turn of a player is continued if the current move includes an *argue_extend* move, even if this move makes the player the current winner of the game. In this case, $m_k$ is said to *anticipate* an attacking reply. The next modification concerns the opening move and allows the proponent to open the game with an *argue_extend* move:

- $P_7$ (original): If $d_{k-1} = d_0$, then $\beta_k$ has the type *argue* or *claim*.
- $P_7'$ (modified): If $d_{k-1} = d_0$, then $\beta_k$ has the type *argue*, *argue_extend* or *claim*.

To ensure a coherent and reasonable interaction, a set of additional rules is required that regulate the response to an anticipated attack within the game framework:

- $P_9$: If $\beta_k = argue\_extend(\Phi_i)$, then $\exists \Phi_j \in Args$ that extends $\Phi_i$.
- $P_{10}$: If $\beta_{k-1} = argue\_extend(\Phi_i)$, then $\beta_k = argue\_extend(\Phi_j)$ or $\beta_k = argue(\Phi_j)$, with $\Phi_j$ extends $\Phi_i$ and $\tau_k = 0$.
- $P_{11}$: If $m_k$ responds to $m_i$ with $\beta_k = why(\varphi_j)$, $m_i$ is of the type *argue_extend* and $\beta_{i+1} = argue(\Phi_l)$ or $\beta_{i+1} = argue\_extend(\Phi_l)$, then $\tau_{i+1} = k$.

$P_9$ defines that a move including the speech act *argue_extend* can only be played if an extending argument is available in $\mathscr{L}$. This ensures that an attack is only extended if additional information is available. $P_{10}$ demands that the extending argument has to be introduced to the dialogue directly after an attacking reply was anticipated. In addition, it defines that the extension itself can also be extended and that all extending moves have an empty target, as they do not explicitly reply to a move in $d_k$. The empty target is required to preserve the restriction of the communication language to attacking and surrendering replies. Finally, $P_{11}$ says that if a player decides to play a previously anticipated attack, the corresponding extension replies to this attack in $d_k$. As a direct result of this rule, the anticipated attack is assigned the status *out* and requires the corresponding player to play an additional move.

In the following, implications and changes in the game that result from this modification are discussed. When introduced, each move in a chain is *in*. The first move in a chain is also a relevant target since a (successful) attack on it changes the status of the initial move. The remaining moves on the other hand are no relevant targets as an attacking reply to them does not affect the status of the initial move. Moreover, challenging the relevant target in a chain only switches the status of the initial move if this challenge is not anticipated in the chain. Otherwise, the responding move in the chain becomes a relevant target and the current player is obliged to play another move.

An example of a chain consisting of two moves, including status and relevance is shown in Figure 5.2. Successful attacking replies to a chain can have multiple forms. They are illustrated at abstract example dialogues with three turns ($t_1$, $t_2$ and $t_3$) in Figure 5.3 and discussed in detail in the following.

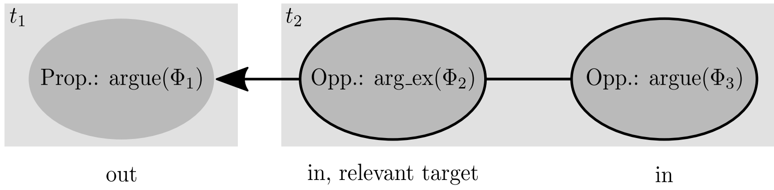


Figure 5.2: Illustration of a chain consisting of one *argue_extend* and one *argue* move. Grey boxes indicate the corresponding turns ($t_1$ and $t_2$) in the game. Both moves in the chain are *in* (indicated by black margins) but only the first one is a relevant target.

It can be seen that possible responses are divided into three groups:

- A chain can be attacked by a series of anticipated *why* moves, followed by a *why* move that is not anticipated (Response 1).
- A chain can be attacked by a combination of anticipated *why* and *argue(_extend)* moves (Response 2).
- A chain can be attacked by an attacking reply to its first move if this reply is not anticipated in the chain (Response 3).

Generally, responses to a chain may also include a (new) chain, thus giving the players far more freedom in their choices. Nevertheless, since the legality of moves is still determined by the same principles as in the original framework, the resulting dialogues have the same formal consistency.

### 5.2.2 Evaluation

To evaluate the discussed extensions, artificial dialogues between two virtual agents Alice and Bob are generated and assessed in a user study. The evaluation setup is chosen in order to compare the results directly with the ratings for the original framework discussed in Chapter 4. To ensure a fair comparison, the setup is as similar as possible to the original one. Thus, the same multi-agent setup is employed, including the same argument structure, a similar rule-based agent strategy and a similar natural language generation (NLG). In addition, the same questionnaire is used for the survey. The relevant details regarding the setup are recalled in the context of the modified framework in the following.

#### Multi-Agent Setup

The utilized argument structure is composed of 72 annotated argument components and their relations derived from the transcript of a debate with the topic *Marriage is an outdated institution* (see Chapter 4). Consequently, arguments again have the form $\Phi_j = (\varphi_j \Rightarrow \varphi_i)$ if $\varphi_j$ supports $\varphi_i$ and $\Phi_j = (\varphi_j \Rightarrow \neg\varphi_i)$ if $\varphi_j$ attacks $\varphi_i$.

During the interaction, the agents select their next move from the list of available options provided by the dialogue game. In order to ensure a competitive but reasonable strategy, each

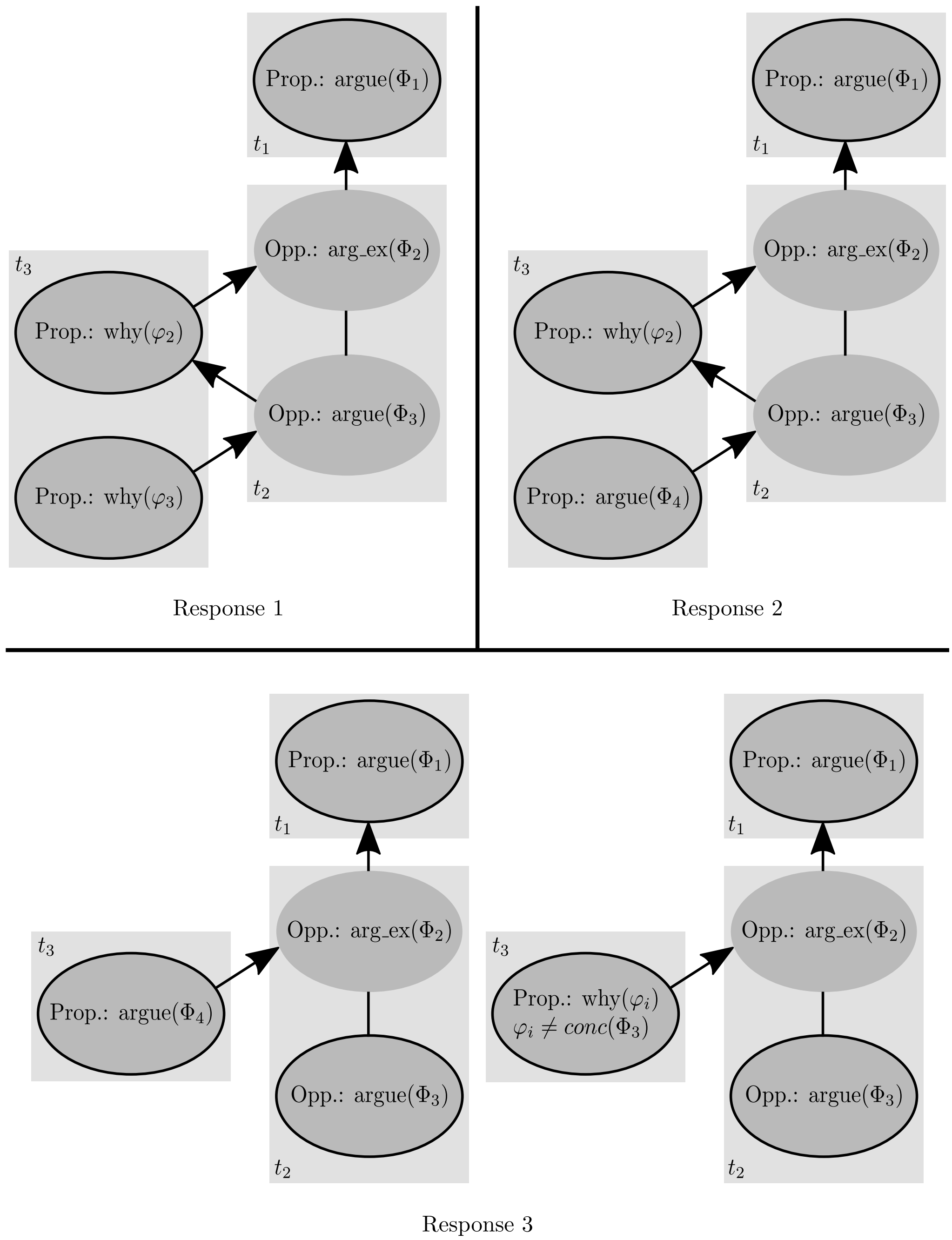


Figure 5.3: Illustration of three possible attacking replies to a chain. The arrow from *argue*$(\Phi_3)$ to *why*$(\varphi_2)$ in Response 1 and 2 indicates the implicit response of the argument chain to the anticipated attack.

Table 5.4: Artificial dialogue between the agents Alice and Bob with the modified dialogue game. Italic parts in the utterances refer to the annotated sentences in the argument structure.

| Player | Utterance | Speech Acts |
|---|---|---|
| Alice | *Marriage is an outdated institution.* | claim($\varphi_0$) |
| Bob | In my opinion *the remarriage rate shows that even people who go through failed marriages retain faith in the institution of marriage.* | argue_extend($\varphi_1 \Rightarrow \neg\varphi_0$) |
| | I would like to go into that a little further. You see, *there are still such huge numbers of people who practice religions to which marriage is integral.* | argue($\varphi_2 \Rightarrow \varphi_1$) |
| Alice | Unfortunately I didn't find that entirely convincing. Would you mind elaborating a little further? | why($\varphi_1$) |
| | In particular, there's one aspect of your argumentation that I have some doubts about. You said that *there are still such huge numbers of people who practice religions to which marriage is integral.* It seems to me that *religion as a whole is becoming less important and, with it, marriage is becoming less important.* | argue($\varphi_3 \Rightarrow \neg\varphi_1$) |
| Bob | Well, maybe we can extend this line of argumentation later. But if you don't mind, let me add something to our overall topic. It seems to me that *marriage is an important institution to religious people.* | argue_extend($\varphi_4 \Rightarrow \neg\varphi_0$) |
| | I would like to go into that a little further. You see, *there are still such huge numbers of people who practice religions to which marriage is integral.* | argue_extend($\varphi_5 \Rightarrow \varphi_4$) |
| | For example *nearly 50% of people in the UK identify as being part of some religion according to the British Social Attitudes Survey of 2007* | argue($\varphi_6 \Rightarrow \neg\varphi_5$) |

agent first prefers attacking moves over surrendering moves, then *argue(_extend)* moves over *why* moves and finally immediate to postponed responses. If there are multiple options with the same preference, the selection between them is random. Moreover, both agents extend their line of argumentation as long as possible. As in the original setup, an *argue*($\varphi_i, \Rightarrow \varphi_0$) attacking reply to *claim*($\varphi_0$) is allowed to cover all available arguments. This attack can also be extended in the modified framework.

The NLG is done turn wise, meaning that all moves of a turn are merged into one utterance. As in the original setup, the natural language representation of arguments is gained from the annotated sentences of the argument components. Postponed *argue(_extend)* replies include the premise and the conclusion. In the case of a direct reply, the conclusion is left implicit. For the remaining moves, a list of templates is used from which the system selects randomly. Again, the natural language representation of the argument component in the move is left implicit for direct replies and explicitly included in the case of a postponed reply. In addition, a new list of connecting and opening phrases is generated in order to concatenate multiple moves into a single utterance. This part of the NLG is an extension to the original version and may influence the user perception of the resulting dialogues. However, since this extension is only possible due to the extended framework, the advanced NLG template is a direct result of the formal extensions. An excerpt from an example dialogue[2] is shown in Table 5.4, including speakers, speech acts and utterances.

#### User Study

To compare the proposed approach with the original framework, ten virtual discussions between the agents Alice (proponent) and Bob (opponent) were generated with the new framework. The corresponding transcripts were then evaluated in a user survey with the same study setup as in the referenced work. In the original case, 20 dialogues were required to cover a majority of the available arguments. This was mainly due to the extensive use of isolated *why* moves. As those are merged into a single utterance within the modified framework, ten dialogues are sufficient to present a similar amount of arguments.

The questionnaire consists of ten statements related to the strategy, the line of argumentation and the naturalness of the dialogue. Each statement was rated on a five-point Likert scale from *completely disagree* (1) to *completely agree* (5) by 61 participants from the UK with an age between 18 and 99. The survey was realized by clickworker[3] and each participant was assigned a single randomly selected discussion in order to avoid a bias. The wording of statements that are relevant for the herein discussed topic together with the corresponding median (original and modified framework) as well as the $p$ value achieved with a Mann-Whitney-U test are shown in Table 5.5. It can be seen that the four statements related to the logical consistency of the argumentation show no significant difference to the original results, whereas the $p$ value for both statements related to the naturalness is below the threshold of 0.05. For the sake of completeness, it is reported that no significant difference was found for the statements omitted in Table 5.5. This indicates that the introduced modifications significantly improve the perceived naturalness of the resulting dialogues without lowering the consistency.

### 5.2.3 Discussion

The introduced modification of the investigated dialogue game for argumentation allows for chained arguments from both players while preserving the relevance criterion that ensures consistency. As a consequence, the modifications are not limited to the investigated framework of *relevant dialogues* and can be combined with other extensions as well. Examples are the framework for *weakly relevant* dialogues and the inclusion of commitment rules discussed in (Prakken, 2005). In particular, the reformulation of dialogue games as markov game discussed in Section 5.1 is directly applicable, although a relaxed version of dialogues $d_k$ is used in the modified framework. However, as the reformulation as markov game does not depend on this definition, the corresponding reformulation is analogous to the original framework. The application of reinforcement learning in the modified framework is discussed in Chapter 8, where the framework is utilized in a complete multi-agent system.

[2] Material reproduced from www.idebate.org with the permission of the International Debating Education Association. Copyright ©2005 International Debate Education Association. All Rights Reserved

[3] https://marketplace.clickworker.com (last accessed 29 August 2021)

Table 5.5: Results for the original framework (Original) and extended one (Modified). Bold lines indicate a significant difference.

| Statement | Original | Modified | p |
|---|---|---|---|
| The arguments presented by Bob are logically consistent responses to the utterances they refer to. | 4.0 | 4.0 | 0.36 |
| The arguments presented by Alice are logically consistent responses to the utterances they refer to. | 3.5 | 3.0 | 0.81 |
| Bob's line of argumentation is not logically consistent. | 2.0 | 2.0 | 0.85 |
| Alice's line of argumentation is not logically consistent. | 2.0 | 2.0 | 0.74 |
| **It was difficult to follow the line of argumentation throughout the debate.** | **3.0** | **2.0** | **0.02** |
| **The whole debate is natural and intuitive.** | **2.0** | **4.0** | **0.02** |

For evaluation purposes, artificial discussions between two virtual agents were generated. The corresponding dialogue transcripts were then rated in a user study to confirm the desired effect of the introduced modifications. In a direct comparison to the results achieved with the unmodified dialogue game (see Chapter 4), a significant improvement in ratings related to the naturalness and intuitiveness of the dialogue is observed. On the other hand, the perceived consistency of the dialogue does not change significantly. It can hence be concluded that the proposed modifications result in the intended effect, i.e. improve the naturalness of the resulting dialogues.

## 5.3 Conclusion

This chapter has addressed the development of challenging argumentation strategies (or policies) in the context of dialogue games for argumentation. As the possible strategies directly depend on the utilized framework, their development was discussed alongside framework properties and modifications for the herein considered dialogue game instantiation.

For the strategy, a general reformulation of the dialogue game formalism as a markov game was proposed. It allows for an optimization of the corresponding agent policy within the game via multi-agent reinforcement learning. The feasibility of the approach was validated in a proof-of-principle setup with a known optimal policy. In addition, a modified winning criterion based on argument quality dimensions was introduced for the investigated dialogue game instantiation to enable the learning of a more complex strategy. Moreover, the previously identified issue of the naturalness of agent-agent dialogues generated with the utilized framework was addressed through modifications of the game formalism.

It can be seen that the individual contributions enable flexibility on different levels: The proposed approach of multi-agent reinforcement learning depends only on the dialogue game formalism and can hence be applied to any dialogue model that can be expressed in terms of the corresponding formal definitions. In particular, it is not limited to a specific dialogue game instantiation or topic-specific argument structure. From a practical perspective, it was observed that the utilized learning methods gradually reach their limits when applied to larger argument structures. Consequently, more complex methods such as deep actor-critic algorithms are necessary for an application in actual systems. However, recent research in the field of reinforcement learning has yielded substantial progress in providing methods for large-scale problems as for example the approach introduced in (Wang et al., 2016). Therefore, a variety of algorithms are available to overcome this limitation. The use of a deep reinforcement learning algorithm in the context of the investigated framework is discussed in more detail in Chapter 8, where the proposed approaches are integrated into a multi-agent system.

In contrast, the proposed modifications for the utilized dialogue game – winning criterion and argument chains – depend directly on the utilized formalism. Although similar concepts can be applied to other dialogue game instantiations as well, the respective modifications need to be developed in the context of the individual framework to preserve the desired properties (in the present case the relevance criterion). However, the evaluation results indicate that in the investigated case, the desired effect of increasing the naturalness of dialogues generated with the framework could be accomplished. It is hence concluded that the herein proposed methods contribute to the development of challenging agent strategies in two ways: First, by introducing a general approach to policy optimization without dependencies on conversational training data, pre-defined opponent strategies or specific dialogue game instantiations and second by addressing pending issues in the specific investigated setup. The second contribution makes the investigated framework more applicable in actual systems and also enables the use of the proposed optimization approach in this system.

# 6 Enabling Topic Flexibility in Argumentative Dialogue Systems

This chapter addresses the second major aspect of the thesis, i.e. the flexibility with regard to the topics that can be discussed by an argumentative conversational agent. There are two main challenges in this regard: The first one is the acquisition of arguments that are suitable for the task, i.e. the question of where the arguments come from. The second one originates from the complexity of the task for conversational agents. It leads to a dependency of many systems on structured knowledge about available arguments and relations between them. This can hence be summarized in the question of how the individual arguments relate to each other. Whereas the first challenge can be addressed in a general manner and mostly independent of specific system requirements, the second challenge depends (to a certain extend) on the system and the therein utilized argumentation framework. However, this separation is mostly on a conceptual level, as many technical approaches address both aspects or at least parts of both aspects jointly. Throughout this chapter, the (modified) formal framework for persuasive dialogues discussed in Chapter 5.2 is used as a reference, whenever system requirements have to be taken into account. Nevertheless, the overall goal is to keep the proposed approaches flexible in this regard.

Two separate approaches are investigated herein. They differ not just in their technical realization but also in the kind of topics they address. In order to distinguish both approaches conceptually, the topics for which arguments are acquired are divided into the two broad groups *opinion* and *non-opinion* based. *Opinion based* topics can also be referred to as controversial topics and cover general subjects on which most people already hold a certain opinion. Examples are political issues like Brexit but also social topics like minimum wage or nuclear energy. The appeal of this kind of topic for conversational agents lies in their subjectivity, which is at the core of argumentation. More precisely, the result of an argumentative interaction concerned with such a topic depends directly on the knowledge of the interlocutors and their ability to defend their point of view. It therefore poses an interesting and challenging task for argumentative dialogue systems and artificial intelligence in general. To acquire arguments for this kind of topics, the applicability of argument search engines as a general approach to argument retrieval in dialogue systems is investigated. Subsequently, a mapping of the retrieved arguments into a tree structure that is compatible with the requirements of the aforementioned formal framework for persuasive dialogues is proposed.

*Non-opinion based* topics on the other hand are topics to which most people have a neutral position. The motivation for the investigation of this topic group comes from an evaluation perspective, as for example the persuasive effectiveness of a system is hard to assess in scenarios where test subjects have a strong opinion towards the discussed topic (i.e. a bias). To this end, the acquisition of arguments from opinions towards hotels and restaurants expressed in reviews is investigated. Given that a user or study participant is not familiar with the discussed establishment, a neutral position prior to the experiment can be assumed. Technically, this is approached through a semi-automatic procedure that generates argument structures (i.e. retrieves arguments and their relation towards each other) from reviews annotated for sentiment analysis. The inference of the structure is completely automatic and combined with manual corrections of mainly the annotated sentences. The manual step is required to ensure high-quality structures that are necessary for a successful evaluation.

The chapter addresses *opinion based* topics first, starting with the evaluation of argument search engines in the context of argumentative dialogue systems. The reported work was published in (Rach et al., 2020a) in cooperation with the TU Darmstadt and during the research period at NAIST. Subsequently, the mapping of arguments retrieved by an argument search engine into a tree structure is discussed and evaluated, followed by an extensive user study in which the complete pipeline from argument search to argumentative dialogue is evaluated. This part of the work was also conducted in cooperation with TU Darmstadt, builds on the results of a bachelor's thesis (Schindler, 2020) and is published in (Rach et al., 2021b). The final part of the chapter addresses the extraction of argument structures from reviews and is published in (Weber et al., 2020a) in cooperation with Augsburg University.

As for notation, a single search result from an argument search engine is denoted as *argument* throughout this chapter and its polarity regarding the overall topic as *stance* (pro or con).

## 6.1 Argument Search for Argumentative Dialogue Systems

As indicated by the discussion of related work in Chapter 3, most argumentative systems so far operate on a carefully but also strictly designed argument database that perfectly matches their requirements. However, to increase their flexibility with respect to the range of discussed topics, automatized and topic independent approaches to acquire arguments are required. Argument search engines (Ajjour et al., 2019) on the other hand have recently emerged from the field of argument mining (Lawrence and Reed, 2020; Lippi and Torroni, 2016) and provide users with a ranked list of arguments corresponding to a given search query. Hence, they are of particular interest for argumentative dialogue systems as they allow to search a wide variety of sources for arguments and are not restricted to specific topics. However, in order to utilize argument search techniques for argumentative dialogue systems, certain quality standards for the retrieved arguments are crucial.

In a cooperative work (Rach et al., 2020a) with TU Darmstadt conducted at NAIST, we investigate these quality aspects by means of an argumentative dialogue system that evaluates arguments retrieved by different search approaches in direct interaction with users. This is realized by enabling the user to give specific ratings in the categories *interesting*, *convincing*, *comprehensible* and *related* as direct feedback to each system utterance. To ensure a setting that is representative of dialogue system applications, the arguments are presented by means of a virtual avatar and synthetic speech. The approach is motivated by the difficulty of argument quality assessment from a purely logical perspective (Habernal and Gurevych, 2016b; Wachsmuth et al., 2017a) as well as the common approach to evaluating dialogue systems from the user perspective (Deriu et al., 2020). Especially the subjective nature of argumentation and the effects of system modalities (virtual avatar and synthetic speech) in the present scenario render approaches that do not explicitly consider the user perception impractical. The proposed setup is applied in a user study to compare two state-of-the-art argument search engines, namely ArgumenText (Stab et al., 2018a) and args.me (Wachsmuth et al., 2017c), to each other. In addition, we introduce an argument retrieval system based on conventional web search to provide a suitable baseline. In order to exclude topic dependencies, the comparison is done over three different controversial topics. The results show significant differences between the investigated approaches for three of the four categories and both search engines outperform the baseline in one category. In addition, both search engines outperform each other in a different category, thereby reflecting the different strengths and drawbacks of the underlying technological approaches.

### 6.1.1 Evaluation Criteria

As a first step in developing the actual evaluation setup, we introduce the criteria utilized throughout the evaluation process to assess the presented arguments. Their definition is driven by the goal of providing a set of evaluation dimensions suitable for an assessment during the interaction as well as general enough to provide valuable insights for the application of the compared approaches in argumentative dialogue systems. Therefore, we introduce four categories that cover the following quality aspects:

- The structural argument properties that are influenced by the different technological approaches (i.e. identification of arguments and stance) of the search engines.
- The suitability of the retrieved arguments for the different tasks of an argumentative system.

To address the first aspect, we rely on *argument quality* dimensions that are strongly influenced by the technological differences between the search engines. For the second aspect, we start from the general notion of *task success* as a common approach to assessing task-oriented dialogue systems. Since the success of an argumentative dialogue system depends on the individual user and is therefore hard to measure objectively, we identify argument properties that facilitate the completion of possible tasks and assess them in separate categories. The following subsections provide a detailed discussion of both aspects and include the notion of the categories utilized within our system.

#### Argument Quality Related Criteria

Based on the taxonomy discussed in Section 2.2.3, overall argument quality includes the three dimensions of *logical*, *rhetorical*, and *dialectical* quality. *Logical* quality is related to the structure of an argument, i.e. the question if an argument is logically sound. *Rhetorical* quality is reflected in the persuasive effect of an argument and hence also depends on aspects like the presentation of the argument and the credibility of the person presenting it. Lastly, *dialectical* quality is related to the contribution of an argument to the resolution of an issue or different opinions on a topic.

The task of an argument search engine includes the mining of arguments from relevant sources as well as the recognition of the respective stance (Ajjour et al., 2019). Errors result in a system output that is not perceived as argumentative, not related to the topic or presented with the wrong relation to the topic, which is all part of the logical quality dimension. In order to distinguish between the different errors, we assess the logical quality of the retrieved arguments with the two evaluation categories *comprehensible* (Does the argument make sense by itself?) and *related* (Is the presented argument related to the topic and is the presented relation correct?). The first category is binary, whereas the second one allows a choice among one positive and two negative options (related, not related, wrong relation) to enable a distinction between the different errors that can occur.

#### Task Related Criteria

The task success rate is a common evaluation criterion for task-oriented dialogue systems (Deriu et al., 2020) and measures to which extend and how often a system provides correct information/responses to user requests. However, in contrast to conventional setups like providing information in the restaurant domain (Schatzmann et al., 2007), the system output of an argumentative dialogue system cannot clearly be divided into right and wrong responses. Nevertheless, the capability of a dialogue system to solve an argumentative task depends on whether it is able to select suitable arguments and hence on the output of the respective retrieval system. Based on the types of argumentative dialogue (Reed and Norman, 2003), the different tasks of argumentative dialogue systems can be broadly divided into competitive and cooperative tasks.

In competitive tasks like persuasion or negotiation, the overall goal is to *convince* the opposite site of for example a certain point of view (persuasion) or to accept a specific offer (negotiation). Consequently, the relevant property of the utilized arguments is their overall likeliness to convince the opponent, in short, their *convincingness*. The respective evaluation category in our setup is hence *convincing* (Does the argument convince me?) for competitive setups.

Cooperative setups (for example deliberation) on the other hand aim for a mutual solution of an issue by exchanging arguments that contribute to this task. In contrast to competitive setups, the goal of the involved parties is not to convince the other participants of a certain point of view but to find the best common ground. Therefore, the suitability of an argument for these tasks depends on its ability to contribute to this solution. However, in an argumentative application, this common ground should satisfy the user's needs and hence depends on the user perspective on the presented arguments. Consequently, we condensate this property in the question if an argument is *interesting* for the user with respect to the discussed topic and assess it by means of a category with the same name.

Both of these categories are also related to argument quality: the *convincingness* of an argument is an aspect of rhetorical quality whereas *interesting* reflects the user's personal view on the overall relevance of an argument and is hence related to dialectical quality. Nevertheless, each category only covers a part of the respective quality dimension, as both dimensions can be further influenced by other modules of the argumentative dialogue system (behaviour of the avatar, natural language generation). As the focus herein is on the evaluation of different retrieval systems, we do not explicitly evaluate different approaches for these modules and consequently focus on the above-discussed aspects instead of the complete quality dimension.

### 6.1.2 Argument Search Approaches

In the next step, the compared argument search approaches are discussed. To be included in the evaluation, the respective search engine has to meet the following two requirements: It has to be accessible by an API and provide information about the stance of the retrieved arguments. The first requirement is necessary for the technical applicability of the search engine within argumentative applications whereas the second requirement is motivated by the need for stance information in the majority of the desired tasks of an argumentative system.

The aim of an argument search engine is to retrieve a ranked list of arguments related to a given search query. Different systems introduced so far follow different paradigms in order to accomplish this goal (Ajjour et al., 2019) and include the one developed in the scope of IBM project debater (Levy et al., 2018), TARGER (Chernodub et al., 2019), PerspectroScope (Chen et al., 2019), args.me (Wachsmuth et al., 2017c) and ArgumenText (Stab et al., 2018a). Out of this list, only ArgumenText, args.me and TARGER provide an API to access retrieved arguments and only the first two also include information about their stance. Consequently, we focus our evaluation on these two and discuss the underlying approaches in detail in the following subsections. In addition, we propose a novel system utilizing a conventional web search approach in order to generate baseline results for the evaluation.

#### args.me

The args.me search engine (Wachsmuth et al., 2017c) allows users to search arguments related to a search query from a corpus with over 300k arguments retrieved from different debating websites. The indexing of arguments is done offline and independently of the search query. In contrast to the other approaches discussed herein, the underlying algorithm exploits the specific debate setup of the source pages in order to identify argument stance and boundaries. Consequently, arguments are defined as a set of premises related to a conclusion, as shown in the example[1] in Table 6.1. Although the arguments therefore include more contextual information than sentential arguments, it is difficult to use this output directly in dialogue systems. Especially in scenarios that do not adhere to a clear structure regarding speaking time and turn-taking, extensive utterances are hard to follow and understand if presented only by synthetic speech (Wilcock and Jokinen, 2021).

To determine a reasonable maximum number of words for an argument in our setup, we investigate the length of manually annotated arguments from an online debate. We use the argument structure described in Chapter 4 since

- it is based on data from one of the source pages of args.me (idebate.org) and
- was annotated for use in dialogue systems.

We find that over 97% of the manually annotated arguments consist of less than 60 words, and we therefore set this value as the maximum number of words included in the system output. Based on this threshold, we re-rank the arguments retrieved by args.me in order to prefer arguments with an overall length of the premises smaller than 60, given that the complete search query is present

[1] Material reproduced from www.idebate.org with the permission of the International Debating Education Association. Copyright ©2005 International Debate Education Association. All Rights Reserved

Table 6.1: Exemplary search result of args.me for the search query *marriage*, including all utilized information (premises, stance, conclusion).

| Information | System Return |
|---|---|
| Premises | *If marriage's main function is to protect against bereavement and divorce then it is essentially protecting against harms that it itself brings. Without marriage, bereavement and divorce would cease to be as serious harms as they currently are.* |
| Stance | con |
| Conclusion | Marriage represents a legal bond which protects both parties in a relationship. |

in the conclusion of the argument. In addition, the premises of longer arguments are truncated to include only the first $m$ sentences, whereas $m$ is the maximum number of sentences that leads to an overall number of words $\leq 60$.

### ArgumenText

The ArgumenText search engine (Stab et al., 2018a) extracts sentential arguments (Stab et al., 2018b) on arbitrary topics. To that end, it retrieves relevant documents from a large web crawl (CommonCrawl[2]) and detects for each sentence extracted from the most relevant documents whether it constitutes a supporting or opposing argument with regard to the query (i.e. topic under consideration), or no argument at all. The underlying algorithm uses an attention-based neural network as described in (Stab et al., 2018b), trained on annotated sentences from web documents for more than 40 topics. ArgumenText has been shown to yield a coverage of almost 90% compared to expert-curated collections of arguments on given controversial topics (Stab et al., 2018a). While the system's drawback is its lower precision compared to expert annotations (argument and stance detection), it can detect arguments on virtually any topic of interest and it retrieves content from many different sources. For a given search query, the API returns a list of arguments and their corresponding stances. In addition, confidence scores of both argument ($conf_a$) and stance ($conf_s$) detection are provided for each argument. We compute the overall confidence as $conf = conf_a \times conf_s$ and rank the retrieved arguments according to this score in order to equally take into account both aspects of the search. Table 6.2 shows an exemplary search result including the utilized information[3].

In addition, ArgumenText provides a Cluster API to group the retrieved arguments thematically. To this end, it determines similarity scores for argument pairs. These scores are then applied to form clusters based on aspects addressed within the arguments. The Cluster API relies on an optimized version of the Sentence-BERT method (Reimers and Gurevych, 2019) that makes use of an efficient bi-encoder that has been trained with additional samples ("Augmented SBERT")

[2]http://commoncrawl.org/ (last accessed 29 August 2021)

[3]https://guncontrolfacts.org/category/gun-control-pros-and-cons/ (CommonCrawl February 2016 crawl)

Table 6.2: Example search result of ArgumenText for the search query *gun control*, including all utilized information (argument, stance, confidence scores).

| Information | System Return |
|---|---|
| Argument | *Countries that have implemented strict gun control laws have been able to reduce the incidence of gun death.* |
| Stance | pro |
| Confidence | $conf_a = 0.999, conf_s = 0.997$ |

from a cross-encoder (Thakur et al., 2021). The utilized supervised approach to learn argument similarity was shown to outperform unsupervised approaches based on BERT embeddings by 10 percentage points (Reimers et al., 2019). Although the Cluster API is not used in the general comparison of argument search approaches, it will be applied for the automatic generation of argument structures from argument search results in the following section.

### Baseline

To compare the discussed argument search approaches to a suitable baseline, we introduce an architecture that utilizes the results of a conventional web search in order to find web pages that contain arguments related to the search query. To ensure reproducible results, we employ the ChatNoir search engine (Bevendorff et al., 2018) on CommonCrawl data with the search query *arguments <TOPIC>*. The text blocks of the websites with the highest ranking for each topic are then searched for sentences that contain topic-specific key words. The list of key words consists of substrings of the topic (for example *nuclear* and *energy*) as well as a WordNet (Bird et al., 2009; Miller et al., 1990) synonym either of the complete search query or (if none was found) the substrings. The final arguments are then sampled from this list with a fixed random seed, also to ensure reproducibility. For each retrieved argument, the stance is determined by a separate stance classification module. The pipeline of the complete baseline system is shown in Figure 6.1.

In order to determine the stance of the retrieved argument, we train a classifier on the IBM stance classification data (Bar-Haim et al., 2017). The corpus consists of 2394 claims annotated with an overall sentiment label $sent_c \in \{1, -1\}$, the stance towards the related topic and a sentiment label for the topic $sent_t \in \{1, -1\}$. Similar to the baseline approaches in the original work, we train a model to estimate the sentiment of each claim and assume that the target towards which the estimated sentiment is expressed is consistent with the target of the topic. The corresponding stance can then be derived as $sent_c \times sent_t$.

Our approach utilizes a pre-trained BERT (Devlin et al., 2019) model[4] to get sentence embeddings for all claims in the corpus. The corresponding sentence vectors are then used as input for a support vector machine (SVM) classification. The parameters of the SVM are optimized in a systematic grid search in order to match the specific task. The performance of our model is evaluated by averaging the results for five different random train/test splittings of the data with

[4] We utilize the *base* model, available at https://github.com/hanxiao/bert-as-service (last accessed 29 August 2021)

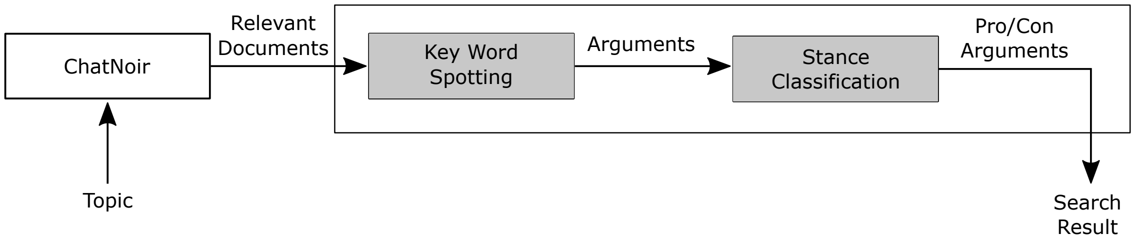


Figure 6.1: Sketch of the pipeline architecture of the proposed baseline search engine.

the same characteristics provided in the original work (training: 25 topics and 1039 claims, test: 30 topics and 1355 claims). Since the overall system requires an estimate of the stance for all retrieved arguments, we do not consider different coverage rates (as in the original work) and only investigate predictions on the complete test set, independently of the confidence of the classifier. For the sentiment classification, we report an average accuracy of 0.80 and an F1 score of 0.77. The final stance classification results in an average accuracy of 0.68 (F1 = 0.70), which is clearly above the baseline in the original work and slightly higher than the values provided for other therein discussed methods.

### 6.1.3 System

The final component of the complete evaluation setup is a dialogue system that enables users to apply the proposed evaluation criteria in an intuitive way and during the ongoing interaction. The utilized system is designed specifically for this task and allows the user to select his or her ratings as a direct response to the system utterance. The rating for each category can be given once for each argument and cannot be changed. In addition, the user is able to start the conversation, request the next argument, go to the previous one and repeat the latest utterance that includes an argument. If requested, the system selects the next argument randomly from the pool of available ones but each argument can occur only once during the interaction. It is important to note that the system requirements regarding the utilized arguments are as liberal as possible to enable the comparison of multiple different search approaches. The only information that has to be provided is the content of the argument as well as a notion of the respective stance towards the main topic. The overall interaction is stopped by the system after a fixed time to ensure the same conditions for each user.

The interface is based on the Charamel™ avatar[5] which presents the system utterance via synthetic speech by utilizing Nuance TTS and Amazon Polly voices[6]. Besides the avatar, the interface also includes buttons for the ratings in each category as well as the remaining user options (repeat, next, previous, start). If an option is not available in the current state of the interaction, this is indicated by the appearance of the respective buttons. A screenshot of the interface including buttons and avatar is shown in Figure 6.2.

[5] https://www.charamel.com/competence/avatare (last accessed 29 August 2021)
[6] https://docs.aws.amazon.com/polly/latest/dg/voicelist.html (last accessed 29 August 2021)


Figure 6.2: Screenshot of the system, including avatar and navigation buttons.

The system utterance is generated by a modified version of the template-based NLG introduced in Chapter 4 and receives the list of retrieved arguments from a certain argument search approach as input. Each argument is presented with an explicit notification of the assigned stance regarding the discussed topic. The specific formulation is selected randomly from a list of available options. In addition, the system starts each interaction by giving a topic-specific introduction including a short repetition of the task and the general claim *<TOPIC> is/are good*. Moreover, statements to notify the user if no further arguments are left and if the time for the discussed topic is over are also included. Table 6.3 shows the transcript of an interaction with the system, including the different utterance types of the dialogue system, the corresponding system utterance and exemplary user ratings. The arguments in this example were created manually for demonstration.

### 6.1.4 Experiment and Results

We applied the complete evaluation setup in a user study with 19 participants at NAIST. In order to minimize topic dependencies, we included the three common controversial topics *nuclear energy*, *self-driving cars* and *animal testing* into the evaluation and retrieved the best ranked 20 arguments for each topic using the three compared argument search approaches. In the course of the experiment, participants were able to interact with the system via the introduced interface and to rate the presented arguments in the discussed four categories *interesting*, *convincing*, *comprehensible* and *related*.

Table 6.3: Dialogue transcript for the topic *Veganism*, including excerpts of the introduction and the closing statement, three arguments and the corresponding exemplary user ratings.

| | System Utterance | Ratings |
|---|---|---|
| Intro | [...] The topic I want to discuss first is: *Veganism is good.* | - |
| $Arg_1$ | In contradiction to my initiating claim I found the following opinion: *Eating animals is in our nature, therefore it cannot be wrong.* | interesting<br>not convincing<br>comprehensible<br>related |
| $Arg_2$ | Okay, let's continue with a new aspect. The next argument is a supporting one. The content is: *There is a nice vegan restaurant, just around the corner.* | not interesting<br>not convincing<br>comprehensible<br>not related |
| $Arg_3$ | All right, let me see what else I found. Next in line is a support argument with the following content: *A key motivation for veganism is to prevent the exploitation of animals.* | interesting<br>convincing<br>comprehensible<br>related |
| End | Thank you for your ratings. Unfortunately, we reached the time limit for this topic. [...] | - |

### Setup

The experiment was divided into the three stages *introduction*, *interaction/rating* and *feedback*. During the introduction, each participant received written and oral instructions including an explanation of the interface, meaning of the categories and purpose of the experiment. In addition, a test trial with a separate small argument pool was offered in order to clarify the task. Since the experiment was conducted with non-native speakers, participants were not obliged to rate each argument in each category and instructed to skip a rating, if undecided. In addition, each participant rated the following statements/question on a five-point Likert scale from *completely disagree* (1) to *completely agree* (5) before starting the experiment:

- I'm in favour of <*TOPIC*>.
- How often do you use speech based devices/applications?

During the interaction phase, participants only interacted with the system and were not allowed to ask additional questions. Each participant listened to arguments for the topics *nuclear energy*, *self-driving cars* and *animal testing* retrieved with one of the three compared argument search approaches. In order to investigate the agreement between participants, a fourth topic (*death penalty*) was added. The pool of arguments for this topic was the same for each participant and included the top eight arguments retrieved with both args.me and ArgumenText. For each of the four topics, the system stopped the interaction after a fixed time of five minutes.

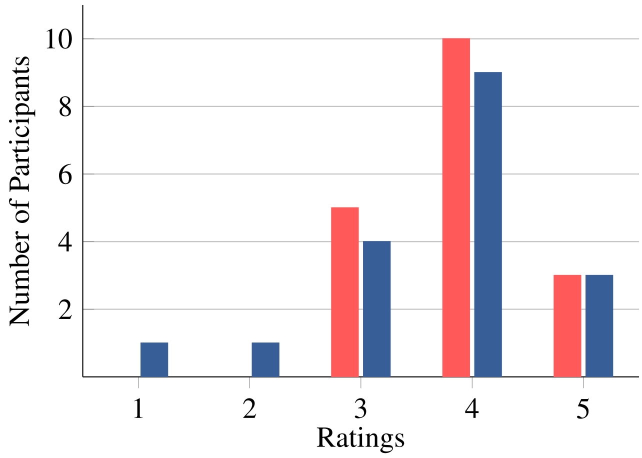


Figure 6.3: Responses on a five-point Likert scale from completely disagree (1) to fully agree (5) for the statements *The synthetic speech was easy to understand* (blue bars) and *All in all I had no problems understanding the system utterances* (red bars).

After the interaction, participants were asked to anonymously provide feedback about their own English proficiency with respect to the task, if they could understand the synthetic speech and the overall understandability of the system on a five-point Likert scale. Moreover, the opportunity to give written and/or oral feedback was provided. In case the language barrier hindered the completion of the task, respective ratings were excluded from the evaluation (this case occurred only once).

### Results

The experiment resulted in a total of 2,407 ratings distributed over all four topics. We start the evaluation by assessing the participant responses regarding the understandability of the system provided after the interaction to rule out errors related to the presentation of the arguments. Responses were given on a five-point Likert scale from *completely disagree* (1) to *fully agree* (5) for the statements

- The synthetic speech was easy to understand.
- All in all I had no problems understanding the system utterances.

Figure 6.3 shows the distribution of the participant responses. We see that only two participants disagreed with the statement that the synthetic speech was easy to understand and no participant explicitly reported problems with understanding the system utterances.

In the next step, we evaluate the agreement between the participants by means of Krippendorf's alpha (Hayes and Krippendorff, 2007) for the topic *death penalty* rated by all participants. This method is chosen as it allows to compare multiple raters, is able to handle missing data, and allows for a comparison with existing work. We analyse the ratings given for the same arguments in the

Table 6.4: Results of the statistical comparison of the ratings for the topics *nuclear energy*, *self-driving cars* and *animal testing*.

| Category | args.me | ArgText | Baseline | $p$ | args.me/ ArgText | ArgText/ Baseline | args.me/ Baseline |
|---|---|---|---|---|---|---|---|
| Interesting | 0.81 | **0.88** | 0.72 | $\leq 0.01$ | 0.10 | $\leq$ **0.01** | 0.10 |
| Convincing | 0.42 | **0.59** | 0.48 | 0.01 | **0.01** | 0.13 | 0.30 |
| Comprehensible | **0.85** | 0.83 | 0.78 | 0.32 | - | - | - |
| Related | **0.89** | 0.69 | 0.66 | $\leq 0.01$ | $\leq$ **0.01** | 0.70 | $\leq$ **0.01** |

Left part: Ratio of positive ratings and overall ratings for each category and architecture including the $p$-value derived with Fisher's exact test. Right part: Resulting $p$-values for all three pairings and categories that showed a significant difference derived by Fisher's exact test with Benjamini-Hochberg correction. Values for the category *comprehensible* are not included since the prior testing showed no significant differences between the three compared approaches.

same category, resulting in a maximum alpha of 0.15 for the agreement between all participants in the category *comprehensible*. Since participants were instructed to rate based on their own opinion and not according to objective guidelines, these results are as expected and emphasise the highly subjective nature of the task. In addition, we investigate the agreement between participants with the same personal stance towards the discussed topic and report that no increased agreement (consistent for pro or con) can be observed. Consequently, we proceed with a statistical evaluation of the complete ratings for each search approach and category (rather than an evaluation on the basis of individual arguments) to derive conclusions from the results.

For the statistical analysis, the ratings for all three approaches corresponding to the topics *nuclear energy*, *self-driving cars* and *animal testing* are compared for each category with Fisher's exact test (Sprent, 2011). The resulting $p$-values and the ratio of positive ratings and all ratings are shown in the left part of Table 6.4. We see that the categories *interesting*, *convincing* and *related* yield a $p$-value smaller than $\alpha = 0.05$ whereas for the category *comprehensible* no statistically significant difference between the investigated approaches is found. For a more detailed discussion, we compare the three approaches pairwise (also with Fisher's exact test) and utilize the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) in order to correct the $p$-values accordingly for multiple hypothesis testing. The results for all three pairings of the utilized systems are shown in the right part of Table 6.4.

The baseline shows no significant advantage over the other two approaches in any category. In contrast, the results for the *related* category indicate a clear advantage of args.me over both baseline and ArgumenText. In addition, ArgumenText shows a statistically significant advantage over args.me in the category *convincing* and over the baseline in the category *interesting*.

### 6.1.5 Discussion

This section introduced an evaluation setup for argument search approaches in the context of argumentative dialogue systems. The approach assesses the users' opinions and perceptions regarding arguments presented by an avatar with synthetic speech. During the interaction with the system, users are able to rate the arguments presented by the avatar in the categories *interesting*, *con-*

*vincing*, *comprehensible* and *related* as a direct response to the system utterance. The approach was applied in a user study in order to compare two argument search engines (ArgumenText and args.me) to each other and to a baseline architecture. Our results show a statistically significant advantage of both search engines over this baseline in one of the categories. Moreover, each search engine also shows a significant advantage over the other in a certain category which reflects the strengths and disadvantages of the underlying techniques. In the following, we discuss the results and implications of our findings. We will focus on the comparison of the investigated search approaches as well as the subjectivity of the evaluation task. Moreover, we address the implications of our findings for the use of argument search in dialogue system applications.

The inability of the baseline architecture to outperform the two investigated argument search engines emphasises the need for argument search in general to retrieve suitable arguments as it can clearly not be substituted with a conventional web search. Regarding the two different search engines, we argue that the advantage of args.me in the category *related* is due to its different approach to stance detection. Args.me utilizes the specific structure of debating websites to determine the stance of an argument as well as its boundaries and thereby ensures a very precise estimate of the argument relation to the discussed topic. In contrast, classifier based detection of stance and arguments as utilized in ArgumenText yield, at the current state of the art, a lower performance. On the other hand, classifier-based approaches allow for a search in a broader variety of sources and hence for a richer pool of arguments. This is in our opinion reflected in the advantage of ArgumenText in the *convincing* category. Moreover, the modified selection of arguments from the args.me search results (re-ranking and shortening) is likely to influence the perception of the users – but is (at the current time) unavoidable in order to use the respective arguments in a speech-based dialogue system. It is hence concluded that this is a current limitation of the system and can be overcome by a more fine-grained system output that includes a list of premises for each argument.

In coding and annotation tasks, a low agreement between coders usually indicates misunderstandings regarding the task or the guidelines. The low agreement between participants reported earlier hence raises the question if more restrictive guidelines for the participants are required. From our perspective, the goal of an argument-wise comparison of the participant ratings with expert annotations would justify this conclusion. However, we pursue the goal of assessing the user perception of the arguments based on his or her personal opinion, since users of argumentative applications are most likely not instructed on how to interpret the system output. Consequently, the observed disagreement is in our opinion a result of the different views of participants on the discussed topics and argumentation as a whole, i.e. the subjective nature of the task. The results are also in line with existing studies on argument quality where only slightly higher alpha values were achieved between seven annotators (Wachsmuth et al., 2017b), and no effect of the annotators' stance could be measured (Potthast et al., 2019).

The study results also enable some general conclusions for the development of future argumentative dialogue systems that aim to exploit argument search to retrieve arguments. Firstly, the reported subjectivity of the argument perception stresses the need for a careful selection of arguments based on the target audience or user. Consequently, adaptation and user modelling approaches investigated for use in dialogue systems (Casanueva et al., 2015; Mo et al., 2018; Ultes et al., 2019) are also required in the domain of argumentation, although it is not clear from the results which user traits are the most relevant. In addition, the reported high user comprehension of the system's utterances (Figure 6.3) allows the conclusion that arguments retrieved by argument search engines can generally be understood. However, several participants reported that spelling or grammar errors in the arguments lead to an unnatural system output, which, although generally comprehensible, was not natural and intuitive. Consequently, more advanced approaches to NLG in combination with paraphrasing and grammar correction are also of interest in order to improve the user experience (Kwon et al., 2015; Wen et al., 2015). It should also be noted that many dialogue systems require a more fine-grained argument structure (Aicher et al., 2021; Rach et al., 2018c; Rosenfeld and Kraus, 2016; Sakai et al., 2018b) that includes not just the general argument stance but also their explicit relations to each other. Hence, additional processing of the search results to structure the retrieved arguments like clustering of arguments (Reimers et al., 2019) may be required to enable systems of this kind to exploit argument search engines.

Several of the identified pending issues for the application of argument search in argumentative dialogue systems are addressed in the remainder of this thesis: The subjectivity of the user perception and the resulting need for user-adaptive approaches in argumentation is addressed in Chapter 7 in the form of a preference model which captures the individual perspective of the user on individual arguments. In addition, the chapter also introduces an approach to automatically assess the subjective user opinion regarding individual arguments from non-verbal cues shown during the interaction with the system. An approach to adapt the utterance selection, i.e. the system strategy to user feedback is discussed in the context of an application system in Chapter 8. Finally, the dependency of existing systems on argument structures that include relations between the arguments is addressed in the following section through a supervised-learning based mapping of argument search results into argument structures.

## 6.2 From Argument Search to Argument Structures

Based on the results of the previous section, the next step is now to apply argument search directly in argumentative systems. This section addresses the second task on the way to a topic flexible conversational system through the automatic generation of argument structures from search results, i.e. by answering the question of how the retrieved arguments relate to each other. Since the specific requirements regarding this structure differ from system to system, preliminary conceptual decisions regarding the argumentation framework are required to enable an evaluation in an application-oriented setup. Building on the results of the previous chapter, the agent-agent setup of Section 5.2 is utilized in the following to assess the proposed approaches. The work in this section was again conducted in cooperation with TU Darmstadt, builds on the bachelor's thesis of Schindler (2020) and is published in (Rach et al., 2021b).

To enable the desired automatic generation of argument structures, we propose an approach that maps the list of pro and con arguments retrieved with an argument search engine for a given topic into a general tree structure that encodes bipolar relations (support and attack) between the individual arguments (see Figure 6.4). In doing so, the approach combines the strong points of both data-driven and formal models for argumentation and enables a corresponding system to discuss literally any topic on which the search engine can find suitable arguments. Based on the comparison of search approaches in the previous section, the argument search engine

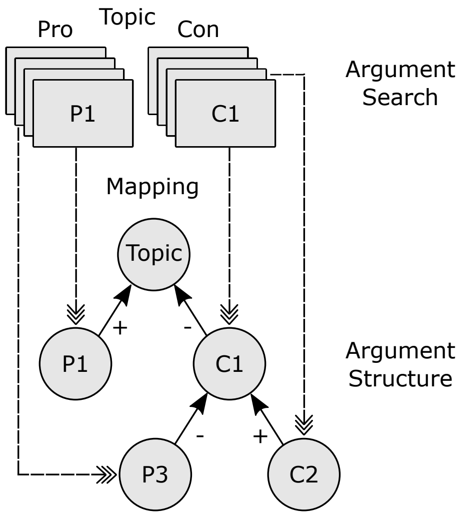


Figure 6.4: Mapping of argument search results to a tree structure with support (+) and attack (-) relations.

ArgumenText is used to retrieve pro and con arguments for a given topic from a large web crawl. This choice is due to the provided retrieval of arguments on a sentence level (which is preferable in a dialogue context), the reliable performance in comparison with the investigated baseline and the possibility to cluster the retrieved arguments thematically using the Cluster API. We train and compare two classifiers to detect relations between pairs of the retrieved arguments which subsequently enables the mapping into an argument structure.

For evaluation, artificial agent-agent dialogues are generated with the modified dialogue game introduced in Section 5.2. The resulting dialogues are then assessed in an extensive user survey with respect to their *coherence* and compared to the results achieved with an annotated structure. Although the annotated structure yields (as expected) an advantage over the automatically generated ones, the results are in some instances fairly close to each other. Besides, we observe varying results for the investigated topics, indicating a dependency of the approach on the available data. In summary, the contributions of this section are:

- An approach to automatically generate argument structures from argument search results.
- An extensive evaluation of this approach in a challenging dialogue setup.

### 6.2.1 Automatic Generation of Argument Structures

In the following, the mapping of the retrieved arguments into an argument structure is discussed. Although some structural argument representations utilized by the systems discussed in Chapter 3 differ to a certain extent, they all require information about the relations between the individual arguments. We hence pursue a modular pipeline approach that first determines possible relations between the arguments and subsequently maps them into a specific structure. In case the required structure cannot be inferred from the herein discussed one, the second module can be adapted accordingly.

The recognition of relations is a sub-task of argument mining as discussed in Section 2.2.4. It is usually utilized to extract the structure of a specific source in the context of *discourse-level argument mining*. However, the goal in the present context is to set arguments retrieved in a *information-seeking argument mining* task into relation and hence not to extract an already existing structure but instead to generate a new one for use in argumentative systems. The following subsections first introduce the target structure utilized in the present setup, discuss the proposed pipeline architecture and present results of a preliminary evaluation, thereby building on the work in (Schindler, 2020). The code of the proposed pipeline is publicly available[7].

### Target Structure

The herein considered target structure is again based on the argument annotation scheme in (Stab and Gurevych, 2014a) introduced in Section 2.2.1. It distinguishes three different types of argument components (*Major Claim, Claim, Premise*) and two directed relations between them (support and attack). Each component has one unique relation towards another component but can be targeted by multiple others. To keep the structure as general as possible, we abstract from this framework in the sense that we are not distinguishing different component types for the retrieved arguments and only focus on finding the best fitting relation of each component towards another (or the main topic, i.e. the search query). Details regarding this abstraction were discussed in Chapter 4. As a direct consequence, the resulting structure can again be represented as a directed tree with the retrieved arguments as nodes, the relations as edges and the main topic as root (as depicted in Figure 6.4). To prevent isolated circles, we further assume that each argument is (directly or indirectly) connected to the root.

With the herein used (informal) notion of arguments as search results, the resulting tree structure can be seen as a specified instance of bipolar argumentation frameworks (BAFs, see Section 2.2.1). From the perspective of the utilized dialogue model however, the nodes in the tree structure are still treated as components to construct new arguments as discussed in Chapter 4. Nevertheless, as ArgumenText retrieves search results on a sentence level that is similar to the granularity considered for annotating argument components (Stab and Gurevych, 2014a), the difference between the two perspectives is mainly the notation. In addition, the unique target in the tree structure and the specification of arguments in the dialogue game (one premise, see Section 4.1.3) allows for translating one perspective directly into the other. In particular, each argument from the search results serves as a premise for one (unique) argument in the dialogue game.

### Pipeline

Next, the retrieval pipeline is introduced. It takes arguments from an argument search engine (here ArgumenText) as input and outputs the above-discussed tree structure in an OWL file (Bechhofer, 2009). To this end, it first predicts relations between pairs of arguments and infers the final ar-

[7] https://github.com/csacro/From-Argument-Search-to-Argumentative-Dialogue (last accessed 29 August 2021)

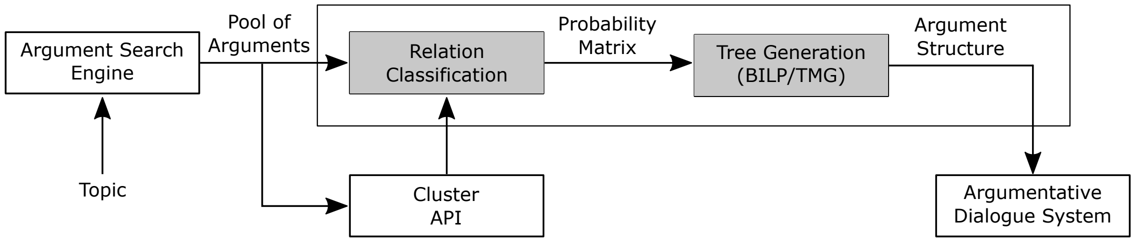


Figure 6.5: Sketch of the complete pipeline.

gument structure from them in a second step. We consider two configurations for the relation classification: The first one predicts relations between all possible argument pairs and hence imposes no restrictions on the shape of the resulting tree. The second one utilizes the ArgumenText Cluster API (Thakur et al., 2021) to group the retrieved arguments prior to the relation classification. Consequently, only argument pairings within a cluster are considered during relation classification and each cluster forms a new branch of the tree. The relation classification is trained on a balanced subset of the corpus[8] by Carstens and Toni (2015) as it labels sentence pairs with directed supportive, attacking or no relation. For the present task, the labels supportive and attacking are combined to a new label *relation* as the polarity can be inferred from the stance information provided by the search engine. Multiple classifiers including support vector machine (SVM), random forest and decision trees are compared on different feature sets with respect to their performance on the corpus. In addition, a BERT model (Devlin et al., 2019) is fine-tuned on the task. The detailed results are included in (Schindler, 2020) and we only include the best performing model as well as a strong baseline into the pipeline. The best performing classifier is the fine-tuned BERT model, reaching an average accuracy of 80.0% in a five-fold cross-validation. We select the SVM trained on BERT embeddings of argument pairs as the baseline due to its robust performance on a minimal feature set. The corresponding average accuracy in the five-fold cross-validation setup is 77.4%.

For the generation of the tree, we utilize the classifier confidence to compute probabilities for all estimated relations. Subsequently, we pursue two approaches to eliminate circles between arguments and derive the final tree structure: Binary Integer Linear Programming (BILP) optimizes the sum of the probabilities of the relations holding in the resulting tree under the structural constraints (Stab and Gurevych, 2017). In addition, we introduce Traversing and Modifying Graphs (TMG) which firstly identifies the most probable relation for every argument to another and connects them accordingly. Afterwards, it searches for circles as all resulting graphs which are not at least indirectly linked to the root contain exactly one such circle. In these circles, the node with the most probable relation to any node outside its graph is determined and the respective relation is redirected to this node. The complete pipeline is shown in Figure 6.5.

[8] https://www.doc.ic.ac.uk/ ft/softwareArg.html (last accessed 29 August 2021)

### Preliminary Evaluation

To compare the above-selected approaches on the actual task, a preliminary annotation study was conducted. To this end, we retrieved 20 arguments from ArgumenText for the topics *nuclear energy is good* as well as *animal testing is good* and compared different combinations of the approaches to create the tree structure. Clustering prior to the relation classification was not considered in this step, as it is investigated thoroughly in the final evaluation. Five annotators without task-related background were asked to label each argument pair with a relation in the resulting tree structure in each of the annotation categories *contradiction*, *entailment*, *specificity*, *paraphrase* and *local relevance* with *yes* or *no*. The first four categories are based on an investigation of the interactions between semantic relations by Gold et al. (2019), the last category was proposed in (Wachsmuth et al., 2017b). As in this latter work, we use the labels of the three most agreeing annotators for each category in order to eliminate outliers.

The Fleiss' kappa (Fleiss, 1971) values yields a substantial (0.66) up to perfect (0.82) agreement (Landis and Koch, 1977). A pair of arguments is concluded to actually hold a relation if it is rated with *yes* in at least one category by majority vote. For our baseline (SVM), this is the case with BILP as well as with TMG for 62.5% of the argument pairs. The BERT model correctly relates 75.0% of the argument pairs with TMG and 77.5% with BILP and we hence select the fine-tuned BERT model for the subsequent evaluation. It should be noted that BILP is highly time-consuming for large structures due to the underlying optimization problem. Since both approaches show similar performances, we only consider TMG in the final evaluation.

## 6.2.2 Evaluation

In this section, the evaluation of the complete approach for argument retrieval is addressed. We utilize the modified formal framework for persuasive dialogues discussed in Section 5.2 to create artificial dialogues from the automatically retrieved structures. The written transcripts of these dialogues are then assessed in a crowdsourcing study with respect to dialogue coherence. For the evaluation, we first introduce the study setup and discuss the results subsequently.

### Setup

The first step in the evaluation of the generated dialogues is the selection of a meaningful set of evaluation categories. The ones utilized herein are based on the notion of dialogue coherence for conversational agents as discussed in (Venkatesh et al., 2017). The authors define a coherent response as one that is neither *irrelevant*, *incorrect* or *inappropriate*. However, a direct application of these criteria is difficult in argumentative settings as for example the correctness of an argument is hard to assess. Therefore, each category is adapted into a yes/no question which directly evaluates utterance properties that are influenced by the retrieval pipeline. The resulting categories are as follows:

- Do you understand what the speaker wants to say (comprehensible)?
- Does the utterance address its reference (reference)?
- Does the utterance contradict the speaker's position (polarity)?

For the study, we implemented a web interface that presents the dialogues utterance-wise to the participants. In the beginning, participants received written instructions about the purpose of the survey and each of the above questions. In addition, a detailed example with manually generated arguments and explanations for the therein included ratings was provided to make the participants familiar with the setup and the evaluation categories. Help icons that displayed the question description from the introduction were included in the web interface to make the explanations available during the study. To test the attention of the participants, we also included test utterances that asked the participant for a specific rating into some of the dialogues. Participants that failed a test utterance were excluded from the study. Each participant assessed three dialogues and was asked to rate the statement *The explanation/definition provided for the question was clear* for each evaluation category on a five-point Likert scale from *totally disagree* (1) to *totally agree* (5). In addition, participants were able to provide written feedback on the survey.

We generated argument structures for seven different topics, namely *nuclear energy*, *abortion*, *self-driving cars*, *school uniforms*, *death penalty*, *animal testing* and *marriage*. The first six topics are used to compare the two pipeline configurations (with and without clustering) and for a general assessment of the artificial dialogues. The topic *marriage* on the other hand is used for a comparison to an annotated structure. The utilized reference structure is discussed in detail in Chapter 4. It includes 72 manually annotated arguments and relations between them from an idebate.org debate on the topic *Marriage is an outdated institution* (Rach et al., 2019a). For each topic, we retrieved a pool of 60 arguments for the query *TOPIC is/are good* with ArgumenText and generated two structures per topic (with and without clustering). For each of the 14 automatically generated structures as well as the annotated one, we generated one reference dialogue for the evaluation and five additional dialogues. From the five additional dialogues, the one that has the least amount of arguments in common with the reference dialogue was added to the evaluation. Consequently, we arrived at a total of 30 dialogues which were divided into ten groups of three dialogues each. To ensure similar conditions for all groups, the dialogues had a fixed length of 20 game moves. Participants were assigned to one of the ten groups in order of appearance and we investigated seven raters per group, resulting in a total of 70 participants. The study was realized via clickworker[9] with participants from the UK (55) and the United States (15). The participants were aged between 18 and 67 years, 31 of them were female and 39 male.

### Results

The study resulted in a total of 10,122 ratings over all three categories. We start the assessment of the results by computing the agreement over all three questions in each group with Fleiss' kappa (Fleiss, 1971). The resulting agreement is rather low with a maximum of 0.46 (group 3) and a minimum of 0.14 (group 4) and hence indicates problems in the comprehensibility of the task. Consequently, we investigate the participants' self-report on the clarity of the task next. The corresponding results are shown in Figure 6.6. Although the majority of the ratings is either neutral or positive, there is also a certain percentage of negative ratings, especially for the *polarity* question. In total, 29 participants rated at least one category with *disagree* or *totally disagree*. Thus, we again consider the best agreeing three participants to derive the final score. The group-

[9] https://marketplace.clickworker.com (last accessed 29 August 2021)

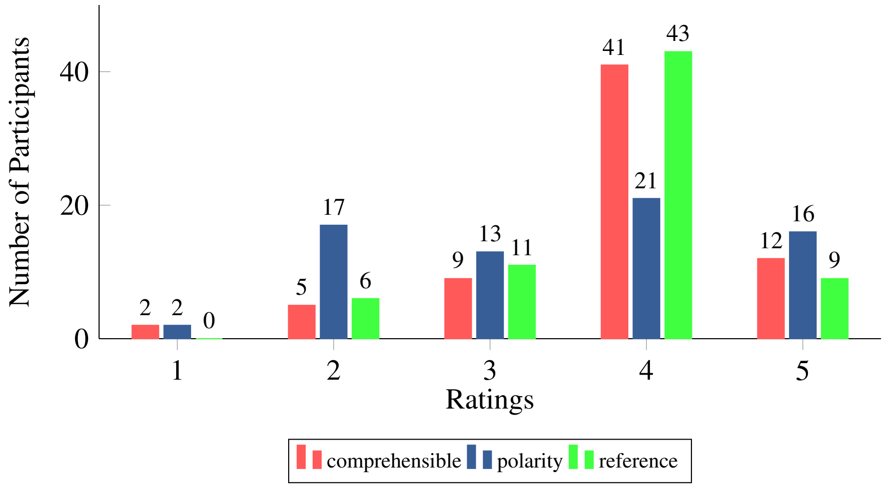


Figure 6.6: Responses on a five-point Likert scale from completely disagree (1) to completely agree (5) for all three evaluation questions and the statement *The explanation/definition provided for the question was clear*.

Table 6.5: Agreement derived with Fleiss' kappa for all ten groups (G1 - G10) and all annotators (all) as well as the three most agreeing annotators (best 3).

| | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | G10 |
|---|---|---|---|---|---|---|---|---|---|---|
| all | 0.30 | 0.43 | 0.46 | 0.14 | 0.26 | 0.28 | 0.19 | 0.40 | 0.32 | 0.28 |
| best 3 | 0.73 | 0.72 | 0.76 | 0.36 | 0.45 | 0.50 | 0.44 | 0.65 | 0.64 | 0.50 |

wise agreement for all and the best agreeing three participants is shown in Table 6.5. It can be seen that now all groups show a fair or better agreement (Landis and Koch, 1977). Given the subjective nature of the task (Wachsmuth et al., 2017b), we consider this a sufficient agreement for our evaluation and use the majority vote of the best agreeing three annotators in the following.

We proceed with a comparison of the two investigated pipeline configurations (with and without clustering) and subsequently compare the results of the automatically generated structures for the topic *Marriage* to the ones achieved with the annotated structure. We investigate each category/question separately and also compute the utterance-wise *coherence*. An utterance in the dialogue is fully coherent if it is comprehensible, addresses its reference and does not contradict the speaker's position, i.e. if it is rated with *yes*, *yes*, *no*. Example majority ratings for a dialogue excerpt generated with an automatically retrieved structure on the topic *school uniforms are good* is shown in Table 6.6.

For the comparison of the two pipeline configurations, we consider all topics with only automatically generated structures in the survey, namely *nuclear energy* (NE), *abortion* (A), *self-driving cars* (SDC), *school uniforms* (SU), *death penalty* (DP) and *animal testing* (AT). The cor-

Table 6.6: Artificial dialogue between the agents A and B generated with the dialogue game for argumentation and an automatically retrieved argument structure.

| Speaker | Utterance | Majority Ratings | Coherent |
|---|---|---|---|
| A (pro) | *School uniforms are good.* (Topic) | - | - |
| A (pro) | *The uniform can help ease the problem about the fashion in school especially in a high school or a university that have a lot of teenagers together.* (esl-lab.blogspot.com) | Comprehensible: *yes*<br>Reference: *yes*<br>Polarity: *no* | *yes* |
| A (pro) | *Wearing the same school uniform creates an image of unity and helps student feel accepted in the school.* (ericmacknight.com) | Comprehensible: *yes*<br>Reference: *yes*<br>Polarity: *no* | *yes* |
| A (pro) | *It also shows unity and children feel included and it helps them to work as a team.* (www.createdebate.com) | Comprehensible: *yes*<br>Reference: *yes*<br>Polarity: *no* | *yes* |
| B (con) | You seem quite confident about that point. What makes you so sure? | Comprehensible: *yes*<br>Reference: *yes*<br>Polarity: *no* | *yes* |
| B (con) | In particular, there's one aspect of your argumentation that I have some doubts about. You said *the uniform can help ease the problem about the fashion in school especially in a high school or a university that have a lot of teenagers together. Rude rules to wear school uniform is killing students' personality and it can be influenced for their future developing as bright personality.* (www.createdebate.com) | Comprehensible: *yes*<br>Reference: *yes*<br>Polarity: *no* | *yes* |
| B (con) | *School Uniforms are not safe to wear.* (www.createdebate.com) | Comprehensible: *yes*<br>Reference: *no*<br>Polarity: *no* | *no* |
| B (con) | *Uniform itself is also uncomfortable and during the winter it is very cold.* (www.createdebate.com) | Comprehensible: *yes*<br>Reference: *no*<br>Polarity: *no* | *no* |

Left part: Speaker and natural language utterances. Right part: Majority answers to the evaluation questions and corresponding coherence results. Automatically extracted content in the respective argument structure is shown in italic font. All sources are provided by ArgumenText from the CommonCrawl February 2016 crawl.

responding ratio of positive and overall ratings is shown in Table 6.7. The first thing to notice is that the results are highly topic dependent, in direct comparison to each other and also in the effect of the clustering: Whereas for example the topic *abortion* reaches a coherence of 0.87 without clustering, the counterpart with clustering shows only a coherence of 0.50. On the other hand, we see that the coherence result for the topic *self-driving cars* is 0.47 without and 0.70 with clustering. The average over all topics (Overall) indicates a slight advantage of the group without clustering in all categories. However, a category-wise statistical comparison of the overall results with Fisher's exact test (Sprent, 2011) shows no significant difference between the two groups,

Table 6.7: Results for the structures with (lower table) and without (upper table) clustering and the topics *nuclear energy* (NE), *abortion* (A), *self-driving cars* (SDC), *school uniforms* (SU), *death penalty* (DP) and *animal testing* (AT).

| | NE | A | SDC | SU | DP | AT | Overall |
|---|---|---|---|---|---|---|---|
| Comprehensible | 0.89 | 0.92 | 0.94 | 0.94 | 0.66 | 0.88 | 0.87 |
| Reference | 0.80 | 0.97 | 0.90 | 0.78 | 0.91 | 0.85 | 0.87 |
| Polarity | 0.71 | 0.97 | 0.59 | 0.97 | 0.94 | 0.97 | 0.86 |
| Coherence | 0.60 | 0.87 | 0.47 | 0.75 | 0.53 | 0.76 | 0.67 |
| Comprehensible | 0.96 | 0.78 | 0.90 | 1.00 | 0.77 | 0.93 | 0.89 |
| Reference | 1.00 | 0.67 | 0.87 | 1.00 | 0.65 | 0.76 | 0.82 |
| Polarity | 0.82 | 0.78 | 0.87 | 0.48 | 0.94 | 1.00 | 0.82 |
| Coherence | 0.79 | 0.50 | 0.70 | 0.48 | 0.55 | 0.72 | 0.62 |

Table 6.8: Comparison of the results for the annotated structure with the automatically generated ones and the topic *marriage*.

| | annotated (a) | cluster (c) | no cluster (nc) | a/c | a/nc | c/nc |
|---|---|---|---|---|---|---|
| Comprehensible | 1.00 | 0.68 | 0.83 | $< 0.01$ | 0.04 | 0.24 |
| Reference | 1.00 | 0.68 | 0.86 | $< 0.01$ | 0.08 | 0.13 |
| Polarity | 1.00 | 0.82 | 0.49 | 0.01 | $< 0.01$ | 0.01 |
| Coherence | 1.00 | 0.43 | 0.34 | $< 0.01$ | $< 0.01$ | 0.60 |

Left table: Ratio of positive and overall ratings. Right table: p-values of pairwise comparison with Fisher's exact test and Benjamini-Hochberg correction.

indicating that (on average) both configurations perform equally well. Finally, we compare the results achieved with the annotated structure to the results achieved with the automatically generated ones. We conduct a pairwise comparison of the three groups again with Fisher's exact test and a Benjamini-Hochberg correction (Benjamini and Hochberg, 1995) of the p-value. The corresponding results for all three structures are shown in Table 6.8. The annotated structure yields a perfect score of 1.00 for all categories, which is not surprising since it is tailored to dialogue setups. The comparison further indicates that the annotated structure outperforms the automatically generated ones, except for the reference category where no significant difference is found between the annotated structure and the automatically generated one without clustering. However, we can also see that the results of the automatically generated structures are lower than for the other investigated topics (see Table 6.7), which indicates that the topic *marriage* is more challenging than the others for the proposed pipeline.

### 6.2.3 Discussion

This section has addressed the automatic generation of argument structures from argument search results for their use in dialogue systems. To this end, a pipeline was introduced that estimates relations between the retrieved arguments and maps them into a general tree structure. We explored two different configurations, namely with and without a prior clustering of the retrieved arguments and utilized a supervised learning-based relation classification to identify related argument pairs. For evaluation purposes, we generated 30 artificial dialogues over seven different topics and assessed them in a crowdsourcing setup with respect to their *coherence*.

The findings and implications of the conducted evaluation are discussed in the following. As already mentioned, the results vary between the investigated topics for all four categories. The difference between the individual topics can be attributed to the different sources the arguments are retrieved from and the resulting performance difference of the utilized techniques. The effect of the clustering on the other hand is not so clear as both structures for a topic are based on the same pool of arguments. However, as the position of the arguments in the tree is directly influenced by the relation classification (and hence by the clustering as well), it varies a lot between the structures with and without clustering. Therefore, the individual arguments can appear in a different context, which arguably also leads to a different perception through the study participants. On average, no significant difference between the two approaches could be found and the choice of the optimal configuration hence depends on the available data for each topic.

The direct comparison with an annotated structure revealed room for improvement, especially with respect to the overall *coherence*. However, we also found that for the individual categories *comprehensible* and *reference*, the results achieved without clustering are fairly close to the performance of the annotated structure. Especially for the *reference* category, which is directly influenced by the herein introduced pipeline, the found difference between the annotated and the automatically generated structure is not statistically significant. In addition, the *coherence* results of the automatically generated structures on the topic *marriage* were lower than for the other investigated topics, indicating that this was the most challenging topic for our approach. Although the above-discussed data dependency renders generalizations difficult, this *coherence* difference between the topic *marriage* and the others indicates that the overall pipeline performance is closer to the one with annotated structures than suggested by the direct comparison.

As for the written feedback, multiple annotators reported confusing formulations of the argument as the major difficulty of the task. Since this is a direct consequence of the heterogeneous sources the arguments are retrieved from, it is hard to address in the pipeline. Therefore, approaches to automatically summarize or reformulate arguments (Bar-Haim et al., 2020; Schiller et al., 2021) could be beneficial to improve the performance.

Regarding applications, it can be seen that the proposed approach is quite flexible: Although a specific multi-agent setup was chosen for evaluation, the proposed pipeline itself has no dependency on this particular setting or the corresponding domain of persuasive dialogues. Therefore, it can be directly applied in other domains and scenarios as well if the respective dialogue system operates on structures of the retrieved kind. This includes for example systems in the opinion building domain (Aicher et al., 2021) or systems that combine argumentation with other types of dialogue like question answering (Sakai et al., 2018a). In addition, the proposed pipeline can be combined with methods that build on the investigated representation of arguments. In particular, the probabilistic rule-based strategy used in the evaluation setup can be extended or replaced with more sophisticated ones in compliance with the desired application. Examples in this regard are strategies optimized via reinforcement learning (see Chapter 5) as well as argument selection based on semantics (Cayrol and Lagasquie-Schiex, 2005) or user concerns (Chalaguine and Hunter, 2020). In light of the evaluation results, the main task for future work with respect to applications is hence the improvement of the pipeline performance to fully meet the quality requirements of the individual systems. However, as the proposed approach relies on argument search engines, it directly benefits from future developments in this area. Moreover, the addition of weights to arguments in the structure could further broaden the range of possible applications. The corresponding weights can for example be derived from the confidence scores of the pipeline components or through automatic approaches to assess argument quality (see Section 3.2).

In summary, the results indicate that the proposed pipeline depends on the quality of the available data but yields promising results for the majority of the investigated topics and at least one of the two investigated configurations (with and without clustering). In comparison to an annotated structure, we observed a similar performance for individual categories but also the expected room for improvement regarding the overall coherence. The proposed approach can be seen as a first step towards fully automatized argument acquisition for argumentative dialogue systems. Since it is based on argument search engines, it benefits directly from future improvements and developments in this area.

## 6.3 Argument Structures from Reviews

In this final section, the semi-automatic generation of argument structures from hotel and restaurant reviews is addressed. The procedure utilizes sentiment analysis labels to induce relations between the annotated sentences which are subsequently revised manually to ensure high-quality structures. The motivation of this approach originates in the question of how the persuasive effectiveness of a system can be assessed through experiments with human participants. The main problem in this regard are personal opinions towards the topic, i.e. the bias of the participant. The effect of this bias on the perception of the system and its performance is hard to isolate. It is therefore desirable to focus on *non-opinion based* topics for which the user has a minimal (or no) bias for evaluation studies.

The herein proposed approach is tailored to this requirement and was first published in a cooperative work with Augsburg University (Weber et al., 2020a). Again, the utilized argument representation is based on the annotation scheme introduced in (Stab and Gurevych, 2014a) and discussed in Chapter 4. In contrast to the previous section, the proposed approach also distinguishes between component types (*Major Claim*, *Claim*, *Premise*) during the generation of the argument structures. Whether or not this information is used in the corresponding evaluation depends on the specific scenario. As discussed earlier, the formal difference between the argument component types is only their allowed relations, which means that each completed structure can be transformed into an abstract directed tree without specified component types, if required.

### 6.3.1 Experimental Requirements

The addressed experimental setups aim at assessing the persuasive effectiveness of an argumentative system in the interaction with human users. Example components that are candidates for such an assessment are argumentation strategies of the system and the multimodal behaviour during an interaction. An evaluation conducted with the discussed *non-opinion based* structures is included in Chapter 7. In addition, an application system that relies on the retrieved structures to compare different persuasion strategies is introduced in Chapter 8. For the generation of the argument structures, the following requirements regarding the envisioned areas of application are identified:

- The argument components should be concerned with a *non-opinion based* topic, i.e. a topic for which a minimal bias of the user can be assumed (Req. 1).
- The argument components considered during evaluation should be similar to the argument components in the actual applications. In particular, if the persuasive effectiveness of a system on an emotional level is evaluated, the argument components should include a reasonable range of different emotions (Req. 2).
- The structure should be balanced with respect to the number of positive and negative arguments to avoid a bias based on one-sided information (Req. 3).
- The quality of the structure should not affect the evaluation outcome negatively through structural errors e.g. repeated relations with wrong polarity (Req. 4).

The first two requirements are addressed jointly through the extraction of arguments from hotel reviews: If the individual user is not familiar with the discussed hotel, a minimal (or even no) bias can be assumed, which satisfies the first requirement. Moreover, the use of real-world data ensures that the argument components express actual opinions. They are hence closer to the actual arguments employed in most systems than manually generated arguments on a fictitious topic would be. In addition, the use of real-world reviews ensures that different emotions are included in the formulation of the argument components. The third requirement can be satisfied by comparing the number of pro and con arguments in the structure and, if a more detailed analysis is required, by comparing the number of arguments on each level of the resulting argument structure. As for the last requirement, an annotated data set and a semi-automatic procedure to extract the argument structures are considered. The overall quality of the result is ensured through the use of robust rules (instead of data-driven approaches) as well as through human supervision.

### 6.3.2 Generation of Argument Structures

Based the discussed requirements, hotel and restaurant reviews from the annotated *SemEval-2015 Task 12* Test Datasets (Pontiki et al., 2015) are utilized. The argument structures are then inferred from the labels with a procedure adapted from the argument mining approach described in (Cocarascu and Toni, 2016). The data set choice is due to the fact that the corpus provides high-quality annotations on real-world reviews and includes all the required information:

- An identifier (id) that associates the annotated sentence with a specific review and indicates the order of sentences within a review.
- An aspect category *Ent.#Att.* consisting of an entity (e.g. *hotel*, *service*, *location*) and an attribute (e.g. *price*, *quality*).
- A polarity (*positive*, *negative*, *neutral*) of the annotated category.
- An opinion target expression that indicates an explicit reference to the entity within the annotated sentence (if present).

The annotation is illustrated by the following example that includes the original sentence as well as the annotated labels.

- *"Vending machines were out of everything except in the lobby."*
  - polarity: *negative*, category: *facilities#general*, target: *Vending machines*

It should be noted that multiple annotations are possible for one sentence if for example multiple entities are addressed. The argument structures are derived from these annotations hotel-wise, meaning that each argument structure is concerned with one particular hotel.

The generation of the argument structures is divided into the four steps *filtering*, *template generation*, *argument component assignment* and *supervision*, which are described in detail in the following. During the first (filtering) step, all sentences from reviews concerned with the investigated hotel are retrieved from the annotated database. Subsequently, the annotations of each sentence are filtered to exclude all annotations that fulfil one (or more) of the following criteria:

- The polarity annotation is inconsistent with another polarity annotation for the same entity or the polarity annotation is neutral.
- The annotated sentence is marked as *OutOfScope*.
- The annotation has no opinion target expression.

The remaining annotations are called valid in the context of the argument structure and are used to include the corresponding sentence into the structure in a later step. It is possible that a sentence has multiple valid annotations.

During the second step (template generation), a template for the desired structure is generated consisting of the *Major Claim "This hotel is worth a visit"* and one *Claim* for each entity in the annotation scheme. The polarity of each *Claim* is determined by comparing the number of positive and negative valid annotations for the respective entity. For example, if there are more negative than positive valid annotations for the entity *facilities*, the respective formulation of the *Claim* is *"The facilities are bad"*.

In the third step (argument component assignment), the annotated sentences are associated with argument components in the structure. To this end, it is assumed that sentences within one review and with the same entity label build on each other. This means that the first sentence targets the respective *Claim* and the following build a chain of arguments. For the remaining argument components, a direct relation to the respective *Claim* is assumed unless the component in question has the same target expression as a predecessor. In this case, it is directly related to the respective preceding component. Sentences with multiple annotations for different entities (and hence also different *Claims*) are included in all related sub-structures. If a component has multiple labels regarding the same entity, it is only included once for this entity. In all cases, the relation is determined by the polarity. If a component has the same polarity as its target, the corresponding relation is support and attack otherwise. The complete automatic procedure is shown in Algorithm 2. As an example, we look at the following two annotated sentences from the same review and the corresponding annotations:

- *"Lots of services were promised and not provided."*
  - polarity: *negative*, category: *facilities#general*, target: *services*
- *"The shuttle wasn't running, and the restaurant was closed down - both of those explained to me with because business is low right now around the holidays."*
  - polarity: *negative*, category: *facilities#general*, target: *shuttle*
  - polarity: *negative*, category: *facilities#general*, target: *restaurant*

Since the first sentence has the entity *facilities*, a relation to the corresponding *Claim* is assumed. For the above example (*"The facilities are bad"*), the relation is support, since the polarities of the annotated sentence and the *Claim* are consistent. The second sentence has two labels that differ only in the target expression but have the same entity. Consequently, it is only included in the *facility* branch of the tree as an extension to its predecessor, since they are from the same review. Again, the relation is support since the polarities are the same.

So far, the procedure is completely automatic. In the last step (supervision), human supervision is employed to correct the retrieved structure to ensure the high quality required in the targeted evaluation setups. First, all sentences that are included repeatedly (i.e. have multiple valid annotations for different entities) are split into separate utterances (if possible). The guiding principle for this separation is to maintain the original formulation as much as possible. For example, the sentence *"The staff was very friendly and helpful and the rooms were large and comfortable"* is divided into a component *"The staff was very friendly and helpful"* for the entity *service* and a component *"The rooms were large and comfortable"* for the entity *rooms*. In addition, sentences with the same content are merged into one component and all relations that target the original components are redirected to the new one. Finally, severe spelling and grammar errors are corrected to enable the use of synthetic speech. The resulting structure is encoded in an OWL ontology (Bechhofer, 2009), including component type, relations and natural language sentences. Four argument structures were extracted for use in later chapters, two in the hotel domain and two in the restaurant domain.

### 6.3.3 Discussion

The generation of argument structures from hotel and restaurant reviews was investigated to meet the experimental requirements of evaluation studies. To this end, requirements for studies of this kind were identified and addressed through the selection of a suitable data set and a tailored semi-automatic procedure to extract argument structures from it. Although the extraction

**Algorithm 2:** Argument Structure Generation

```
Init: compID = 0, compPolarity = 'positive', compText = 'This hotel is worth a visit.',
  compType ='MajorClaim', compTarget = None, ArgStrucutre, components
retrieve data for the respective Hotel from the database
add Arg(compType, compText, compPolarity, compTarget, compID) to ArgStructure
foreach sentence in data do
    if sentence not OutOfScope then
        listOfEntities ← Array()
        text ← text of sentence
        foreach annotation in sentence do
            get entity, target, polarity, review from annotation
            valid ← True
            valid ← check polarity (consistent and not neutral) and target (not NULL)
            if entity in listOfEntities then
                valid ← False
            if valid then
                add Comp(text, entity, target, polarity, review) to components
                add entity to listOfEntities
foreach entity do
    compType ← 'Claim'
    compare number of positive and negative polarities for category to get compPolarity
    generate compText for Claim based on compPolarity and category
    compTarget ← 0
    compID ← compID + 1
    add Arg(compType, compText, compPolarity, compTarget, compID) to
      ArgStructure
foreach comp in components do
    compType ← 'Premise'
    pred ← None
    pred ← predecessor in ArgStructure with same target, if present
    pred ← predecessor in Argstructure with same review and category, if present
    if pred not None then
        compTarget ← compID of pred
    else
        compTarget ← compID of Claim corresponding to entity of comp
    compPolarity ← polarity of comp
    compText ← text of comp
    compID ← compID + 1
    add Arg(compType, compText, compPolarity, compTarget, compID) to
      ArgStructure
∀Arg ∈ ArgStructure derive relations from polarities
```

is in the presented version rule-based and includes human supervision, it can be combined with data-driven approaches to enable automatic extraction of argument structures from reviews in the future: As the utilized data set is introduced for the development of sentiment analysis models, it can be directly utilized to train a classifier for predicting the annotated labels. Given successful training, this model is then able to extract labelled sentences from which the corresponding argument structure can be derived with the herein proposed method. In addition, approaches to automatize the supervision step are also available. The identification of repeated components can be addressed through the use of argument similarity measures (Thakur et al., 2021) which have seen a lot of progress in recent research. For the splitting of components with multiple different annotated entities, two different approaches are identified: The first one is to utilize paraphrasing approaches based on the recognized target expression, whereas the second is the use of an extended argument representation that includes multiple relations of a single component to other components. Applications for which the data-driven approach might be preferred over the robust but less flexible approach considered herein are for example recommendation systems that identify the best restaurant or hotel for users based on their individual preferences and concerns. However, as the motivation for the introduced approach is to enable the evaluation of persuasive systems, this is not addressed in the present thesis and left for future work. Nevertheless, the argument structures generated with the discussed procedure are utilized in the following to evaluate the predictive power of a preference model in Chapter 7 and in the multi-agent system that aims at displaying the effect of different persuasion strategies in Chapter 8.

## 6.4 Conclusion

This chapter was concerned with providing flexibility for argumentative conversational agents with respect to the topics that can be addressed. To this end, topics were divided into the two broad categories *opinion* and *non-opinion* based and approaches to acquire arguments for both of them were introduced.

For *opinion-based* topics, the use of state-of-the-art argument mining applications, namely argument search engines, was investigated. As the goal of these search engines is, similar to web search engines, to provide a list of pro and con aspects regarding a specific topic in written language, their application in dialogue systems is not trivial. To enable the desired combination of the two technologies, two main challenges were addressed: The first one is the suitability of the retrieved arguments for dialogue systems, in particular for the use of synthetic speech. The second one is the dependency of many argumentative dialogue systems on structured knowledge about the arguments and their relations towards each other.

The first aspect was addressed through an extensive evaluation of argument search engines in the context of dialogue systems. To this end, reasonable evaluation categories were identified based on argument quality dimensions as well as common approaches to dialogue system evaluation. In addition, an evaluation system was introduced that presents the retrieved arguments to human users and enables them to rate the presented content in the provided evaluation categories. The setup was used to compare two state-of-the-art search engines to a baseline system in an extensive user study thereby indicating strengths and weaknesses of the underlying technological approaches.

Based on this evaluation, a suitable candidate for addressing the second challenge was identified and an approach to map the retrieved arguments into an argument structure with bipolar relations was proposed. It was evaluated through the generation of artificial dialogues using the modified framework introduced in Chapter 5.2 and an assessment of the resulting agent-agent dialogues in an extensive user study. The results indicate the general feasibility of the approach but also the expected room for improvement in comparison to an annotated structure.

As the main motivation for *non-opinion* based topics is their suitability for evaluation studies, the need for robust methods resulting in high-quality structures out-weights the requirement of full flexibility. Consequently, a robust procedure that meets the requirement of evaluation studies and includes human supervision was discussed. It extracts arguments from reviews that are annotated with sentiment analysis labels.

In conclusion, the goal of providing topic flexibility was accomplished especially in the case of *opinion* based topics where the proposed approach allows a corresponding system to discuss literally any topic on which the search engine can find suitable arguments. Although this comes at the price of reduced dialogue coherence in comparison to annotated structures, the dependency on argument search engines enables the approach to directly profit from future developments made in this area.

A limitation of the proposed mapping approach is that it includes (in the current version) only one specific argument representation, i.e. the herein investigated tree structure. However, the modular architecture of the retrieval pipeline allows for straightforward adaptation of the approach to other representations, given that they rely on similar relations between argument components. Although some models (like the Toulmin model discussed in Section 2.2.1) consider a more fine-grained representation, it is evident from the discussion of related work in Chapter 3 that most of the current argumentative dialogue systems utilize representations that include either bipolar or only attack relations and do not strictly distinguish between multiple component types (beyond claim and premise/evidence). Consequently, the proposed approach is concluded to meet the conceptual requirements of current systems.

Finally, the proposed methods are compatible with the other approaches discussed in the scope of this thesis, namely the strategy optimization in Chapter 5 (as previously discussed) and the user-adaptive approaches discussed in the following chapter. In particular, all application systems introduced in Chapter 8 can directly utilize the argument structures retrieved with one of the proposed methods without additional modifications.

# 7 Adapting Artificial Argumentation to Human Users

This chapter covers the third and final aspect of flexible argumentation addressed in this thesis, namely the adaptation of an argumentative system to human users. As it has become clear in the previous chapter that the perception of arguments is highly subjective, the task is now to equip conversational agents with the capacity to recognize as well as react to the subjective view of individual users. However, although it is intuitive *why* a system should be able to adapt, it is a priori not obvious *how* this adaptation is to be approached. Whereas adaptation is natural for humans, it poses several technical challenges for conversational agents. In the context of the present work, these challenges are categorized into two broad aspects:

- Efficient approaches to model the user opinion and perform the corresponding adaptation during the interaction, i.e. in real-time.
- Methods to assess the subjective opinion of a user during the interaction, if possible automatically.

With the adaptation of dialogue systems being an active research field in general, the herein presented approaches make no claim of completeness. Nevertheless, both of the above aspects are addressed, if required in the context of a specific application scenario or setting. The first part of the chapter covers adaptation techniques by introducing a fine-grained model of the user opinion and an approach to utilize it for real-time adaptation of conveyed system emotions during interaction. The second part of the chapter is dedicated to the automatic recognition of the user opinion through an approach from the field of affective computing. Again, both parts build on approaches from previous chapters. In particular, the automatic assessment of the user opinion is based on the same experiment (and therefore uses the same data) as Section 6.1.

Due to the interdisciplinary nature of the addressed tasks, this chapter includes results of multiple cooperations: The discussed user model was developed with co-workers from both Ulm University and Augsburg University and is published in (Aicher et al., 2021; Rach et al., 2019b). The real-time adaptation to users is also a result of the cooperation with Augsburg University and published in (Weber et al., 2020a). The part on the automatic assessment of user opinions constitutes the main contribution of the chapter. It was conducted at NAIST and is published in (Rach et al., 2021a).

## 7.1 Modelling of User Preferences

The task of equipping dialogue systems with the capacity to adapt to an individual user was addressed in multiple domains and with multiple goals. Examples are an adaptation of referring expressions to match the knowledge of the user (Janarthanam and Lemon, 2010) and an adaptation of communication styles to increase the naturalness of the interaction (Miehle et al., 2018). Depending on the specific task, different information about the user is utilized and typically encoded into a user model. For argumentative tasks, the main information of interest is the user opinion regarding the overall topic and the aspects under discussion. This includes the general opinion before the interaction, i.e. the bias of the user, but also changes of this bias as a result of the interaction. The first step in the development of adaptation techniques for argumentative systems is hence to investigate models that represent the individual user opinion and can be dynamically updated during the interaction. This aspect is addressed in the following section through the introduction of a *preference model* in the context of two different application scenarios.

### 7.1.1 Formal Model

The general opinion of an individual towards a specific claim or topic depends on multiple aspects, i.e. different arguments concerning this topic. However, it has become clear from the experiments in Section 6.1 that being for or against a general topic does not imply that the corresponding person is in favour of all (or the majority) of the arguments that support his or her opinion. Instead, the opinion is arguably a result of an intuitive ranking of the known arguments, i.e. the assumption that although aspect $\varphi_i$ might be acceptable, it is less relevant or important than another aspect $\varphi_j$. This can be illustrated with the following example for the topic *animal testing* and the three aspects

- Animal testing violates animal rights ($\varphi_1$).
- Animal testing is required for an efficient testing of cosmetics ($\varphi_2$).
- Animal testing is important for the development of medications ($\varphi_3$).

Assuming that a person is now generally in favour of animal testing, it can still be the case that he or she rejects $\varphi_2$ and values the well being of animals used in $\varphi_1$ as more important than the testing of cosmetics ($\varphi_1 > \varphi_2$). However, the same person can also assume that the development of medication is required to save human lives, that this benefit out-weights the violation of animal rights ($\varphi_3 > \varphi_1$) and hence that he or she supports animal testing for this reason. A sophisticated model of the user's opinion thus has to take into account the different aspects and their influence on each other. On the other hand, for topics with a certain complexity, it is not practical to ask the user for a specific rating or judgment on each aspect and especially its effect on previous ones, as this is neither efficient nor user-friendly.

The approach proposed within (Aicher et al., 2021; Rach et al., 2019b) to address this issue is based on the abstract tree structure discussed in Chapter 4. It encodes the dependencies of argument components in bipolar relations between nodes in a graph. Thus, each component is related to one other component by either backing it up or attacking it. Consequently, it is possible to assign a level to each component. For an arbitrary component, the corresponding level is the number of connected nodes between it and the root (including the respective component itself). Throughout this chapter, components $\varphi_j$ that target another component $\varphi_i$ are called its child nodes and the set of all child nodes of $\varphi_i$ is denoted with $\mathbb{A}_i$. In its original form, the tree structure

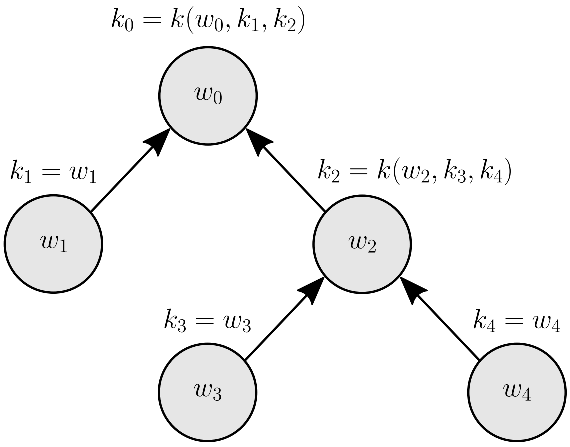


Figure 7.1: Sketch of the preference model with $k_i$ the strength of the respective node and $w_i$ the corresponding weight.

only captures the dependencies of components and makes no distinction in their strength, validity or weighting. To model the user's opinion, the structure is hence extended analogous to *bipolar weighted argumentation graphs* (BWAGs, see Section 2.2.1) in which a *weight* is assigned to each argument to determine its overall *strength*. In the preference model, the weights are updated incrementally, i.e. during the interaction with the corresponding system, based on (explicit or implicit) user feedback to represent the user opinion on the discussed aspects. It should be noted that the tree structure is, in contrast to a BWAG, formally comprised of argument components instead of arguments. However, the formal similarities between the two frameworks can be exploited to apply concepts of BWAGs to the herein considered tree structure. Similar to the discussion in Section 6.2, the (weighted) tree structure can be translated into a specific instance of a BWAG with each component $\varphi_i$ specifying a (unique) argument $\Phi_i$ of the form $\Phi_i = (\varphi_i \Rightarrow \varphi_j)$ (support) or $\Phi_i = (\varphi_i \Rightarrow \neg\varphi_j)$ (attack), respectively. In the following, the introduced approaches are discussed in the context of an argument structure with argument components to ensure compatibility with the previous chapters. However, they can be directly applied to general acyclic BWAGs as well without additional modifications.

Analogous to the formal definition of semantics for BWAGs in Section 2.2.1, the strength of an argument component $\varphi_i$ is then a function of its weight $w_i$ and the strength of its child nodes $\varphi_j, ..., \varphi_l \in \mathbb{A}_i$

$$k_i = k(w_i, k_j, ..., k_l). \tag{7.1}$$

Thus, the strength considers both the weight of the component itself as well as the implications arising from connected components. If a component has no child nodes, its strength equals its weight. For the sake of simplicity and without loss of generality strengths and weights are assumed to be real values in the interval $[0, 1]$. The idea is sketched in Figure 7.1. In the following, the notation $k(\varphi_i)$ is used analogous to Equation 7.1 to indicate the strength of a component $\varphi_i$. The approaches for computing strength values in the referenced literature assume an acyclic argumentation graph, which is ensured in the herein discussed scenario by the tree structure. For any model of this kind, a prediction of the users *stance* towards the topic under discussion is defined as follows (Weber et al., 2020a).

**Definition 21** (Stance Prediction)**.** *Let $k_0$ be the strength of the root component $\varphi_0$ in the utilized argument structure, then the user's stance regarding the topic under discussion is predicted as*

$$stan_{user} := \begin{cases} \text{pro} & k_0 \geq 0.5 \\ \text{con} & else \end{cases} . \tag{7.2}$$

The threshold of 0.5 is a direct consequence of the above stated assumption $k_i \in [0,1]\ \forall \varphi_i$ and can be adapted for strength values in different ranges. In addition, the definition of the prediction does not include an option *neutral*, as the value of 0.5 is assigned to a positive stance. However, the distance of the strength $k_0$ to the threshold value can be used to derive a *confidence* of the model in its prediction. The specific instantiation of this model, i.e. the update of component weights and strengths, depends on the respective application scenario and the therein available information.

### 7.1.2 Instantiations

In the following, two instantiations in different application scenarios are discussed in detail, thereby also including approaches to get the required user feedback to build the model. The first one is the cooperative setup the preference model was originally introduced for. It relies on explicit user feedback that is given as part of the interaction. In contrast, the second instantiation considers a multi-agent scenario based on the experimental setup discussed in Chapter 4. In addition to the application of the preference model, an implicit approach to assessing user preferences is discussed.

#### Opinion-building Domain

In (Aicher et al., 2021), we introduced the argumentative system BEA (Building Evidence-based Argumentation) in the opinion-building domain. The task of the system is to provide extensive information regarding the topic under discussion and to model the corresponding user opinion. Due to this goal, the system introduces all child nodes of the current component under discussion in a single system turn. Consequently, a component-wise rating by the user would be time-consuming and is hence not practical. Instead, we rely on *preferences* between the components, from which we compute a numerical hierarchy that is reflected in the weights.

The preferences are determined between components that are related to the same parent node, thus enabling incrementally forming the model. In the scope of the interaction, the user is able to express this preference regarding a specific aspect (or reject it). In addition, he or she is able to request new argument components, request information encoded in the preference model and end the discussion. The system outputs the response in the form of a written utterance whereas the user is able to select the desired response from a drop-down menu. The interface of the system is depicted in Figure 7.2.

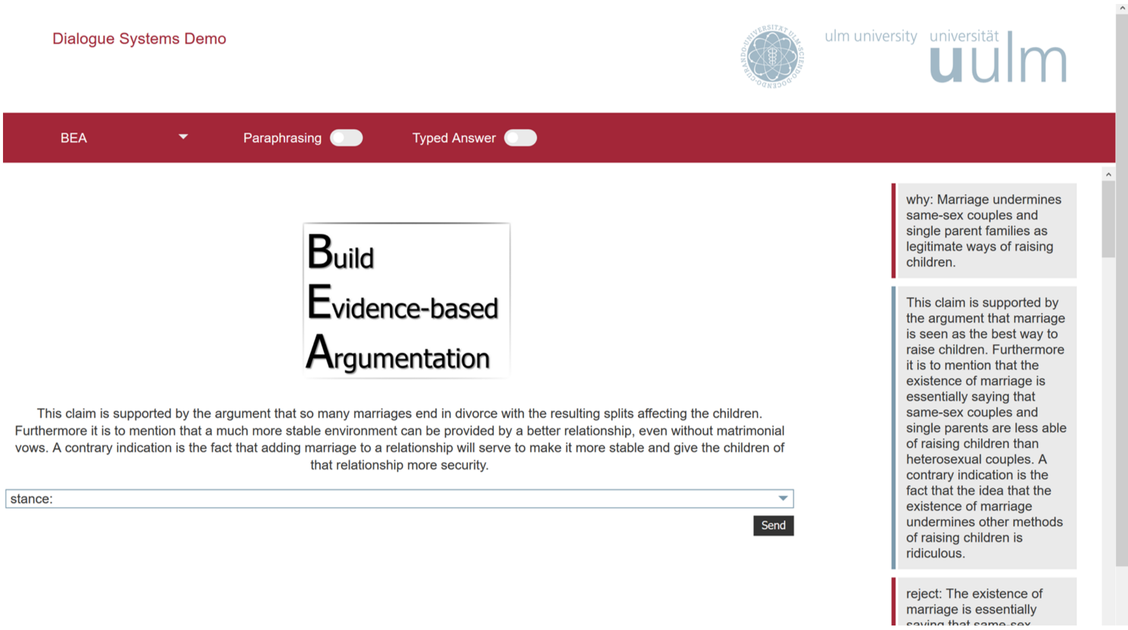


Figure 7.2: BEA interface for the example topic *Marriage is an outdated institution*, including dialogue system output, drop-down menu and dialogue history.

For updating the preference model, we adapt the Euler-based restricted semantics introduced for BWAGs in Section 2.2.1. To this end, the energy $E_i$ of an argument component $\varphi_i$ is defined as (Amgoud and Ben-Naim, 2018)

$$E_i = \sum_{\varphi_j \in supp(\varphi_i)} k(\varphi_j) - \sum_{\varphi_j \in att(\varphi_i)} k(\varphi_j), \tag{7.3}$$

where $supp(\varphi_i) \subseteq \mathbb{A}_i$ and $att(\varphi_i) \subseteq \mathbb{A}_i$ are the sets of supporting and attacking child nodes of $\varphi_i$. Hence, the stronger or more-numerous the supporting argument components are, the greater and more-likely-positive is the energy (and vice versa for attackers). The strength of an argument component $\varphi_i$ is then a function of its initial weight $w_i$ and the energy in Equation 7.3:

$$k_i = 1 - \frac{1 - w_i^2}{1 + w_i e^{E_i}}. \tag{7.4}$$

Consequently, $k_i$ considers the weight of the component itself as well as the influence from related components. As mentioned earlier, the model considers preferences between components related to the same parent node. After a preference is expressed, the corresponding weight is adjusted according to an update function and iterated through all connected arguments following Equation 7.4. Preferences can be expressed as either *prefer* or *reject*, where the first option means that the argument component is preferred over its siblings and the latter means that it is considered invalid. Consequently, the strength of a preferred component has to be greater than the ones of its siblings and is thus set to

$$k_i = k_{\max} + 0.5\,(1 - k_{\max})\,. \tag{7.5}$$

**Algorithm 3:** Preference Model Update

get *userFeedback* for current component $\varphi_i$
**if** *userFeedback is prefer* **then**
  get $k_{\max}$ from $\varphi_i$ and all siblings
  $k_i \leftarrow k_{\max} + 0.5\,(1 - k_{\max})$
**else**
  $k_i \leftarrow 0$
$E_i \leftarrow \sum_{\varphi_j \in supp(\varphi_i)} k(\varphi_j) - \sum_{\varphi_j \in att(\varphi_i)} k(\varphi_j)$
$w_i \leftarrow \frac{e^{E_i}(1-k_i)}{2}\left(-1 \pm \sqrt{1 + \frac{4k_i}{e^{2E_i}(1-k_i)^2}}\right) \in [0,1]$
**while** $\varphi_i \neq \varphi_0$ **do**
  $\varphi_i \leftarrow$ parent node of $\varphi_i$
  $E_i \leftarrow \sum_{\varphi_j \in supp(\varphi_i)} k(\varphi_j) - \sum_{\varphi_j \in att(\varphi_i)} k(\varphi_j)$
  $k_i \leftarrow 1 - \frac{1-w_i^2}{1+w_i e^{E_i}}$

Here, $k_{max}$ denotes the maximum strength of all siblings. As we formally update only weights in order to consider later preferences, the new weight is determined (solution of Equation 7.4) as

$$w_i = \frac{e^{E_i}(1-k_i)}{2}\left(-1 \pm \sqrt{1 + \frac{4k_i}{e^{2E_i}(1-k_i)^2}}\right). \tag{7.6}$$

Due to the square root, there are two solutions to this equation but only one is in the required interval $[0,1]$. If a component is rejected, its weight (and thus strength as well) is set to 0. The whole opinion model is updated after each expressed preference according to the Algorithm 3. As no information about the user's preferences is known before the interaction, weights are initialized with the same value ($w_i = 0.5$) upon introduction to the dialogue.

For the previously discussed example, the topic *animal testing* is the root ($\varphi_0$) and the three aspects $\varphi_1 - \varphi_3$ are its child nodes. In the introduced system, the corresponding preferences can be expressed by rejecting $\varphi_1$, and preferring $\varphi_3$. Consequently, the weight of $\varphi_1$ is then set to $w_1 = 0$. As all components are initialized with $w_i = 0.5$, the strength of $\varphi_3$ is updated as $k_3' = 0.5 + 0.25 = 0.75$. Inserting the strength into Equation 7.6 yields the new weight 0.75, which equals its strength due to the absence of child nodes. Consequently, the energy of the root node is $E_0 = 0.25$, resulting in a strength $k_0 = 0.54$ and the model hence correctly predicts that the corresponding user is in favour of the topic.

### Virtual Discussion

In addition to the previously discussed system, we propose an approach to integrate the preference model into a multi-agent setup where two virtual agents discuss controversial topics (Rach et al., 2019b), again in cooperation with Augsburg University. It is motivated by investigating the use of implicit feedback in an argumentative setup, including its estimation and combination with the preference model. Although the system is not fully implemented and evaluated in the proposed form, the corresponding considerations provide the basis for the estimation of subjective user opinions in Section 7.3. Moreover, the proposed application of the preference model can be directly transferred to the therein considered feedback.

The utilized multi-agent setup is based on the models discussed in Chapter 4, namely a dialogue game for argumentation to model the interaction and a tree structure to represent knowledge about the available arguments and their relations. Therefore, agents exchange arguments $\Phi_i$ that are built from the underlying tree structure. As discussed before, each included argument is again fully defined by the argument component that serves as premise $\varphi_i = prem(\Phi_i)$ which enables the direct application of the component-based preference model. To provide an intuitive interface, the envisioned system is represented by virtual avatars that interact via synthetic speech. In the scope of a virtual interaction, each agent takes a stance and tries to win the discussion in terms of the dialogue game. The user's emotional response can then be utilized to gain insight into the respective preferences and update the preference model accordingly. In order to focus on this aspect only, the user assumes the role of a (passive) audience. In the following, the proposed estimation of the user preferences based on emotion recognition is discussed, followed by the corresponding instantiation of the preference model.

An intuitive way to judge whether or not a user agrees with an opinion or a single aspect is looking at different multi-modal behavioural patterns, such as head nod or gaze. Users, for example, tend to nod if they agree. However, human behaviour is arguably too complex to get an opinion or preference by just looking at one or two of such patterns. More precisely, humans can be excited, bored, happy or disappointed, etc. These affects are part of the well-known circumplex model of affect (Russell, 1980). Nevertheless, taking all these different affects into account for user preference modelling is also challenging as the association of a specific emotional status with a preference is likely to be subjective. This is because people behave very differently and do not always show the same affects – not even to the same extent – for the same stimulus. Therefore, a summarizing approach is required that considers the different affects while at the same time enables to generally define whether a user feedback corresponds to the preference or rejection of an argument.

Matsuda et al. (2018) proposed a mapping for the emotional affects of the emotion status model (Figure 7.3) into the three classes *Positive*, *Neutral*, *Negative*, where *Positive* contains the emotional statuses *Excited*, *Happy/Pleased*, *Calm/Relaxed*, the *Neutral* class corresponds the status *Neutral* and *Negative* includes *Sleepy/Tired*, *Bored/Depressed*, *Disappointed*, *Distressed/Frustrated*, *Afraid/Alarmed*. The three classes can then be estimated in a supervised learning task from social cues or derived from an emotion recognition model that estimates valence-arousal values or affect states. To utilize this approach for learning the user's preference, we associate *preference classes* with each group:

- *Positive* → *Prefer:* The weight of the component is increased.
- *Neutral:* The weight of the component is not modified.
- *Negative* → *Reject:* The weight of the component is decreased.

Upon introduction, each argument component in the user model is initialized with weight 0.5 corresponding to the preference *neutral*. Whenever an argument is presented by one of the avatars,

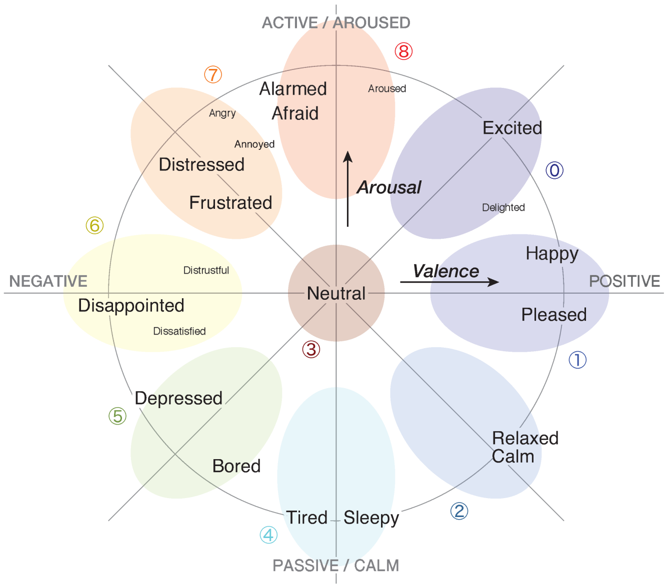


Figure 7.3: Two dimensional emotion status model based on Russell (1980) and taken from (Matsuda et al., 2018).

the emotional feedback is used to determine the respective preference class and the weight of the corresponding argument component is adjusted accordingly. The update formula for this scenario and an increase of the weight is defined as

$$w_i^{t+1} = w_i^t + \xi_p(1 - w_i^t). \tag{7.7}$$

and respective formula for a decrease is defined as

$$w_i^{t+1} = \xi_r w_i^t\,. \tag{7.8}$$

Here, $t$ denotes the temporal identifier of the corresponding turn in the dialogue game and $\xi_p, \xi_r \in [0,1]$ are weighting factors. The updates of the component strength is adapted from the previous scenario (Equation 7.4) and therefore again relies on the energy $E_i$ of the component.

For demonstration purposes, the example topic *animal testing*, weighting factors $\xi_p = \xi_r = 0.5$ and the following user feedback are considered:

- Animal testing violates animal rights ($\varphi_1$): *negative*
- Animal testing is required for an efficient testing of cosmetics ($\varphi_2$): *neutral*
- Animal testing is an important for the development of medications ($\varphi_3$): *positive*

The corresponding weights are then updated to $w_1 = 0.25$, $w_2 = 0.5$ and $w_3 = 0.75$. Consequently, the strength of the root node is updated to $k_0 = 0.59$ and the model (correctly) predicts the user stance as *pro*.

### 7.1.3 Discussion

In the following, the strengths and shortcomings of the previously introduced approaches are discussed. The introduced model enables an argumentative system to encode the opinion of users regarding the topic under discussion in a fine-grained and detailed way by considering not just individual aspects but also their effect on the opinion towards related ones. In addition, the (continuous) numerical scale of the weights and strengths also allows for encoding preferences between supported (or rejected) aspects on the same level. In the discussed scenarios, a tree structure was considered which assumes a unique relation of every argument component towards another. This assumption is reasonable when the argument structure is used as a knowledge base for argumentative systems since it prevents the repetition of aspects. However, in the context of the preference model, it can be a limitation as the effect of a component on the user opinion towards another, not related component is not captured (unless this other component is indirectly connected to the current one on a higher level of the tree). nevertheless, the introduced model does not depend on the assumption of unique relations but only on the condition of directed, acyclic graphs and is therefore capable of considering additional relations (if they are known and included in the structure).

In addition, the capability of the model to predict the user stance was demonstrated on a small example with two levels. Although this illustrates the general functionality of the model, an extensive evaluation of its capability to correctly predict the user stance is required. This evaluation is addressed alongside an adaptation approach that utilizes the model in the following section. Similar to the first discussed instantiation in the opinion building domain, the evaluation depends on explicit user feedback, which is reasonable in the desired context to minimize errors.

Nevertheless, actual applications of the model will rely on implicit feedback, especially in competitive scenarios where the user actively participates in the discussion and explicit feedback hinders the interaction. To this end, an approach based on emotion recognition was discussed in the context of a multi-agent scenario. Whereas the strength of the reliance on emotion recognition techniques is the extensive research in this area to build on, it also faces some conceptual issues in the current context: Even if the emotion can be predicted with high accuracy, the mapping of emotions into preferences is prone to errors that arguably compensate the benefits of an individual user model. Therefore, Section 7.3 introduces an approach to assess the user opinion directly and implicitly from non-verbal cues that are shown by the individual user during the interaction. Although it builds on the herein discussed concepts in the sense that it focuses on summarizing aspects of the user opinion, it does not rely on the explicit recognition of emotions or their mapping into preference classes. However, the herein proposed use of the preference model can be applied to this direct estimation as well.

## 7.2 User-adaptive Presentation of Arguments

The previous section was focused primarily on modelling the user opinion through updating the component weights in the argument graph structure. The next step is hence to utilize this information for adaptation, i.e. to change the system behaviour according to the modelled opinion. Since adaptation depends on the goal of the system, the herein investigated approach is discussed in the context of the application setup it is built for, and an evaluation system tailored to the specific task is introduced. In addition to the discussion and a proof-of-principle evaluation of the adaptation approach, the system is also utilized to evaluate the capacity of the underlying user model to correctly predict the user stance.

The considered adaptation is focused on the emotions displayed by the system with the aim to make the human audience aware of the subliminal effects of an emotionally appealing presentation. The approach consequently does not consider the selection of the argumentative content but only the emotions with which it is displayed. To minimize the effect of a pre-defined user opinion, i.e. a bias, the system utilizes an argument structure from the hotel review domain acquired with the semi-automatic procedure discussed in Section 6.3. The system then uses an instantiation of the previously discussed preference model to determine its persuasive effectiveness and adapts the displayed emotions according to an emotional policy learnt by reinforcement learning during the interaction. The section introduces the system architecture first and subsequently discusses the adaptation approach and the utilized instantiation of the preference model. The last part covers the evaluation, including setup, results and their discussion.

The work in this section was done in cooperation with Augsburg University and published in (Weber et al., 2020a). Major parts of the work, namely the development of the adaptation approach, the system implementation and the evaluation were conducted by co-workers at Augsburg University. They are herein reported in full for the sake of completeness and to provide the necessary foundation for the discussion of system implementations in Chapter 8.

### 7.2.1 System

To enable an application of the investigated approaches in direct interaction with users, an evaluation system is proposed. The system is inspired by the one introduced in Section 6.1 for the evaluation of argument search engines and is visualized in Figure 7.4. The system includes a conversational agent that is again represented by a virtual avatar and presents the system utterances via synthetic speech and multimodal emotions (mimics and gestures). The agent's task is to talk about the topic encoded in the root node of the utilized argument structure by providing argument components that are either *for* or *against* the topic. To do so, the system first assigns one of the two *stances pro* or *con* to the agent. Then, the agent presents both pro and counter components from the argument structure in different ways to convince the user to either change their belief (if the user is against the agent's stance) or to increase their beliefs towards the corresponding stance. To influence the user, the agent can underlie its utterance with appropriate emotions. At each turn, a random argument component $\varphi_i$ is selected from the subset of components that have a direct link to any component introduced before. Then, an appropriate emotion is selected depending on the agent's policy $\pi$ and presented to the user in combination with the natural language utterance that includes the selected component.

Similar to Section 6.1, the utterances are generated by a NLG that relies on the annotated sentences in the argument structure. In addition, connecting phrases between the arguments were included in a separate template to ensure a fluent interaction. These additional statements include a notification about the stance of the presented argument component (if it attacks its target), a notification if the presented argument component refers not to the immediate predecessor in the conversation, as well as an introduction and closing statement. Apart from the introduction and

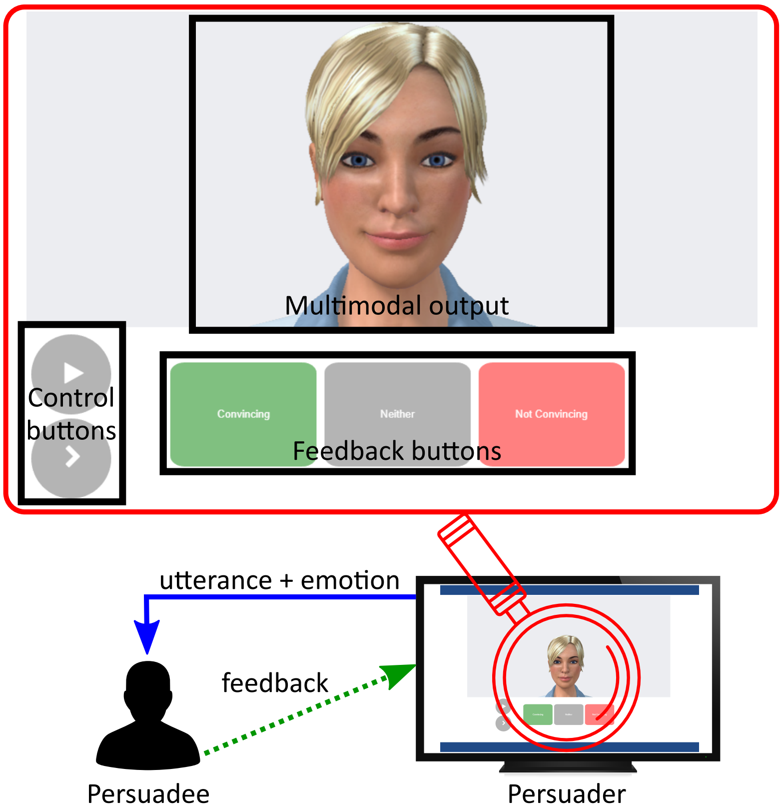


Figure 7.4: Illustration of our web interface consisting of an agent presenting her arguments to a user along with emotional behavior. The user gives feedback (*convincing*, *neutral*, *not convincing*) after each introduced argument component, which is used to train the strategy.

closing statement, multiple formulations for each case were included in the template from which the system selects randomly for each utterance. The following exemplary utterance includes a topic switch (ts), the referenced component ($\varphi_1$), a notification about the stance (st) and the new argument component ($\varphi_2$): *"The next argument is related to something I mentioned earlier. I said (ts): All in all, it is a nice and affordable spot for sightseeing in the area ($\varphi_1$). I also found an opinion that disagrees with this aspect (st). The respective author wrote: I think all in all the price was way too high for such a poor accommodation ($\varphi_2$)."*

After each utterance, the user provides the agent with explicit feedback using the feedback buttons (*convincing*, *neutral*, *not convincing*) as illustrated in Figure 7.4. This feedback is subsequently used to determine the agent's persuasive effectiveness and to learn the policy $\pi$ that works best for the current user.

To present the system utterances via multimodal output to the user, we employ the Charamel™ 3D character rendering engine[1]. We herein use the Gloria avatar capable of performing social-

[1]https://www.charamel.com/competence/avatare (last accessed 29 August 2021)

Figure 7.5: Example emotions: Angry, happy, and sad (from left to right).

cue based interactions with the user as illustrated in Figure 7.4. The avatar can perform lip-sync speech output using the Nuance TTS along with the Amazon Polly voices[2]. Further, the avatar comes with more than 50 motion-captured gestures as well as a set of 14 facial expressions, including the basic emotions defined by Ekman (1992). Some examples of the avatar's possible emotions are given in Figure 7.5.

### 7.2.2 Adaptation Approach

In the following, the adaptation approach is described in detail. It consists of two components, namely the RL method to optimize the agent strategy with respect to user feedback and the preference model instantiation to predict the user stance. As stated by Ritschel et al. (2017) and Weber et al. (2018) an adaptation to human preferences should work during the interaction. Thus, some simplifications are needed for our system in terms of state and action space to ensure that the underlying learning problem has a complexity that allows for real-time adaptation.

#### Reinforcement Learning-based Adaptation

To enable the agent to adapt its multimodal output to the user's preferences during the interaction, we apply reinforcement learning based on linear function approximation as described in Section 2.3.2 along with a basis transformation using the Fourier basis as proposed by Konidaris et al. (2011). This approach has three advantages:

- Reinforcement learning allows for learning a sophisticated policy $\pi$ based on trial and error (during interaction).
- A linear function approximation bears the advantage that multiple similar states can be learned at the same time while allowing for quick adaptation compared to non-linear methods.
- Using the Fourier basis allows for learning explicit linear and non-linear dependencies between state-input parameters.

[2]https://docs.aws.amazon.com/polly/latest/dg/voicelist.html (last accessed 29 August 2021)

Every RL state $s \in S$ is determined based on the current argument component $\varphi \in L_t$ with $L_t$ the topic language, i.e. the set of all components in the argument structure. As aforementioned in Section 6.3, each component has an annotated polarity, in the following referred to as *stance* $\in \{pro, con\}$ as well as a *relation* $\in \{attack, support\}$. To enable the agent to further optimize its behaviour with respect to the conveyed sentiment of the utterance, we use the sentiment analysis tool *vaderSentiment*[3] by Hutto and Gilbert (2015) to compute information about the negativity, neutrality and positivity and the respective compound score of an argument component. A state $s \in S$ for an argument component $\varphi \in L_t$ is then defined as:

**Definition 22** (State Space $S_{emotion}$)**.** *Let $\varphi_i \in L_t$ be the argument component chosen at time step t by the system, stan : $L_t \rightarrow \{pro, con\}$ the stance function that assigns each argument component to its stance towards the root node and rel($\varphi_i$) $\in \{attack, support\}$ the relation of $\varphi_i$ towards its target. Further let score : $L_t \rightarrow [-1,1]$ be a normalized compound score of the sentiment analysis of a component. Then the state $s_{emotion}$ is defined as follows:*

$$s_{emotion} := (stan(\varphi_i), rel(\varphi_i), score(\varphi_i)). \tag{7.9}$$

As described above, we make use of the agent's provided emotions. Thus, the discrete action space $A$ consists of different emotions (facial expression and gestures), such as *happy*, *sad*, *angry*, *disappointed*, etc., that can be displayed by the agent with different discrete intensities $\in \{low, medium, high\}$.

### Prediction Model

As aforementioned, the user in the considered setup assumes the role of the audience and provides feedback $f(\varphi_i)$ for each introduced component. To ensure a fluent interaction and interrupt the discussion in a minimal way, the feedback includes the three options *convincing*, *not convincing* and *neutral*. It is translated into a numerical value by the feedback function $f : L_t \rightarrow \{1.0, 0.5, 0.0\}$, as:

- *Convincing*, i.e., positive feedback ($f$ = 1.0)
- *Neutral* ($f$ = 0.5)
- *Not convincing*, i.e., negative feedback ($f$ = 0.0)

Since the goal of each agent is to convince the user of the assigned *stance* the feedback signals cannot be used directly as reward. Instead, it has to be inverted respectively, if the argument's stance does not match the agent's stance:

$$inv_{\varphi}(f(\varphi)) = \begin{cases} 1.0 - f(\varphi) & if\ stan(\varphi) \neq assigned\ stance \\ f(\varphi) & else \end{cases}. \tag{7.10}$$

Even though, in theory, using $inv_{\varphi}(f(\varphi))$ as a reward would generally allow the agent for adaptation, this approach does not enable the agent to observe to what extent the goal of convincing

[3] https://github.com/cjhutto/vaderSentiment (last accessed 29 August 2021)

the user of the *assigned stance* has already been achieved compared to the *non-assigned stance*. In addition, it does not consider the effect of relations between the individual components and their distance to the root. However, as argued in the previous section, the opinion of the user does not only depend on how many components could be weakened/strengthened by others but also on the relations between the components.

To account for these aspects, the reward is derived from an instantiation of the previously introduced preference model that computes the strengths of argument components considering their own weight $w$ and the strengths $k$ of their child nodes. As the explicit user rating in the present scenario is absolute, the weights for each argument component are assigned as follows:

$$w_i = \begin{cases} f(\varphi_i) & \textit{component } \varphi_i \textit{ used by agent} \\ 0.5 & \textit{else} \end{cases}. \tag{7.11}$$

It can be seen that the strength update in Equation 7.4 is not suitable for the present task, as components with a weight equal to zero are considered invalid and, thus, their children have no effect on their strength. To mitigate this problem, an alternative instantiation of the preference model is proposed. To this end, the energy of an argument component $\varphi_i \in L_t$ is re-defined as

$$E_i = \sum_{\varphi_j \in supp(\varphi_i)} k(\varphi_j) + \sum_{\varphi_j \in att(\varphi_i)} 1 - k(\varphi_j), \tag{7.12}$$

where $supp(\varphi_i)$ and $att(\varphi_i)$ again denote the sets of supporting and attacking child nodes of $\varphi_i$. The strength $k_i$ of the argument component $\varphi_i$ is then computed as

$$k_i = \frac{w_i + E_i}{1 + |\mathbb{A}_i|}. \tag{7.13}$$

It can be seen from equation 7.11 that each weight in the argument structure is initialized with 0.5, which consequently yields $k_0 = 0.5$. During the interaction, the strength of each component is re-computed from bottom to top recursively after each feedback through the user as discussed in the previous section. As a direct consequence of the above definitions, a neutral user feedback ($f$ =0.5) does not change the strength of the root component $k_0$ from which the user's current stance can be predicted according to Definition 21. In the context of the adaptation (learning) problem, the strength of a component $\varphi_i$ at a dialogue time step $t$ inverted with respect to the agent stance is used as a measure of its *effectiveness*

$$e_{i,t} = inv_{\varphi_i}(k_{i,t}). \tag{7.14}$$

Based on the current effectiveness, the reward $r : S \times A \to [0,1]$ for the reinforcement learning-based adaptation is defined as:

**Definition 23** (Reward Function). *Let $s_t \in S$ be the current state and $a_t \in A$ the action at RL time step $t$. Further, let $\boldsymbol{e}_{0,t}$ be the current effectiveness after performing action $a_t$ and $\boldsymbol{e}_{0,t-1}$ the previous one. Then, the reward $r_t(s_t, a_t)$ at time step $t$ is defined as*

$$r_t(s_t, a_t) := n(e_{0,t} - e_{0,t-1}). \tag{7.15}$$

Initial tests have shown that $n \approx |L_t|$ works best for the herein considered learning task.

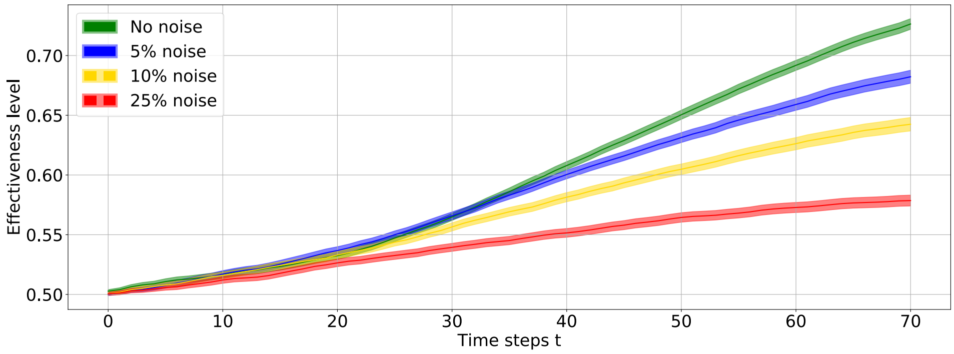


Figure 7.6: Simulation results including 95% confidence intervals showing the cumulative effectiveness level over time with respect to the assigned stance. A continuous increase over time can be observed in all cases and even for high noise.

### 7.2.3 Evaluation

We conducted two separate evaluations with the proposed system. The first one is intended to demonstrate the general feasibility of the reinforcement learning-based adaptation with respect to different users and to investigate the number of epochs that are required for efficient learning. The second evaluation assesses the capability of the utilized preference model to correctly predict the user stance.

#### Numerical Simulation

To demonstrate the general feasibility of the RL approach and to show that our system can handle different types of users, we tested our system in a simulation setup. We have run a simulation of 1,000 simulated users with different behaviour preferences as well as stances (either *for* or *against* the agent's stance) that work best to persuade them individually, e.g., users where the agent was the most effective when looking sad while presenting negative arguments that are against its own stance, and happy when presenting positive arguments that are in favour to the agent's stance and vice versa. Some simulated users preferred higher intensities of the emotions, while others preferred lower ones. Further, we varied the agent's *assigned stance* ensuring that all possible combinations are tested.

Since deterministic user feedback is far from realistic (Rieser and Lemon, 2011), we have run the same simulation with different levels of noise (5%, 10% and 25%), where noise simulates random user feedback, which not necessarily matches the optimal policy $\pi^*$.

Figure 7.6 depicts the results showing the effectiveness levels over time with different degrees of noise where the shaded areas depict a 95% confidence interval. The results reveal that the effectiveness levels increased over time, as the agent managed to gradually move the user towards its assigned stance by learning the most effective behaviour strategy. Even though the performance decreases due to non-deterministic feedback when including noise, it shows that the agent is still able to cope with such non-deterministic feedback and find the strategy that increases the effectiveness level.

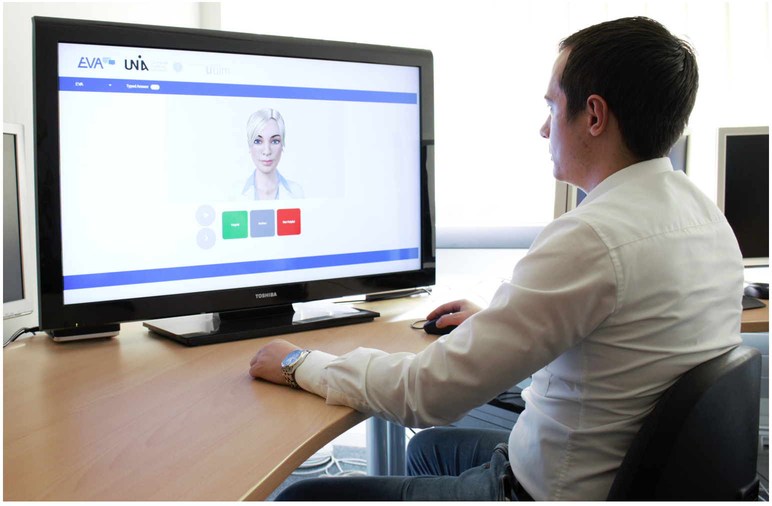

Figure 7.7: Evaluation setup with our interactive agent.

### User Study

To evaluate the utilized prediction model, i.e., to verify that our model is able to correctly predict the user's stance, we conducted a between-subject user study. We recruited 48 participants (32 male, 16 female, 18-30 years old) from the Augsburg University campus. All participants were students. At the start of the study, they were informed about the general procedure and asked to provide the agent with feedback using the interactive system of Figure 7.4. After the session, they were asked, whether or not they like to visit the hotel. To avoid bias effects beforehand, they were not told about the system's overall goal to predict their decision but asked to provide feedback on whether or not they find an utterance helpful. Figure 7.7 depicts the general study setup showing a participant interacting with the agent.

In each session, the agent presented 43 arguments for and against the hotel, which took about ten to 15 minutes. The agent computed the strength $k_0$ based on the given feedback $f(\varphi)$ for every presented argument $\varphi \in L_t$. The agent's *assigned stance* was counter-balanced, i.e., half of the participants were interacting with an agent who was in favour of visiting the hotel and vice versa. During the study, we collected the following data:

1. Directly given user feedback $\boldsymbol{f}(\varphi), \forall \varphi \in L_t$.
2. Computed strength $k_0$ using the given feedback.
3. Subjective decision if users like to visit the hotel (post-study).

### Results

In the following, we first plot the collected data to explore trends and afterward present statistical tests. Figure 7.8 summarizes the results for all participants depicting the final strength $k_0$, the percentage of *user feedback* in favour of the agent's assigned stance $f^+$ and the percentage of

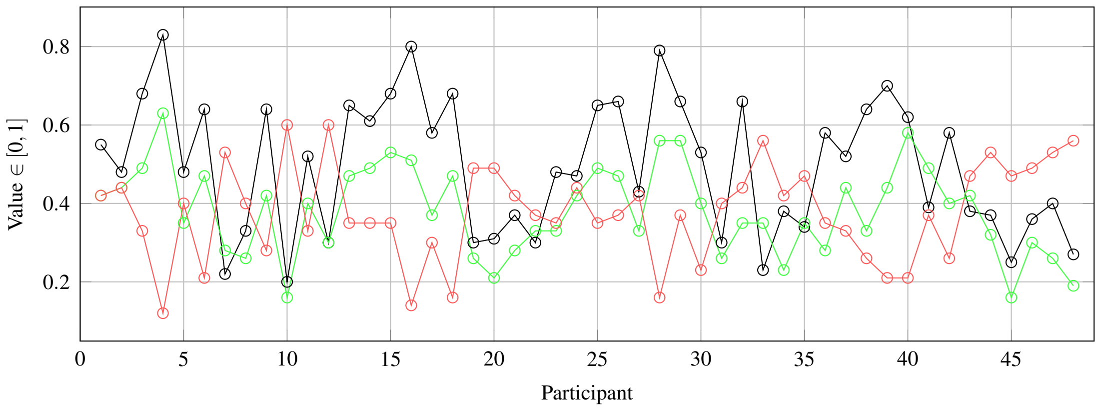


Figure 7.8: Collected data: Strength $k_0$, percentage of positive (negative) feedback $f^+$ ($f^-$) with respect to the assigned stance.

Table 7.1: Correlation between feedback and strength $k_0$.

| | n | PCC | p |
|---|---|---|---|
| Pos. Feedback & strength $k_0$ - Pearson correlation | 48 | 0.92 | <.001 |
| Neg. Feedback & strength $k_0$ - Pearson correlation | 48 | -0.83 | <.001 |

*user feedback* not in favour of the agent's assigned stance $f^-$. Neutral feedback is not depicted as it does not affect the overall strength of the root node (see Section 7.2.2). First, we notice two trends:

1. The higher (lower) the positive feedback, the higher (lower) the strength $k_0$.

2. The lower (higher) the negative feedback, the higher (lower) the strength $k_0$.

Thus, the positive feedback $f^+$ seems to positively correlate with the strength and the negative feedback $f^-$ seems to negatively correlate with the strength $k_0$. As stated, the general idea of assessing the strength $k_0$ is to get a prediction of the user's current stance regarding the topic under discussion. So, we expect that a lot of positive feedback increases this strength value, while a lot of negative feedback decreases it. The trends, therefore, are in line with our expectations.

To verify the expected trends statistically, we computed the correlation between feedback and $k_0$ using the Pearson correlation coefficient (PCC, see Section 2.3.1). The corresponding results are shown in Table 7.1 and show a very strong and significant correlation (*positive correlation for positive feedback, negative correlation for negative feedback*).

We then evaluate to what degree the predicted user's stance $stan_{user}$ (see Definition 21) and the subjective user's decision to visit the hotel match. We compute the agent's confidence in the predictions based on how close they were on the threshold value ($k_0 = 0.5$). To this end, we utilize a modified sigmoid function to ensure that the extreme values of 0 and 1 correspond to

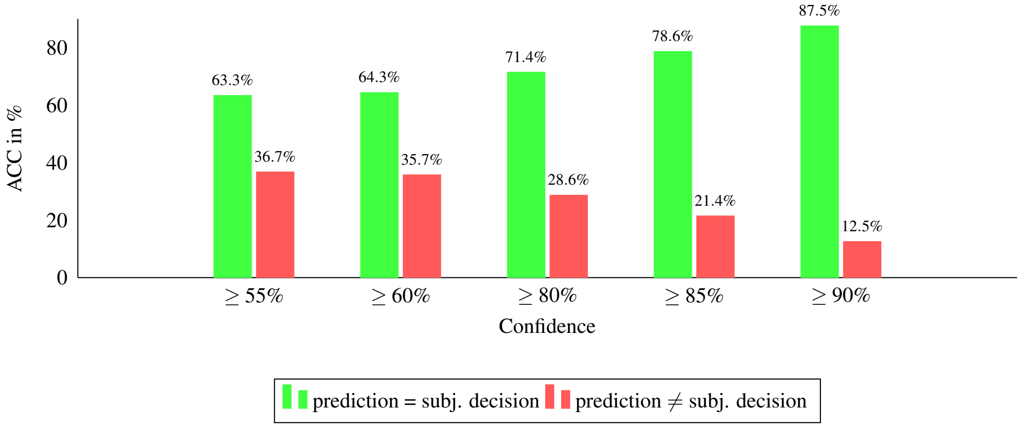


Figure 7.9: Accuracy of predicted user's stance depending on different confidence values.

Table 7.2: F1 score for different prediction confidences.

| | Confidence | | | |
|---|---|---|---|---|
| Stance | ≥ 60% | ≥ 80% | ≥ 85% | ≥ 90% |
| pro | 0.67 | 0.73 | 0.77 | 0.86 |
| con | 0.62 | 0.70 | 0.80 | 0.89 |

a confidence of 100% and obtain a more fine-grained prediction for the most common interval $[0.3, 0.7]$. Consequently, a confidence $\geq 80\%$ means $k_0 \geq 0.64$ or $k_0 \leq 0.36$. The accuracy (ACC) results in Figure 7.9 show that the objective system's prediction is very accurate even for low confidence values, hence indicating the practical potential of the model.

To verify the sensitivity and precision of the predictions, we compute the F1 score for both the positive and negative stance depending on the prediction confidence as summarized in Table 7.2. The results show that the F1 score increases with higher confidences, thus, proving both the sensitivity and precision of the prediction.

### 7.2.4 Discussion

We have presented an adaptive virtual agent capable of learning a behaviour strategy during interaction with a human to increase its perceived persuasive *effectiveness*. The therefore proposed approach utilizes the previously introduced preference model to obtain predictions regarding the current stance of the user and the *effectiveness* of the agent. The adaptive feasibility of the approach was shown in a numerical simulation beforehand. To verify the predictive power of the utilized preference model, we additionally presented a thorough evaluation in a user study with 48 participants. We have found a very strong and significant correlation between feedback and predicted strength $k_0$ and have shown that our system is able to correctly predict the user's stance. Both observations indicate the validity and practical potential of the model.

At the beginning of the chapter, we argued that different argument components can have different effects on the user's stance based on their position in the argument structure. Despite the observed correlation between overall feedback and predicted strength at the end of the interaction, using the proposed fine-grained preference model for predicting the user stance and in the reward function bears several advantages in comparison to models based solely on the statistics of the feedback:

- The feedback for different components is weighted differently in the graph and, hence, differently affects the behaviour strategy during the interaction.
- Additional argument-specific or structure-specific information can be used in combination with the feedback to provide more fine-grained information for learning.
- The proposed learning of multimodal behaviour can be combined with fine-grained logical strategies discussed in Chapter 5.

An approach that combines a previously trained logical strategy with the herein proposed adaptation to user feedback is presented in Chapter 8.

The given user feedback during the study does not necessarily lead to their final decision for or against the hotel. Thus, it would not have been surprising if the predictions of the model were incorrect. However, the fact that we are able to predict the user's current stance using the provided feedback in an agent-user interaction opens a lot of new possibilities for the development of persuasive systems. For instance, the predicted stance can be used to determine when the user is likely convinced and the system can stop the persuasion process. Secondly, in persuasive discussions between several agents/humans, the predicted stance can be used to determine the success of the whole dialogue during interaction and, thus, enables the agents to adopt opponent strategies that are more successful.

However, the discussed application potential also emphasises a current limitation of the proposed approach, namely its reliance on explicit user feedback. Whereas the interface-based rating of individual argument components provides a robust and practical feedback signal for evaluation, actual applications in which the user assumes a more active role requires an implicit assessment of the user opinion regarding the discussed content to enable a fluent and natural interaction. This is addressed in the next section through an automatic estimation of the subjective user opinion regarding individual arguments from non-verbal cues.

## 7.3 Estimating Subjective Argument Quality Aspects from Social Signals

The last section of this chapter addresses the previously discussed need of adaptation approaches for an implicit assessment of the user opinion. It is required in all adaptation approaches that focus on the subjective user perception of arguments and cannot rely on explicit feedback. Although explicit feedback can be used in some instances (as discussed before) it requires the user to assume the role of a judge or audience, i.e. he or she cannot participate in the actual discussion.

In (Rach et al., 2021a), we introduce an approach to address this issue through estimating subjective quality aspects of arguments presented by an argumentative dialogue system from social signals shown by the user during the presentation. The estimated aspects are the users *interest* in an individual argument as well as its perceived *convincingness*. Both aspects were already used in Chapter 6 to assess the suitability of arguments retrieved by means of argument search engines for completing argumentative tasks. In addition, the category *convincing* was also used for adaptation purposes in the previous section. The assessment was in both instances done through explicit ratings by users given after the system utterance was presented by means of synthetic speech and by a virtual avatar. This section builds upon the approaches and results of Chapter 6 by using the explicit user ratings as labels for each argument and estimating them from social signals shown by the corresponding participant during the interaction.

The proposed method is along the line of affective computing approaches that aim at estimating subjective information like interest (Tomimasu and Araki, 2016) or agreement (Bousmalis et al., 2013) from social cues. Some approaches utilize a mapping of the observed cues to a psychological model of emotions and derive the final label from the recognized emotion, as for example in (Yeasin et al., 2006). This was also considered for the present task in Section 7.1.2. However, as the association of emotions with a specific opinion is likely to be subjective as well and this mapping is hence prone to errors, the original approach is modified to directly estimate the quality labels from the observed cues. The utilized social signals include eye movement and facial expressions extracted from video recordings of the interactions. In addition, we perform a human annotation for both aspects on a (randomized) representative test set to compare the automatic estimation to human performance. The comparison shows similar performances with a slight advantage of the human annotation in the *convincing* task whereas the automatic estimation performs slightly better in the *interesting* task. In summary, the contributions of this section are:

- Providing the (to the best of our knowledge) first recognition of argument quality aspects based on individual social signals.
- Discussion of the contributing features and their importance.
- Comparison of this estimation to human performance.

As the work discussed in the following builds on the evaluation categories and the data collection covered in Section 6.1, a concise summary of the utilized material and concepts is included if required for the present section.

### 7.3.1 Evaluation Categories

In the following, the selection of the two categories that are estimated in the presented setup is discussed. Similar to Chapter 6, a search result from an argument search engine is denoted as argument and its polarity towards the overall topic as *stance* (pro or con).

From the four evaluation categories introduced in the previous chapter, we now focus on the task-related ones (*convincing* and *interesting*) and estimate the respective user rating from social signals. This choice is due to several reasons: First of all, for both categories, similar approaches in the field of affective computing exist, namely the estimation of (dis)agreement and interest in other domains. Therefore, it is likely that users show distinct non-verbal cues which can be used to estimate the corresponding rating. Secondly, the categories *related* and *comprehensible* are meant to assess technical differences between argument search approaches and are less likely to be influenced by their presentation through the system. In contrast, the user perception of the categories *convincing* and *interesting* can also be changed by different presentations of the argument. In the case of the *convincing* category, the presentation of the argument was discussed as a quality (sub)dimension in (Wachsmuth et al., 2017b) and an approach to increase the convincingness of an argumentative dialogue system by adapting for example the behaviour of the presenting virtual avatar were already introduced in the previous section. Similarly, the user's interest and his or her engagement in the interaction with a conversational system was shown to be influenced by the system's presentation of the content (Ritschel et al., 2017). Consequently, the information whether or not a user is convinced by or interested in a presented argument can be used for adapting the presentation of the following utterances whereas the information about a correct/incorrect relation and the comprehensibility of an argument cannot be transferred to the next one. Also, since the perception of all categories is highly subjective, the information about the relation and the comprehensibility cannot necessarily be used in following interactions with different users. In summary, we focus on the estimation of the categories *convincing* and *interesting* because we expect indicating signals from users based on existing work and because the estimated information can directly be used for system adaptation.

### 7.3.2 Data Collection

The evaluation experiment was conducted with 19 international students (14 male, five female) at NAIST and approved by the associated Ethics Review Committee. In the following, the relevant aspects regarding the experimental setup are recalled from Section 6.1. In addition, details that are only relevant for the task of the present section are provided. Before the experiment started, participants received instructions about the overall purpose of the experiment, the system interface and the meaning of the categories. In addition, a short test trial with a separate topic was offered to exemplify the procedure. During the experiment, participants were seated on a chair, facing a 27inch display that showed the system in full-screen mode. A camera was placed on top of the display to record the user reactions throughout the experiment. A picture of the complete setup is shown in Figure 7.10.

Participants listened to arguments presented by the virtual avatar and gave quality ratings in the four categories *convincing*, *interesting*, *comprehensible* and *related*. Each participant listened to arguments related to the four topics *nuclear energy*, *animal testing*, *self-driving cars* and *death penalty* with a time limit of five minutes per topic after which the system stopped the rating process and introduced the next topic. The arguments for the first three topics were retrieved with one of the three argument search engines, whereas the arguments for the topic *death penalty* included arguments collected with both ArgumenText and args.me and were the same for each participant.

As reported in Section 6.1, the ratings for the first three topics were used to compare the utilized argument search engines whereas the ratings for the fourth topic were used to analyse the agreement between the participants in each category by means of Krippendorff's alpha (Hayes and Krippendorff, 2007). The found agreement was low for all categories, with the highest alpha of 0.15 in the *comprehensible* category. This again indicates the subjective nature of the task and emphasises the need for the herein proposed approach.

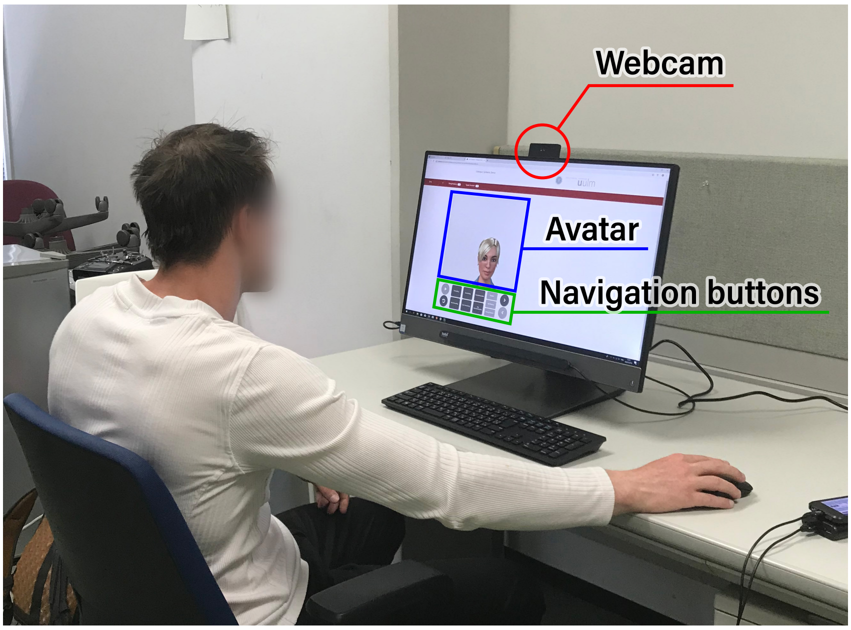


Figure 7.10: Experimental setup.

Besides the rating of the arguments, participants answered two questionnaires (one before and one after the interaction). In the first survey, the participants' stance on the discussed topics and their proficiency regarding dialogue systems were assessed, whereas the questions after the interaction were concerned with the understandability of the system with an emphasis on synthetic speech and language skills of the participants. In case a participant reported substantial difficulties with the language, the corresponding ratings were excluded from the data set (this occurred only once). Although several participants reported that they were sometimes irritated by the synthetic speech, no participant reported that he or she was not able to understand the presented content. The corresponding questionnaire ratings are shown in Figure 7.11 and were also discussed in Section 6.1. Based on these results, we assume that the reactions shown by the users are mainly due to the presented content. Nevertheless, the reported irritation caused by the synthetic voice may introduce some noise for the recognition of the user opinion.

### 7.3.3 Data Statistics and Feature Extraction

The experiments resulted in a total of 1,263 ratings in the two categories *interesting* (653) and *convincing* (610) and approximately six hours of recorded interaction. The difference in the number of ratings for both categories is due to the fact that participants were able to skip a rating in one or more categories if they were undecided. The splitting into the two categories and the respective class balance (positive and negative response) is shown in Table 7.3. We see that in the *convincing* case, the distribution of positive and negative ratings is almost balanced, whereas in the *interesting* case the positive ratings outweigh the negative ones.

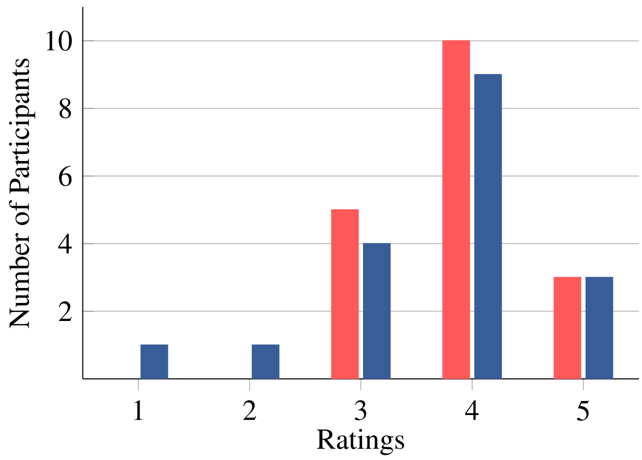


Figure 7.11: Responses on a five-point Likert scale from completely disagree (1) to fully agree (5) for the statements *The synthetic speech was easy to understand* (blue bars) and *All in all I had no problems understanding the system utterances* (red bars).

Table 7.3: Statistics of positive and negative ratings in the recorded data for the two categories *interesting* and *convincing*.

| | Convincing | Interesting |
|---|---|---|
| Positive | 287 | 506 |
| Negative | 323 | 147 |
| Total | 610 | 653 |
| Ratio | 0.47 | 0.77 |

For the classification approach, we assume that the non-verbal signals that correspond to the ratings are shown during or directly after the presentation of the argument since participants are focused solely on the content of the arguments during this time. Consequently, we investigate the time window between the user request for a new argument and his or her first rating for it. If the user listens repeatedly to the argument, the whole time window until the first rating is included. We call a pair of rating and time window a session and denote the time window corresponding to a specific session $i$ with $win_i$.

Due to the passive role of the participants in our setup, we assume that the most informative cues are shown on the participants' faces and focus on social signals related to facial expressions and eye movement. To derive a set of meaningful features for the classification, we extract the intensity of the 17 facial action units (AUs) shown in Table 7.4 and eight eye movement values (x, y, z and angle for both eyes). The extraction is done by means of the OpenFace toolbox (Baltrusaitis et al., 2018), resulting in a data point $q_t = (AU_{1-17,t}, eye_{1-8,t})$ with the dimension $dim = 25$ for each video frame.

Table 7.4: Action units extracted by the OpenFace toolbox (Baltrusaitis et al., 2018), including number and description.

| Action Unit | Description |
|---|---|
| AU01 | Inner Brow Raiser |
| AU02 | Outer Brow Raiser |
| AU04 | Brow Lowerer |
| AU05 | Upper Lid Raiser |
| AU06 | Cheek Raiser |
| AU07 | Lid Tightener |
| AU09 | Nose Wrinkler |
| AU10 | Upper Lid Raiser |
| AU12 | Lip Corner Puller |
| AU14 | Dimpler |
| AU15 | Lip Corner Depressor |
| AU17 | Chin Raiser |
| AU20 | Lip Stretcher |
| AU23 | Lip Tightener |
| AU25 | Lips Part |
| AU26 | Jaw Drop |
| AU45 | Blink |

Based on all data points $\{q_t\}$ within a time window $win_i$, we compute the following statistical features along all 25 dimensions for each session $i$:

- Standard deviation (std)
- Mean value (mean)
- Area under the curve value (auc)

Since the data points of the action units represent intensities, peaks in the respective time series indicate how clear (and how often) the corresponding cue was shown. To include this information into the classification process, we additionally compute the following two features related to maximum values for each action unit intensity:

- Maximum value (max)
- Number of peaks higher than the mean value (peaks)

This results in a feature set with 109 features per window. In addition, we also encode the reported stance of the participant on the overall topic into the feature set. As the stance was reported on a scale from 1 (completely disagree) to 5 (fully agree), the numerical rating directly serves as a feature, resulting in a feature vector $x_i = (x_i^1, \ldots, x_i^n)$ with $n = 110$ entries. As we have included features with different ranges, we re-scale each feature as

$$x^k = \frac{x^k - min(x^k)}{max(x^k) - min(x^k)}, \tag{7.16}$$

where $x^k$ denotes the list of feature values along dimension $k$. A well-known problem of this rescaling method is that extreme outliers lead to a compression of a majority of the data points into a very small range. However, in the present case, this is not an issue as we are using mostly statistical features that already put isolated extreme data points into perspective. The only exceptions are the maximum value of each action unit and the reported user stance. However, the intensities of the action units are provided by OpenFace on a fixed scale from 0 to 5 and the user stance is rated on a fixed five-point Likert scale. Therefore, both features also cannot include extreme outliers.

### 7.3.4 Experimental Setup and Results

We now discuss our approach to estimating the participant opinion based on the extracted features. We see that we have data sets of 610 (*convincing*) and 653 (*interesting*) sessions as well as 110 features per session. Based on these numbers, we expect the following challenges and requirements regarding the underlying machine learning problem:

- In comparison to other machine learning problems and corpora, the available data sets are small and we therefore only consider data-efficient methods, namely support vector machine (SVM) and random forest.
- The ratio of sessions and features indicates a high probability for overfitting during the training.
- The class distribution for the *interesting* task requires additional pre-processing to avoid majority vote classification.

We use two different metrics to measure the performance of our model, namely the percentage of correct classifications, i.e., the accuracy (ACC) and the unweighted average recall (UAR). Especially in the *interesting* case, where the classes are unequally distributed, the UAR will be the most informative metric for the overall performance as discussed in detail in Section 2.3.1.

#### Class Balance

In the case of the *interesting* task, the low number of negative samples is likely to result in a majority class model, i.e., a model that always chooses the majority class regardless of the corresponding feature. In order to prevent that, we utilize two different approaches:

- Class weights adapted to the actual class balance.
- Synthetic Minority Oversampling Technique (Chawla et al., 2002).

The first one is used only in the SVM, whereas the second one is used for both SVM and random forest classification. Both methods are introduced in detail in the following. As discussed in Section 2.3.1, SVMs separate data points by means of the hyperplane that maximizes the margin between the plane and the data points. The soft-margin parameter $\Lambda$ regulates the cost of wrong classifications and represents the trade-off between a large margin and the number of wrongly classified training samples. In the case of balanced data, this value is symmetrical for both classes, meaning that the margin has the same size on both sides of the hyperplane. For imbalanced data, the $\Lambda$ value can be adapted to the class balance in order to give the underrepresented samples additional weight (Yang et al., 2007). We herein use the balanced value

$$\Lambda_{pos,neg} = \frac{\#samples}{2 \times \#samples(pos/neg)} \Lambda, \tag{7.17}$$

where $\#samples$ denotes the number of all samples in the training set, $\#samples(pos)$ the number of samples with a positive rating and $\#samples(neg)$ the number of samples with a negative rating. Therefore, the cost for a wrong classification is increased for samples in the minority class.

The second approach to address the class imbalance is Synthetic Minority Oversampling Technique (SMOTE). The simplest way of oversampling only replicates data points in the minority class and thereby increases its overall numbers of samples. This approach is extended in SMOTE by generating synthetic data points which are close to the original ones in the feature space. In doing so, additional and new data points of the minority class are generated and the overall training set can be balanced. As this approach is applied to the data set, it is independent of the utilized classifier.

#### Results

To avoid overfitting as much as possible, we utilize repeated k-fold cross-validation to evaluate the different machine learning models. More precisely, we averaged the results of five ten-fold cross-validations to compute the final validation score for each investigated configuration. The model parameters of both SVM and random forest are selected with the complete feature set in a systematic grid search for each task separately. Afterwards, we divide the features into three groups, namely the complete feature set (full, 110 features), a feature set without action units (eye, 25 features) and a feature set without eye movement (AUs, 87 features) and investigate the model performance for each feature and task set separately. The results for both models, the *convincing* task and all three feature sets is shown in Table 7.5.

We see from the results that the SVM performs slightly better than the random forest classifier and that the performance of the SVM is actually increased on the limited feature set that includes no information about the AUs. In addition, we see that both metrics are very similar for all the cases, which can be attributed to the almost equal class balance in the data.

For the *interesting* task, we compare both machine learning techniques for all three data sets as well as the approaches discussed in Section 7.3.4 to deal with the unequal class balance. The results can be seen in Table 7.6.

In terms of UAR, the SVM with the adjusted class weight performs best. Since the UAR accounts for both classes equally and independent of the class balance, we consider it the most relevant metric for this task and it is therefore used as the main indicator of the performance. In the next step, we compare the results of the repeated ten-fold cross-validation to a leave-one-participant-out cross-validation in order to investigate if the cues shown by the participants are individually different. The results for both SVM and random forest in the case of the *convincing* task for all three feature sets are shown in Table 7.7. We see a clear decrease in both metrics for the

Table 7.5: Results of SVM and random forest (RF) on the *convincing* task for three different feature sets.

| Model | Metric | Full | Eye | AU |
|---|---|---|---|---|
| SVM | ACC | 0.61 | 0.63 | 0.57 |
| | UAR | 0.61 | 0.63 | 0.57 |
| RF | ACC | 0.61 | 0.61 | 0.57 |
| | UAR | 0.60 | 0.61 | 0.56 |

Table 7.6: Results of SVM and random forest (RF) on the *interesting* task for three different feature sets. The imbalanced data is addressed by SMOTE oversampling (SMOTE) and an adjusted class weight (cw) in the SVM.

| Model | Metric | Full | Eye | AU |
|---|---|---|---|---|
| SVM + | ACC | 0.66 | 0.65 | 0.68 |
| cw | UAR | 0.64 | 0.64 | 0.61 |
| SVM + | ACC | 0.67 | 0.62 | 0.67 |
| SMOTE | UAR | 0.61 | 0.63 | 0.59 |
| RF + | ACC | 0.66 | 0.64 | 0.65 |
| SMOTE | UAR | 0.63 | 0.61 | 0.61 |

facial action unit feature set. Moreover, the SVM again outperforms the random forest classifier and the best results are similar to the ones achieved in the repeated ten-fold cross-validation.

The results for the *interesting* task, all three classifier configurations and all three feature sets are shown in Table 7.8. Again, we see a clear drop in both metrics for the facial action unit feature set but in contrast to the *convincing* task, also for the full feature set. In this case, the performance of the different classifier configurations varies and there is no clear advantage for one approach over all three feature sets. However, the overall best results (in terms of UAR) are again achieved with the class-weighted SVM.

We conclude that the classification is in some instances hindered or at least not supported by a large set of features, which can be attributed to the limited amount of available data. In addition, the eye movement data appears to be more informative than the facial action units, especially in the *convincing* task. For a comparison with the performance of human annotators in the following subsection, we use the SVM with the feature sets eye (*convincing)* and full (*interesting*) as they perform best in the most general ten-fold cross-validation setup.

Table 7.7: Results of SVM and random forest (RF) on the *convincing* task for three different feature sets and leave-one-participant-out cross-validation.

| Model | Metric | Full | Eye | AU |
|---|---|---|---|---|
| SVM | ACC | 0.60 | 0.62 | 0.51 |
| | UAR | 0.60 | 0.62 | 0.51 |
| RF | ACC | 0.61 | 0.58 | 0.52 |
| | UAR | 0.60 | 0.58 | 0.52 |

Table 7.8: Results of SVM and random forest (RF) on the *interesting* task for three different feature sets and leave-one-participant-out cross-validation. Imbalanced data is addressed by SMOTE oversampling (SMOTE) and an adjusted class weight (cw) in the SVM.

| Model | Metric | Full | Eye | AU |
|---|---|---|---|---|
| SVM + | ACC | 0.55 | 0.63 | 0.52 |
| cw | UAR | 0.52 | 0.61 | 0.46 |
| SVM + | ACC | 0.57 | 0.59 | 0.55 |
| SMOTE | UAR | 0.49 | 0.58 | 0.45 |
| RF + | ACC | 0.62 | 0.60 | 0.59 |
| SMOTE | UAR | 0.56 | 0.53 | 0.53 |

### 7.3.5 Human Annotation

To compare the machine learning approaches to human performance, we conducted an annotation on a subset of the recordings (10%). Due to the high variations observed during the individual ten-fold cross-validations, this test set was selected in compliance with the following conditions from a list of randomly generated candidate sets to ensure a representative comparison:

- The class balance for both tasks in the test set is representative of the class balance of the overall dataset.
- The deviation of the machine learning performance on the test set from the average performance is lower than or equal to the average standard deviation of the ten-fold cross-validation.

Annotators watched snippets of the experiment recordings and were asked to rate if the observed person is interested in and convinced by the presented arguments. In addition, annotators were asked to report their confidence in the rating on a scale from 1 (low confidence) to 5 (high confidence). Annotators had the same information available that is used for the classification, namely the video snippet in the period from the beginning of the system utterance up to the first user rating and the reported stance of the respective participant on the discussed topic. In addition,

Table 7.9: Performance of the human annotators (A1-A3) and the SVM model on the test set for the *convincing* category.

| Metric | A1 | A2 | A3 | Majority | SVM |
|---|---|---|---|---|---|
| ACC | 0.64 | 0.67 | 0.64 | 0.67 | 0.64 |
| UAR | 0.63 | 0.67 | 0.64 | 0.67 | 0.64 |

Table 7.10: Performance of the human annotators (A1-A3) and the SVM model on the test set for the *interesting* task.

| Metric | A1 | A2 | A3 | Majority | SVM |
|---|---|---|---|---|---|
| ACC | 0.78 | 0.64 | 0.60 | 0.72 | 0.67 |
| UAR | 0.55 | 0.44 | 0.53 | 0.52 | 0.60 |

annotators answered two questionnaires, one before starting the annotation and one after completion. In the first questionnaire, annotators were asked to name the indicators they assume to be the most influential and to rate the sentence *The task will be difficult* on a five-point Likert scale. After completing the task, annotators were interviewed again regarding the most influential indicators for both tasks and their opinion on the difficulties they encountered. Each annotator also answered the question *The task was difficult* again on a five-point Likert scale.

The annotation was done by three annotators (two male, one female) at NAIST for each session in the test set. We measure the inter-annotator agreement for both tasks by means of Fleiss' kappa (Fleiss, 1971). For the *convincing* category, we report an agreement between the human annotators of $\kappa = 0.24$ with a corresponding p-value of $p = 0.001$ which is a fair agreement. The accuracy and UAR values for all three human annotators, the majority rating and the SVM model are shown in Table 7.9.

In the case of the *interesting* category, the inter-annotator agreement between the human annotators is $\kappa = 0.01$ with a p-value of $p = 0.864$, which means that the agreement is not more than random. The results for each annotator, the majority rating and the SVM model for this task are shown in Table 7.10.

Overall we observe a similar performance of the SVM and the human annotation in both tasks. In order to investigate the differences between the two classifications, we additionally compare the class-wise recall of the human majority rating and the SVM classification for both tasks. The corresponding results are shown in Table 7.11. For the *convincing* task, we see that the SVM has similar performance for both classes, whereas the human ratings are better for negative samples. In the case of the *interesting* task, both approaches yield better results for the positive class. In comparison, the SVM outperforms the human annotation in the negative class whereas the majority rating of the human annotators is better than the SVM for the positive class.

As for the questionnaires, all three annotators rated the expected difficulty of the task (before

Table 7.11: Class-wise recall of the human majority rating and the SVM classification for both tasks (convincing and interesting).

| Category | Class | SVM | Majority |
|---|---|---|---|
| Convincing | positive | 0.66 | 0.59 |
| | negative | 0.63 | 0.75 |
| Interesting | positive | 0.73 | 0.89 |
| | negative | 0.46 | 0.15 |

Table 7.12: Average reported confidence of the human annotators (A1-A3) for both categories (convincing and interesting).

| Category | A1 | A2 | A3 |
|---|---|---|---|
| Convincing | 3.23 | 2.56 | 3.74 |
| Interesting | 3.67 | 2.78 | 3.33 |

the experiment) with 3 and the actual difficulty (after the experiment) with 4. Indicating cues that were mentioned by at least two annotators were eye movement and facial expressions for the *interesting* task as well as facial expressions and head movement for the *convincing* task. For the *interesting* task, these cues are in line with the features utilized in the SVM, whereas the cues for the *convincing* task differ from the eye movement features used in the machine learning approach.

In addition, two of the annotators reported a lack of reactions in some instances and all three annotators said that they would assume more expressions in a human discussion with a more active role of the user. Moreover, two of the three annotators reported that it was helpful to see more than one clip of a person. The average confidence scores for both tasks and all three annotators are shown in table 7.12. It can be seen that despite the lower agreement in the *interesting* case, the average confidence is higher than in the *convincing* case for the annotators A1 and A2.

### 7.3.6 Discussion

We now provide a discussion of our findings and possible implications. As it extends over several aspects that influence the final conclusion, a summary is only provided at the end of this subsection. In general, we can see that all results are clearly above the random guess baseline of 50% and hence, that the tasks can be addressed by machine learning approaches. However, the low amount of (imbalanced) data and the observed high variability in the classification results makes further general assertions difficult. We start the detailed discussion with the comparison of human to machine learning performance, followed by a comparison of our results to literature values. Finally, we look at the results from the perspective of applications in future argumentative dialogue systems.

### Comparison of Annotation and Classification

Overall, we observe similar results for the human annotation and the best machine learning approach (SVM) in both tasks. This indicates that the reported performance is close to the upper bound of the utilized data, which is also in line with the comments of the annotators who reported a lack of cues in several instances. However, the comparison of the class-wise recall shows that there are also differences between the human and the machine classification, i.e. that some instances are classified correctly by one approach and not by the other. This can in parts be attributed to the technical configuration of the SVM, especially in the *interesting* task where the imbalanced data and the adjusted class weight of the SVM lead to a better performance in detecting negative samples at the cost of lower performance in the detection of positive ones. Also, the indicating cues reported most frequently by the annotators for this task are in line with the features used in the SVM classification, which indicates that similar information is used by the annotators and the machine learning model. For the *convincing* task, however, the observed differences in combination with the reported indicating cues suggest, that different information is used in the two classifications for this task and hence, that additional or alternative features might improve the machine learning performance further. If and to what extent such an improvement is possible has to be investigated on a larger data set in future work. As also reported by the annotators and indicated by related work, a more active role of the user is likely to improve the results as well since it allows the use of additional information (like gestures and linguistic features) and leads to a more expressive behaviour of the participants in general. In addition, the results also indicate that a user can hide his or her opinion from the system recognition.

Regarding the two different tasks, the results of the human annotation show that the *interesting* category appears to be more challenging. One reason for the difficulty of this task can be attributed to the imbalanced class distribution and the lack of clearly negative samples. Moreover, the experimental setup is also likely to influence the results, as one annotator (A3) reported the impression that participants try to be generally interested as a consequence of the recording (and not necessarily the arguments). The presence of cues that are related to general (dis)interest and not to the arguments is also an explanation for the higher average confidence scores of A1 and A2 in this task. This is an important factor for applications as well since the reason for the shown response cannot always be identified. From this perspective, it would also be beneficial to have a more active role of the user, as well as confirmation strategies on the system side to clarify the situation.

### Comparison with Literature

As discussed in Chapter 3, similar approaches in other domains/applications were investigated and we compare our results to the setups that are most similar to ours. For the *convincing* task, we compare our results to the assessment of spontaneous (dis)agreement by means of non-verbal cues (Bousmalis et al., 2011). Similar to our case, the authors investigate spontaneous (dis)agreement recognition and address it as a binary classification problem. In contrast to the herein presented approach, the data set is derived from human-human debates in which the observed individuals assume an active role and the data label is based on human annotation. In addition, prosodic features i.e. features extracted from the voice of the speakers and visual features related to hand actions as well as head and body gestures are utilized for the estimation. The reported overall accuracy of 64% is close to the herein reported accuracy of 63%, although it was accomplished with both prosodic and visual features. The reported accuracy without prosodic features is only slightly above 50% (exact values are only provided for the best results) and hence clearly lower than the herein achieved 63%.

In the case of interest recognition, we compare our results to the ones presented in (Tomimasu and Araki, 2016). The similarity to our task is mainly due to the very similar setup of human-machine interaction as well as a turn-wise assessment of the user interest on a binary scale. The differences to the present scenario are the investigated domain, the annotation-based labelling of the data and the use of both prosodic and visual features. The best reported average recall score of 70% is higher than the herein reported 64%. Nevertheless, the authors also report an average recall of 62% achieved without prosodic features which indicates that the herein achieved results are comparable to the ones in the referenced scenario.

We conclude that the comparison with the literature yields similar or better results than achieved in the reference work without prosodic features. However, both comparisons also indicate that the herein reported performance could be improved by using prosodic features in a scenario with a more active user where those are accessible. In addition, a more active role of the user would also allow for an exploration of additional visual features such as gestures and postures for the herein considered tasks.

### Applications and Limitations

For applications, it is evident that the herein discussed approaches need to be enhanced to perform reliably. In particular, additional training data is required to ensure a robust recognition in application scenarios. Also, users are able to hide their opinion which can hinder the recognition in competitive tasks like negotiation even further. Consequently, scenarios in which the user cooperates with the system are suitable candidates for first applications. For example, a system like the one discussed at the beginning of the chapter for the information-seeking domain could use the herein proposed methods in combination with confirmation strategies in direct interaction with human users. In addition, the more active role of the user in such a scenario is likely to facilitate the recognition of both herein discussed quality aspects.

A remaining open question is whether or not a personalized model of the user opinion established over multiple interactions can further improve the recognition. A comparison of the best results for the ten-fold cross-validation and the leave-one-participant-out cross-validation shows only a very small difference. However, these results correspond to a classification solely based on eye movement features, whereas especially for the facial action units, the results clearly decrease in the leave-one-participant-out cross-validation. A possible explanation is that the facial expressions are more individual and the eye movements are more general. This is also in line with the somewhat controversial reports of the annotators: One annotator said that additional clips of a single participant did not help in accomplishing the task as no context information is provided thereby. In contrast, the remaining two annotators reported that additional clips of the participants helped.

### Summary

We discussed the estimation of subjective argument quality criteria from social signals, namely how *convincing* and how *interesting* an argument was perceived by a human user. In order to do so, we utilized data collected for the evaluation of argument search engines with a dialogue system that allows users to rate single arguments in multiple categories. The corresponding ratings were estimated from the facial expressions and the eye movement shown by the participants during the interaction with the system.

A comparison with human annotations yielded similar results, with a slight advantage of the human annotation in the *convincing* task and a slight advantage of the machine learning approach in the case of the *interesting* task (in terms of UAR). Moreover, we reported a fair agreement between human annotators for the *convincing* case, whereas the *interesting* case showed only a random agreement. Finally, we compared our results to the ones achieved in the literature for scenarios that bear similarities to the herein addressed tasks. The comparison showed that our results are similar (*interesting*) or above (*convincing*) the results reported in the compared literature for similar feature sets (visual features). In summary, we draw the following conclusions:

- The overall results are clearly above the random baseline and comparable to performances reported in the literature for similar tasks and feature sets. Hence, a recognition of both subjective argument quality aspects from social signals is feasible.
- The results of the human annotation show that both tasks are nevertheless challenging and additional data in combination with a more active role of the user is likely to improve the performance.
- The machine learning results are in the same range as the performance of the human annotators and it is therefore unlikely that the results can be significantly improved by additional features in the same setup.

Since one current limitation is the amount of training data, the next steps in this research direction should include a more extensive data collection to improve the recognition performance and allow for an application in actual systems.

## 7.4 Conclusion

Throughout this chapter, user-adaptive argumentation with conversational agents was addressed. As stated in the beginning, this task can be divided into the development of adaptation approaches, also including user models, and the (implicit) assessment of user feedback. The first aspect was approached through the introduction of a preference model based on formal concepts of weighted bipolar argument graphs and the utilized argument structure. It is updated incrementally during the interaction based on the user perception of individual argument components in the structure.

Several application scenarios were discussed, including the use of the proposed model to adapt the multimodal behaviour of an argumentative agent to individual user feedback through reinforcement learning. For this adaptation approach, an evaluation system was introduced that presents argument components to human users and enables a rating in the categories *convincing*, *neutral*, *not convincing*. The system was used to evaluate the general feasibility of the reinforcement learning-based adaptation and to assess the predictive power of the preference model in a user study. The results indicate the functionality of the model and its suitability for application systems.

Based on the observations during the first two steps, the need for an implicit assessment of the individual user opinion regarding the argument under discussion was identified and addressed in the last section of the chapter. In contrast to related work discussed in Section 3.2, the task is approached from the affective computing side through an automatic estimation of the aspects *interesting* and *convincing* from non-verbal cues shown by the user during interaction. The proposed supervised learning model was trained on data collected during the evaluation of argument search approaches in Chapter 6 and compared to human annotations. The results indicate that the task is difficult but also that the machine learning performance is similar to the human performance.

In terms of flexibility, it can be seen that the discussed preference model and the approach to automatically assess user feedback both do not depend on a specific application scenario and are hence applicable in multiple different systems with varying tasks. Nevertheless, the preference model depends on the bipolar relations between argument components in the argument structure although this dependency does not include the assumption of individual targets. It is hence applicable to any argument representation that does not distinguish between multiple component types (beyond claim and premise/evidence) and includes bipolar (or only attack) relations. As already discussed in Section 6.4, this is not a strong limitation with respect to argumentative dialogue systems as the majority of the systems discussed in Chapter 3 rely on this (or a similar) kind of argument structure. It is hence concluded that the introduced preference model is flexible with respect to different application systems. This flexibility is also indicated by the different settings to which the model was applied in the present chapter.

As for the automatic estimation of user feedback, there is no dependency on a specific argument representation and it is hence even more flexible on a conceptual level. The main limitation of the proposed approach in the current state is the performance which is, although similar to human performance and literature values, not reliable enough for direct applications without complementary feedback. However, as the general feasibility of the approach was shown, it is likely that this shortcoming can be overcome with a larger set of training data and through the consideration of additional (prosodic) features that can be used in scenarios with a more active user.

Finally, the discussed reinforcement learning-based adaptation was introduced for a specific setup and adjusted, in the present case only the multimodal behaviour of the agent. It is hence limited in the present form to a specific application. An advanced approach that combines the discussed adaptation of the multimodal behaviour with the selection of arguments is discussed in the context of application systems in the following chapter.

# 8 Demonstration of Applicability

This chapter reports on application systems that were implemented using the previously discussed concepts. Although built for different scenarios, all herein discussed versions of the system were developed in the scope of the DFG Ratio project *EVA: Empowering Virtual Agents to improve their persuasiveness* and in cooperation with Augsburg University.

The discussion addresses each system version separately and also distinguish the different application scenarios for which the corresponding version was designed. These scenarios are separated into application-oriented and evaluation-oriented setups. The application-oriented case is focused on implementing a persuasive system that is capable of interacting with a human user and hence on integrating the user into the persuasive discussion as an interlocutor. Therefore, the corresponding version of EVA utilizes a single virtual agent and the user assumes the role of its counterpart in the discussion. In contrast, the evaluation-oriented case focuses on the human perception of agent-agent argumentation and users hence assume the passive role of a judging audience. The main motivation behind this scenario is to isolate and investigate individual aspects of persuasion and their effect on the overall perceived persuasive effectiveness of a virtual agent. Consequently, the persuasive dialogue includes two virtual agents that can be observed and rated by a human audience. For both scenarios, multimodal argumentation is considered and the system is in each instance represented by virtual avatars that interact via synthetic speech, mimic and gestures with their interlocutor.

## 8.1 Single-Agent EVA

In (Rach et al., 2018c), the single-agent version of the multimodal argumentative dialogue system EVA is introduced. It engages with a user in a discussion about a controversial topic. The system is capable of presenting arguments via natural language and supporting them with multimodal emotions. To this end, the setup discussed in Chapter 4 is extended by a virtual avatar that conveys the system's output and a graphical interface enabling the user to choose their input. During the interaction, users assume the role of the second agent and discuss a controversial topic with their artificial counterpart through the selection of appropriate utterances provided through an interface. In addition, the single-agent version of the system utilizes the markov game formalism for dialogue games introduced in Section 5.1 and is capable of discussing any topic encoded in an argument structure derived with one of the methods introduced in Chapter 6.

### 8.1.1 Architecture

The architecture of the system is separated into two modules: the *argumentative dialogue system* which regulates the interaction, and the *virtual avatar* responsible for producing multimodal out-


Figure 8.1: Screen capture of the EVA interface including a dropdown menu with possible answers, avatar and dialogue history.

put. The interface for the user input is realized as a dropdown menu in which allowed responses in the current state of the dialogue are presented. The system's reply is chosen by the dialogue system and presented to the user through the virtual avatar by different modalities including synthetic speech, mimic and gestures.

In the following, both modules of the system are discussed in more detail. A screen capture of the interface including menu, dialogue[1] and avatar is shown in Figure 8.1.

## Argumentative Dialogue System

As previously mentioned, the experimental setup discussed in Chapter 4 is applied in the dialogue system module. Consequently, the interaction between the user and the system is modelled as a dialogue game for argumentation as introduced in (Prakken, 2000; Prakken, 2005) and discussed

[1] Material reproduced from www.idebate.org with the permission of the International Debating Education Association. Copyright ©2005 International Debate Education Association. All Rights Reserved

in Section 2.2.1. Thus, each utterance corresponds to a formal move in this game and the set of allowed moves in each state of the dialogue is given by the game protocol.

In addition, the system relies on the previously discussed tree structure encoding the available arguments as well as their relations in the form of an OWL ontology. For demonstration purposes, the database encodes 72 argument components about the topic *Marriage is an outdated institution* extracted from the corresponding debate from the *Debatabase* of the idebate.org website (see Chapter 4) and 64 argument components concerning the topic *Boxing should be banned* from the argument mining corpus presented in (Aharoni et al., 2014). However, the argument structures discussed in Chapter 6 can be directly applied in this system. In particular, the automatic approach to generate tree structures in Section 6.2 can be utilized to generate a database for arbitrary topics on which the search engine finds suitable arguments.

As for the system strategy, two options are available: The first (rule-based) policy relies on the probabilistic rules discussed in Chapter 4. The second (reinforcement learning) policy is optimized in the markov game framework as discussed in Section 5.1.2. Whereas the rule-based policy can be directly applied to new topics, the reinforcement learning strategy requires training prior to the interaction for the respective argument structure. For the two above mentioned topics, pre-trained policies achieved with Q($\lambda$) as well as SARSA($\lambda$) and linear function approximation are available.

Finally, the natural language generation (NLG) is also adapted from Chapter 4. Consequently, the sentences corresponding to argument components in the utilized database serve as the formulations for *argue* moves and the formulations for moderating utterances are chosen randomly from a template list.

### Web-Based Avatar

To provide our system with multimodal output capabilities, we employ the Charamel™ avatar[2] being freely available for research institutes and students. It makes use of the Nuance TTS and all Amazon Polly voices[3] to enable the user to listen to their virtual interlocutor. The system utterances are presented by the avatar in a natural way using synthetic speech and are additionally displayed as text in the context of the complete dialogue history. This enables the user to both read the utterances while listening and read previous utterances. For the multimodal behaviour, the avatar utilizes a pre-defined template of mimics and gestures. It is generated from the available expressions (mimic and gestures) of the avatar and synchronized with specific formulations of the NLG.

## 8.1.2 Towards Natural Language Understanding of User Input

Despite its flexibility with respect to the strategy and possible topics, the system is still limited regarding the interface that requires users to select their responses from a list of pre-defined options. Although not implemented, yet, we propose an approach to address this issue in (Rach et al., 2018a) for the system under discussion. The main challenge in this regard is that users are not

[2]https://www.charamel.com/competence/avatare (last accessed 29 August 2021)
[3]https://docs.aws.amazon.com/polly/latest/dg/voicelist.html (last accessed 29 August 2021)

limited to arguments that can be constructed from the argument structure nor are they forced to strictly follow the rules of the dialogue game. Concepts to address both challenges are discussed in detail in the following.

The argument structure of the system encodes all the knowledge the system has about the discussed topic. Thus, natural language input from users has to be mapped to known components or integrated (in case of unfamiliar content) into this database to enable an appropriate response of the system. Consequently, the utterance of the user has to be (automatically) processed into a form that fits this representation. In this process, different questions have to be addressed:

- Is the user's utterance allowed/reasonable in the current state of the game? (Q1)
- If so, is the respective content included in the argument structure? (Q2)
- If it is not, is the new content valid in the current state of the game? (Q3)

The first question refers to the identification of the type of move and is usually called *intent recognition*. This sub-task can be addressed by state-of-the-art NLU tools as for example the ones compared in (Braun et al., 2017) or through a language model trained (or fine-tuned) on the task. If the user input includes an invalid type of move (for example a nonsense statement), the system can reject the utterance and ask for an alternative formulation.

The remaining two questions have to be addressed separately for *argue* moves and moderating utterances. For a moderating utterance (*why*, *concede*, *retract*), the system move that the utterance refers to has to be identified (Q2). This can be addressed by employing semantic similarity measures like the one introduced in (Li et al., 2006) and word or sentence embeddings. If no explicit target (similarity over a certain threshold) is found in the dialogue, the system can assume a reference to its latest utterance and determine the validity of the user move accordingly (Q3).

For *argue* moves, it has to be determined if the user utterance corresponds to an argument that can be constructed from the argument structure (Q2). Approaches that can be utilized for this step are again semantic similarity measures as well as techniques to estimate argument similarity, as recently introduced in the context of argument clustering (Reimers et al., 2019). If the utterance includes a known argument, the legality of the user move can be directly determined through the game protocol (Q3). If it is unknown, the system has to determine the validity of the new content in the context of the dialogue game. As the corresponding utterance was already identified as *argue* move (and not as an unrelated statement), the remaining question is how the corresponding argument relates to the ones known by the system. As discussed in Chapter 4, the arguments considered by the system are fully defined through their premise, i.e. the respective component in the argument structure. Consequently, the task at hand is to identify possible relations between the user utterance and the respective components in the utilized structure. This relation estimation is a major task in the field of argument mining (see Section 2.2.4) and at the core of the automatic generation of argument structures discussed in Section 6.2. Consequently, the therein trained relation classification can be applied to identify suitable candidates towards which the freshly introduced argument holds a relation. A successful estimation enables the system to assess the new argument in the context of the dialogue and to respond accordingly (Q3). However, since this mapping has to be done in real-time, i.e. during the ongoing interaction with the user, a computationally efficient approach to this classification is required.

It is worth noting that this approach also enables the system to learn the new content by including it in the database. Consequently, the new argument can be utilized by the system itself in later dialogues. Similar to humans, the system is thus able to remember arguments even if they were not included in the database in the first place.

The procedure is illustrated for a new user argument in an example with the topic *Marriage is an outdated institution* from the annotated sample debate in Chapter 4. In the following dialogue, the user responds to the systems initial claim (S1) with an unfamiliar argument (U1).

S1: *Marriage is an outdated institution.*

U1: *Marriage promotes a better way to raise children.*

The NLU first has to detect the type of move (Q1) – in this case *argue*. As an *argue* move is a valid response to the initial claim, the system now matches the utterance with known arguments (Q2) by for example computing the semantic similarity between the utterance and argument component in the argument structures. As a respective component is not included in the database (i.e. each similarity score is below a certain threshold), the NLU now investigates whether or not the content is related to the content of the initial claim by identifying possible relations to earlier *argue* or *claim* moves (Q3). As the user utterance directly addresses the main topic, it is a legal response and the system accepts the new argument as well as the respective move and selects an appropriate answer. In the herein considered scenario, possible responses are to challenge the user move (S2.1) or concede to it (S2.2).

S2.1: *Why do you think that?*

S2.2: *I agree with you.*

In a later stage of the dialogue, the system can also change the topic by introducing a counter-argument on an earlier (familiar) one of the interlocutor. In all cases, the system extends its own database by including the new argument which is then available in future discussions.

### 8.1.3 Discussion

The introduced multi-modal system is capable of handling complex persuasive dialogues with a human counterpart based on the underlying dialogue game formalization. In addition, the combination with the markov game framework enables the optimization of the strategy for any given topic using multi-agent reinforcement learning. Due to its reliance on tree structures, a direct application of the retrieval pipeline discussed in Chapter 6 is possible and the system can hence discuss any topic on which the utilized argument search engine can find suitable arguments. The main limitation of the system is the reliance on a dropdown menu for the user response. Although it guarantees that only arguments encoded in the argument structure are used and that the user sticks to the utilized game formalism, it is not a natural or intuitive way to discuss for humans. This shortcoming can be addressed through a flexible NLU module that utilizes state-of-the-art classification approaches from the field of natural language processing and argument mining. Despite this current limitation of the system, it enables users to explore a multitude of topics through a dialogue interaction and to explore different aspects of a topic collected from a variety of different sources in a convenient, incremental way.

## 8.2 Multi-Agent EVA

Next, the first multi-agent version of the system for the evaluation-oriented scenario is discussed. The question of how persuasion works was addressed in many different fields, including philosophy (Krapinger, 1999), psychology (Petty and Cacioppo, 1986) and computational argumentation (Wachsmuth et al., 2017b). Despite the differences in all these approaches, it has become clear that the process of persuading a person includes an interplay of multiple different aspects. Especially in the case of dialogical persuasion, it involves not just the rational arrangement of suitable arguments, but also a presentation that is emotionally appealing to the interlocutor. However, isolating and investigating the contribution of these individual aspects is difficult in human dialogues, as humans usually act and react intuitively in a conversation. To address this issue, we introduce the multi-agent version of EVA (Weber et al., 2020b) in which different aspects of persuasion can be displayed, isolated and compared. To ensure a natural and intuitive interaction, each agent is represented by a virtual avatar that interacts with its counterpart through synthetic voice and multimodal emotions.

Within this system version, two aspects of persuasion are considered: *Logical/rational persuasion*, i.e. the persuasion through an appropriate selection of arguments and *subliminal/emotional persuasion*, i.e. persuasion through an utterance presentation that is emotionally appealing to the audience. In this first version of the system, both aspects are treated as separate problems, meaning that the selection of utterances is independent of the selection of the emotion the system is supposed to convey. Whereas the rational strategy is optimized prior to the interaction through multi-agent reinforcement learning, the emotional strategy is adapted during the interaction with respect to user feedback based on the approach in Section 7.2. The formalization and development of the emotional policy were again done by co-workers at Augsburg University. They are included herein for the sake of completeness.

As for the discussed topics, the annotated argument structure on the topic *Marriage is an outdated institution* is used to generate examples but the system is capable of handling argument structures derived with the methods introduced in Chapter 6 without any modification.

### 8.2.1 Architecture

The proposed multi-agent system is comprised of three modules: A *dialogue module* which manages the interaction between the two agents, an *emotion module* which adapts the emotion conveyed by the avatars to user feedback and the *interface* that includes the avatars and buttons for providing feedback. As in the previous section, the dialogue module is based on the setup discussed in Chapter 4. Consequently, the dialogue is again modelled as a dialogue game for argumentation according to (Prakken, 2000; Prakken, 2005) and the available arguments are encoded in a tree structure. Besides the rule-based strategy, the system again includes the reinforcement learning-based strategy. To distinguish between the strategy in the dialogue game and the strategy that selects the emotion the system is supposed to convey, the dialogue strategy (or rational strategy) for agent $p$ is denoted with $\pi^p_{ratio}$.

The emotion module manages the emotional presentation of the system utterances according to an emotional policy $\pi^p_{emo}$ (for agent $p$). It is derived using the reinforcement learning-based adaptation approach discussed in Section 7.2. As the original version of this adaptation algorithm

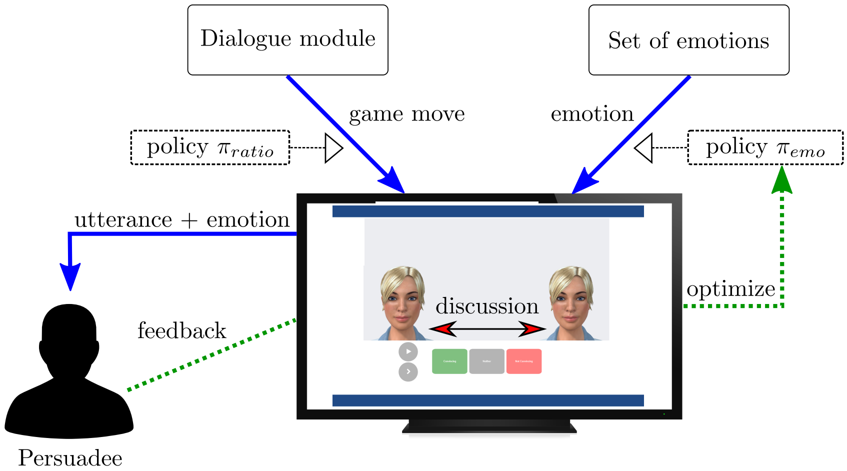


Figure 8.2: Overview of our approach with two virtual agents discussing a controversial topic. Both sides employ a pre-trained logical strategy $\pi_{ratio}$ as well a dynamically learned behavior strategy $\pi_{emo}$ adapted to the user/persuadee based on explicit feedback (*convincing*, *neutral*, *not convincing*).

was investigated in a single-agent scenario, it is herein extended to match the multi-agent setup and the utilized dialogue game. For every interaction step $t$, the player to move in the dialogue game selects an utterance (one after the other) as well as a corresponding emotion and presents it to the user. The user then provides the agents with feedback, which is used to optimize the policy $\pi^p_{emo}$ following the adaptation algorithm.

The interface of the system includes the natural language generation of utterances, two avatars and the buttons for user feedback. The NLG is again based on the setup in Chapter 4. Consequently, the natural language sentences associated with argument components in the argument structures are used to generate argumentative utterances, whereas the formulations for the remaining moves are (randomly) retrieved from a template list. The utterances are then presented by the avatar of the corresponding player via synthetic speech and the multimodal emotion (mimics and gestures) the agent is supposed to convey according to $\pi^p_{emo}$. Again, Charamel™ avatars utilizing Nuance TTS and Amazon Polly voices are used, as in the previous version. As the multi-agent system is designed for evaluation-oriented setups, the feedback options *convincing*, *neutral* and *not convincing* can be selected by the user as in the single-agent evaluation system discussed in Section 7.2. It should be noted, however, that the present system is not designed to evaluate one specific approach (as the one in Section 7.2). Instead, it includes several different models to display a complex discussion between virtual agents. The overall concept is sketched in Figure 8.2.

In the following subsections, we first propose a formalization of the complete decision making (rational and emotional policy) to provide insights into the interplay between the two different policies. Subsequently, the formal extensions of the real-time adaptation approach required for the application in the considered multi-agent setup are introduced.

### 8.2.2 Formalization

To apply the adaptation approach discussed in Section 7.2 to the multi-agent system, we again utilize the markov game framework. As can be seen from Definition 22, the state space in the original adaptation approach is based on features of the utilized argument components. However, in the present case argument components are introduced in the context of the dialogue game which results in a dependency of the adaptation approach on the rational policies of the agents. To account for this dependency, we propose a formalization of the complete decision making problem in a joint agent policy comprised of $\pi^p_{ratio}$ and $\pi^p_{emo}$. The corresponding markov game with two players is then formally defined as 5-tuple $(I, S, \boldsymbol{A}, \boldsymbol{r}, T)$, where $I$ defines the set of players, $S$ defines the merged state space, $\boldsymbol{A}$ the joint action space and $\boldsymbol{r}$ the joint reward function and $T : S \times \boldsymbol{A} \times S \to [0,1]$ determines the transition function. The individual components are defined as follows:

**Definition 24** (Joint State Space)**.** *Let $S_{emo}$ be the sub state space to determine the next emotion and $S_{ratio}$ the sub state space for determining the next dialogue game move, then the system's merged state space $S$ is defined as*

$$S := S_{ratio} \times S_{emo}\,. \tag{8.1}$$

**Definition 25** (Joint Action Space)**.** *Let $A^p_{emo}$ be the emotion sub-action space and $A^p_{ratio}$ the argumentation sub-action space for player $p \in I$. Further, let $A_p := A^p_{emo} \times A^p_{ratio}$ be the respective merged action space for player $p \in I$, then the* joint action space *is defined as*

$$\boldsymbol{A} := \times_{p \in I} A_p\,. \tag{8.2}$$

**Definition 26** (Joint Reward Function)**.** *Let the function $r^p_{emo} : S \times \boldsymbol{A} \to \mathbb{R}$ be the reward given for the emotional part and $r^p_{ratio} : S \times \boldsymbol{A} \to \mathbb{R}$ be the reward given for the rational part for player $p \in I$. Further, let $r_p := r^p_{emo} \times r^p_{ratio}$ be the respective merged reward function, then the joint reward function $\boldsymbol{r}$ is defined as*

$$\boldsymbol{r} := \times_{p \in I} r_p\,. \tag{8.3}$$

**Definition 27** (Joint Policy)**.** *Let $\pi^p_{emo}$ be the emotion policy and $\pi^p_{ratio}$ be the argumentation policy for player $p \in I$. Further, let $\pi_p := \pi^p_{emo} \times \pi^p_{ratio}$ be the respective merged policy, then the joint policy is defined as*

$$\boldsymbol{\pi} := \times_{p \in I} \pi_p\,. \tag{8.4}$$

Both agents in the joint formulation have again opposing goals and it is hence again a zero-sum game. The optimum *joint policy* $\boldsymbol{\pi}$ corresponds to a Nash Equilibrium, which means that a change of any strategy $\pi_p$ does not yield an advantage for the respective agent if the policy of the opposing player is kept stationary.

It can be seen from the definition and the discussion of the individual learning problems in earlier sections, that the two state and action spaces clearly differ in their size and complexity. Since the argumentation policy $\pi^p_{ratio}$ should be able to differentiate between all available arguments, a real-time optimization (during the interaction with a user) is not feasible. On the other hand, research has yielded different objective quality criteria for logical argumentation. As discussed

---

**Algorithm 4:** Adaptation

---

**foreach** $t = 1, ..., n$ **do**
  **foreach** $p \in I$ **do**
    $s_t \leftarrow (s_{ratio,t}, s_{emo,t}) \in S$
    select action $a^p_{ratio,t}$ for $s_{ratio,t}$
    select action $a^p_{emo,t}$ for $s_{emo,t}$
    apply $a^p_t = (a^p_{ratio,t}, a^p_{emo,t})$
    observe state $s_{t+1}$ and $r_p$
    update policy $\pi^p_{emo}$.
  **end**
**end**

---

in Section 5.1.2, these theoretical insights can be utilized to define an appropriate reward. Since this is not necessarily the case for the emotional policy $\pi^p_{emo}$ (due to subjectivity), it has to be adapted in real-time and directly to the user response. Consequently, we split the learning process into two phases (pre-training and real-time) and optimize the two parts of the policy $\boldsymbol{\pi}$ separately. Algorithm 4 sketches the general procedure during the interaction.

Based on this formalization of the overall decision making of the agents, we now introduce the required modifications of the adaptation approach. As discussed above, the state space in Definition 22 is based on features of the presented argument component. The necessity of the modifications consequently results from the dialogue model as it includes not only argumentative but also moderating utterances. In addition, the role of both agents is taken into account, yielding the following definition:

**Definition 28** (State Space $S_{emo}$)**.** *Let $\varphi_i \in L_t$ be the argument component in the argumentative move $m_t$ chosen at time step t by agent $p \in I$, stan : $L_t \rightarrow \{pro, con\}$ the stance function that assigns each argument component to its stance towards the root node and $rel(\varphi_i) \in \{\text{attack}, \text{support}\}$ the relation of $\varphi_i$ towards its target. Further let score : $L_t \rightarrow [-1, 1]$ be a normalized compound score of the sentiment analysis of a component. Then the state $s_{emo}$ is defined as follows considering it is player $p'$s turn:*

$$s_{emo} := (p, stan(\varphi_i), rel(\varphi_i), score(\varphi_i)). \tag{8.5}$$

It can be seen that the state directly depends on the component $\varphi_i$ and not on the dialogue move $m_t$. Consequently, the corresponding parameters in the state are only updated after an *argue* or *claim* move and constant for all remaining moves. As the emotional policy adapts in the current state only the emotion with which the utterances are presented, the corresponding action space is analogous to the single-agent case and defined as follows:

**Definition 29** (Action Space $A_{emo}$)**.** *Let Emo be the set of all emotions that can be displayed by the system, including their different intensities and $a_w$ an await action that does nothing. Then, the action space for agent p is defined as*

$$A^p_{emo} := Emo \cup \{a_w\}. \tag{8.6}$$

As in Section 5.1, the *wait* action $a_w$ is included to enable the multi-move protocol in the utilized dialogue game. The reward function $r^p_{emo}$ for agent $p$ depends on the preference model that provides an estimate of the current user stance as discussed in Chapter 7. As the agents have opposing roles assigned at the beginning of the interaction, each agent derives its reward from its corresponding effectiveness, i.e. the strength of the root node in the argument structure updated with the user feedback and inverted with respect to the agent stance. We herein assume a mapping of the user feedback into numerical values as discussed in Section 7.2 and a corresponding update of the component strengths (according to Equation 7.13), yielding the following reward definition:

**Definition 30** (Reward Function $r_{emo}$)**.** *Let $e^p_t$ be the effectiveness of player p at time t with respect to the topic $\varphi_0$, then the reward $r^p_{emo,t}$ at time step t is defined as*

$$r^p_{emo,t} := |L_t|\,(e^p_t - e^p_{t-1}), \tag{8.7}$$

*with $e^p_0 = 0.5\ \forall p \in I$.*

In the above definition, the subscript specifying the effectiveness component is omitted for the sake of readability and because only the effectiveness regarding the root component $\varphi_0$ is considered for the reward. In addition, the reward definition implies that user feedback is only given for argumentative moves, resulting in a reward of zero for the remaining moves. The use of the effectiveness in the multi-agent setup enables both agents to assess their performance with respect to the one of the opponent as the utilized prediction model includes the feedback regarding all introduced argument components. The advantage of this approach is that an agent can learn to take over the opponent's strategy if it notices that it currently holds the worse position. In the presented version of the system, both agents maintain the same prediction model although it is also possible in the future to use different instantiations for the two agents to compare their effect on the resulting strategies.

In the following, we provide a detailed example of an interaction between both agents and the user to demonstrate the functionalities of the system. The herein presented dialogue between the two agents concerns the topic ($\varphi_0$) *Marriage is an outdated institution*[4] and is shown in Figure 8.3. The stances of the two depicted agents are as follows:

- Left agent: $stan = pro$
- Right agent: $stan = con$

For the sake of simplicity, we assumed that the user is against the claim and finds the right agent (*Agent 2*) convincing if attacking arguments are presented with an *angry* emotion. The left agent (*Agent 1*) is found convincing if the sentiment of the argument is negative. In all other cases, a negative reward is given.

The first argument introduced by *Agent 1* is presented with a neutral emotion and, therefore, *not convincing* leading to a negative reward for the agent. The next argument of *Agent 2* is an attacking argument and presented along with an angry emotion and, therefore, found *convincing*.

[4] Material reproduced from www.idebate.org with the permission of the International Debating Education Association. 

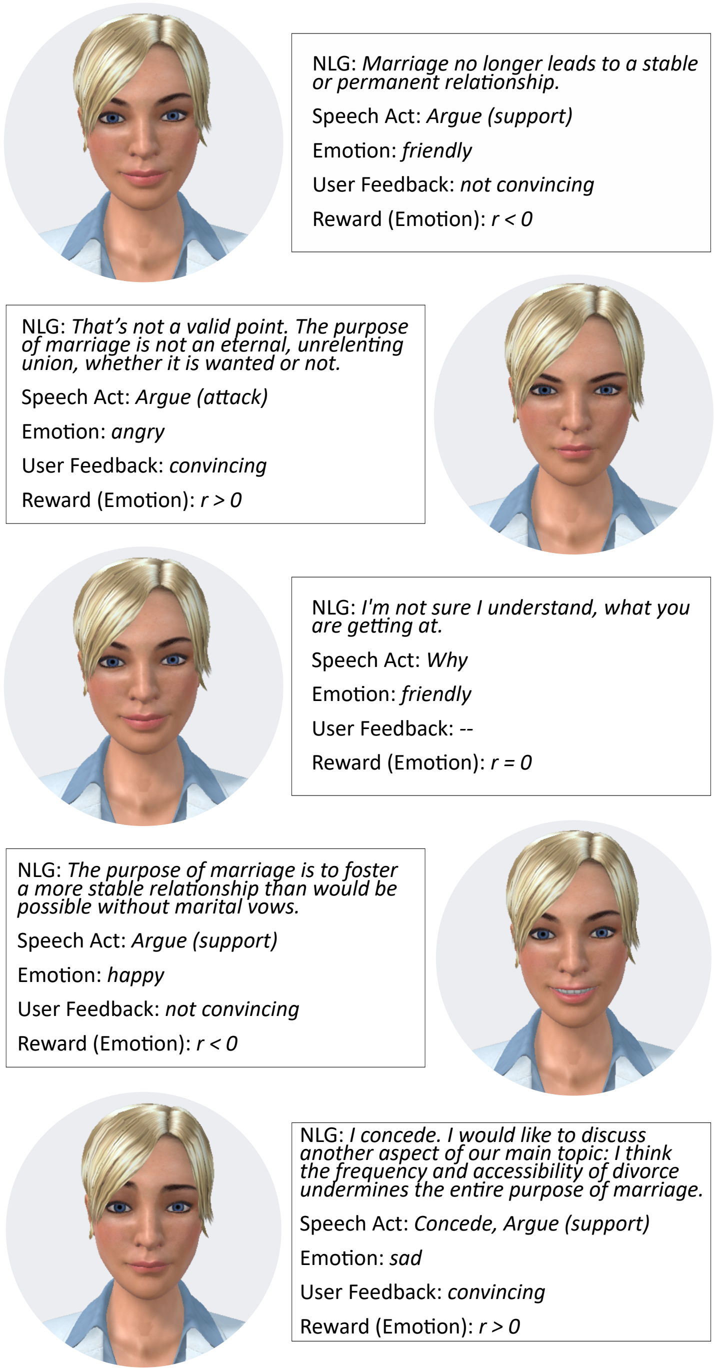


Figure 8.3: Example dialog between the two agents showing different utterances presented with different emotions and the corresponding feedback of a fictitious human user.

*Why* speech acts do not require feedback since they do not contain any argument, therefore, the reward is always zero. The fourth argument is found *not convincing* due to the nature of our fictitious user, while the last argument given by *Agent 1* is found *convincing* again since the correct behaviour is chosen.

### 8.2.3 Discussion

The presented first version of the multi-agent argumentative system combines a rational strategy developed using the methods proposed in Chapter 5 with a user-adaptive emotional policy that changes the emotion the agents convey according to an extended adaptation approach based on the methods discussed in Chapter 7. As the system utilizes the herein considered argument structure, it is capable of discussing any topic for which an argument structure can be derived with the approaches discussed in Chapter 6. Consequently, it combines methods from all three herein investigated sub-tasks of flexible argumentation in a single system. Although fully implemented, it can be seen as a prototype version in the context of this thesis as it is based on the unmodified original dialogue game and does not investigate an interplay between the individual approaches as for example user-adaptive utterance selection.

This is addressed in the following section, where several modifications are applied to the multi-agent system to utilize the full potential of the herein investigated methods.

## 8.3 Multi-Agent EVA 2.0

The updated version of the multi-agent system is published in (Rach et al., 2021c) and extends the previously discussed one in multiple ways. First, the modified dialogue game that includes chained arguments (see Section 5.2) is used to enable a more natural and intuitive persuasive interaction between the virtual agents. In addition, we include a conceptual extension to the decision making of the system: Whereas the previously discussed version separated the emotion the system is supposed to convey from the selection of the next dialogue utterance and treated them as individual problems, both aspects are now combined into a new emotional policy that is adapted in real-time with respect to the individual user response. The next utterance is then selected in compliance with the adapted emotion and based on the emotional wording of the arguments. In addition, an updated version of the original rational strategy is included which is optimized prior to the interaction in self-play with respect to the new formal framework by means of deep reinforcement learning.

The system utilizes argument structures from hotels and reviews extracted with the procedure discussed in Section 6.3. This choice is motivated by the goal to have a minimal bias of the user in evaluation studies. Moreover, the use of reviews ensures that arguments with different emotions are included, as they are based on subjective customer opinions. However, all argument structures retrieved with methods discussed in Chapter 6 can be directly applied in this version of the system as well.

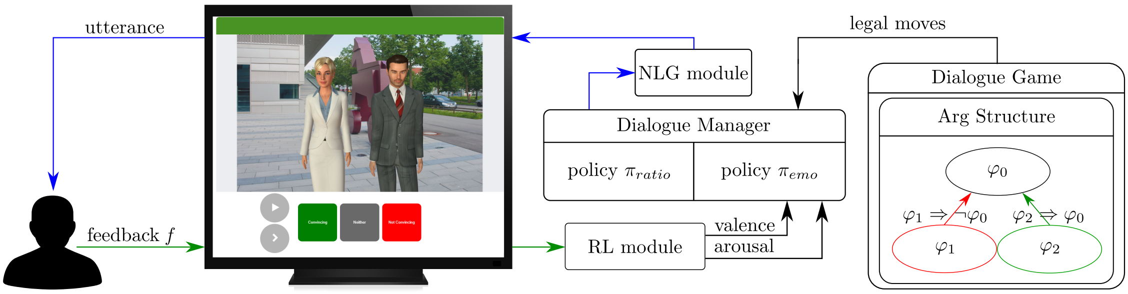


Figure 8.4: Conceptual overview of the combined system with two arguing virtual agents Alice (left) and Bob (right) allowing the user to give feedback using the three feedback buttons (*convincing*, *neither*, *not convincing*) at the bottom.

### 8.3.1 Architecture

The extended system is comprised of five different modules as can be seen in Figure 8.4. The *dialogue game module* regulates the interaction and keeps track of the dialogue history, following the modified dialogue game formalism discussed in Section 5.2. At each stage of the interaction, it provides a set of available game moves from which the *dialogue manager* selects one in compliance with the currently utilized policy. It can be seen that this module now additionally includes an emotional policy, in contrast to the previous version where the emotion the system conveys was not considered in the utterance selection. All selected moves that correspond to a turn in the dialogue game are transformed by the *NLG module* into a system utterance according to the template-based natural language generation also introduced in Section 5.2 for evaluation. The separation of this module from the dialogue manager is due to the merging of multiple moves into a single utterance and hence results from the modified dialogue game.

Each utterance is presented by the avatar of the corresponding agent with synthetic speech and emotion in the interface. As in the previous versions, the interface is based on the Charamel™ avatar and utilizes Nuance TTS in combination with Amazon Polly voices. In addition to both avatars, the interface again includes buttons that enable users to assess the presented turn if it includes an argument. This feedback is used by the *RL module* to update the emotion conveyed by the system which is then utilized by the emotional policy in the next turn. Regarding the avatars, we selected one male and one female avatar for demonstration purposes and the corresponding players are denoted with Bob and Alice for the remainder of this section. In practice, different avatars (male and female) can be chosen in compliance with the desired setup. This also allows for an investigation of gender effects on the perceived persuasive effectiveness (Siegel et al., 2009).

The selection of the next game move is again formally described as a markov game as discussed in Section 5.1 with two players and applied to optimize the two different policies. In addition, the separation of pre-training and real-time learning introduced for the previous version is still utilized in the modified scenario as well as the difference in the complexity of the corresponding state and action spaces still remains. However, as the present version utilizes the emotional policy for utterance selection, the formulation as joint policy discussed earlier is not applicable here. Instead,

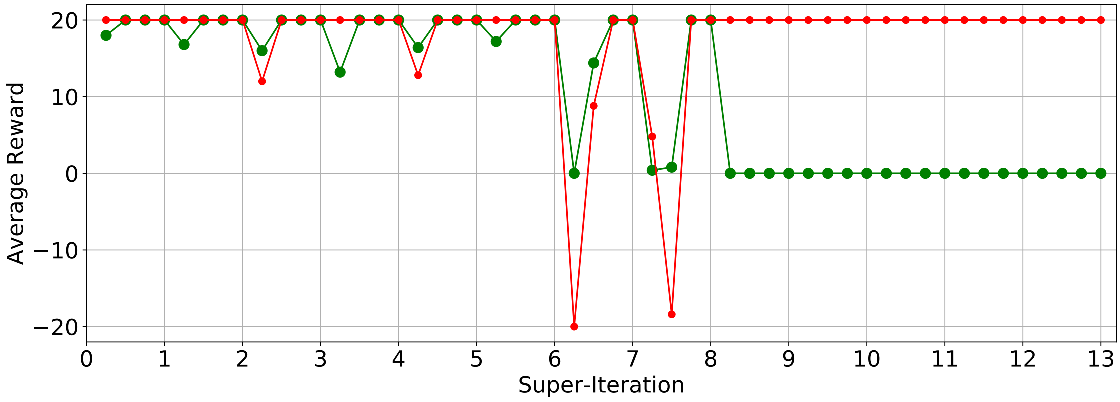


Figure 8.5: Average reward for the training agent as a function of super-iterations. The average over both stances is shown green, the average when assigned the stance of the opponent is shown red.

the two policies are alternative strategies the agents can use. The formalization and adaptations for both the purely rational policy as well as the new emotional policy will be discussed in detail in the following subsections.

### 8.3.2 Rational Policy

For the rational policy, i.e. the policy that is focused on the selection of appropriate arguments to persuade, the reformulation of dialogue games into markov games discussed in Chapter 5 is applied to the modified framework. The modifications do not influence the assumptions made for the reformulation of the original framework, namely that the temporal order of the moves is not relevant for the legality of moves and that commitments are not enforced. Although the reformulation is hence formally analogue to the one in Section 5.1, the addition of a new speech act in the modified framework increases the state and the action space and thus requires an advanced learning algorithm. Consequently, the Actor-Critic with Experience Replay (ACER) algorithm proposed by Wang et al. (2016) and discussed in Section 2.3.2 is used to optimize the rational policy due to its high sample efficiency and stability. Prior to application in the system, the ACER algorithm was evaluated in different testing scenarios with the original dialogue game and the original winning criterion for which the optimal policy is known in the scope of a master's thesis (Yang, 2019). The performance of the algorithm was tested on randomized argument structures with different sizes and on the annotated example structure on the topic *Marriage is an outdated institution*. The results showed that an optimal policy could be found for argument structures of all investigated sizes, although convergence to local optima occurred for argument structures with more than 30 arguments and more frequently with an increasing number of arguments in the structure. For the sake of illustration, the training process has been divided into *super-iterations*, each consisting of 400 dialogue episodes. Every 100 episodes the overall average reward and the average reward for the stance of the opponent was calculated. Figure 8.5 shows the average reward as a function of super-iterations for an example argument structure with 20 arguments.

Table 8.1: Artificial dialogue between the agents Alice and Bob with the rational policy trained with the ACER algorithm on the modified winning criterion. Content that corresponds to a component in the utilized argument structure is shown in italic font.

| Player | Utterance | Speech Acts |
|---|---|---|
| Alice | *You should visit this hotel.* | claim($\varphi_0$) |
| Bob | From my perspective *the facilities are bad.* | argue($\varphi_1 \Rightarrow \neg\varphi_0$) |
| Alice | It seems to me that *the Restaurant was great.*<br>*The restaurant downstairs is the best-kept secret in the area!* | argue_ex($\varphi_2 \Rightarrow \neg\varphi_1$)<br>argue($\varphi_3 \Rightarrow \varphi_2$) |
| Bob | Hm, pardon the topic switch but I'm still thinking about one of our earlier points. You said that *you should visit this hotel.* What makes you so sure? | why($\varphi_0$) |
| Alice | From my perspective *the hotel is very good in general.*<br>*This property has really improved since our last stay.*<br>*All in all it is a nice and affordable spot for sightseeing in the area.* | argue_ex($\varphi_4 \Rightarrow \varphi_0$)<br>argue_ex($\varphi_5 \Rightarrow \varphi_4$)<br>argue($\varphi_6 \Rightarrow \varphi_5$) |
| Bob | It's hard to disagree with that. I see your point there.<br>I'm still thinking about your general claim. From my perspective *the rooms are bad.*<br>*All 4 lower rooms were 80 degrees and above.* | concede($\varphi_4$)<br>argue_ex($\varphi_7 \Rightarrow \neg\varphi_0$)<br>argue($\varphi_8 \Rightarrow \varphi_7$) |

It can be seen that from the eighth super-iteration, the overall average reward dropped to 0 while the average reward for the stance of the opponent remained 20, indicating that any agent with this stance won the game. This convergence indicates a Nash equilibrium where both policies are optimal against each other, which was confirmed in test games against the (known) optimal rule-based strategy.

For the application in our system, we applied the ACER algorithm to the modified dialogue game with both the original as well as the modified winning criterion. We used five randomized argument structures with 20 arguments and the original winning criterion again as a proof-of-principle setup to confirm the feasibility of the algorithm in the new setup. Subsequently, we trained the final rational policy for the hotel structure on the modified winning criterion. For both setups, the minimal state space in Definition 20 and final rewards only according to Definition 16 were used. For the proof-of-principle setup we report that in all investigated instances, the optimal policy was found and confirmed in test games against the rule-based strategy. However, the experiments also indicate that the convergence of the algorithm depends on the random initialization of the utilized neural networks. Hence, repeated optimization with different random seeds is required to deal with the convergence to local optima. However, as the resulting policies can be compared in test games against each other, the best performing optimization can be identified even if no rule-based strategy is available for evaluation. Nevertheless, convergence to a Nash equilibrium cannot be guaranteed in these cases. For the investigated test case with randomized argument structures, convergence to the optimal policy in the first two trials is reported for three of the five structures. The maximum amount of optimizations required to beat the rule-based strategy was eight and occurred only once. This indicates that the shape of the underlying tree structure can influence the performance, i.e. the convergence of the applied algorithm. For the modified winning criterion, we again experimented with different configurations of the parameter set $\{\xi_j, Eff\}$. For the generation of example dialogues, the configuration $\xi_4 = 2$, $\xi_1 = \xi_2 = 1$, $\xi_3 = 0$ and $Eff = 4$ was used. An excerpt of such a dialogue is shown in Table 8.1.

### 8.3.3 Emotional Policy

In contrast to the rational policy, the emotional policy selects the next utterance in compliance with the emotion the system is supposed to convey. Since the effect of a certain emotion differs from person to person and is hence user-dependent, the emotional policy again utilizes real-time adaptation to react to individual users. In contrast to the previous version of the system, the resulting emotion is now also considered in the selection of the next game move.

Although the previously introduced approach is capable of taking the selected content of an utterance into account during the interaction, it has no influence on the actual selection of the moves. However, there is evidence that non-verbal inconsistencies lead to poor first impressions (Weisbuch et al., 2010), which is in line with van Kleef (2014) who showed the importance of appropriateness of emotions.

Therefore, the focus in the next adaptation scenario is on generating an emotion-based dialogue $d = m_1, ..., m_n$ conveying the most-influencing emotional tone by learning a policy $\pi^p_{emo}$ that also selects the next system utterance during interaction with the user. This approach is motivated by several reasons:

- The effectiveness of emotions is highly subjective (Kaptein et al., 2010; O'Keefe and Jackson, 1995).
- Adaptation of the strategy should be the main focus at an individual level (Fogg, 2007).
- There is evidence that persuasive messages are more successful when framed with the emotional state of the recipient (DeSteno et al., 2004).

As per DeSteno et al. (2004) this is mediated by biases induced by the own emotions of the recipient. Thus, we focus on learning the most effective affective state by modifying the agents' pleasure (valence) and arousal dimensions (Russell, 1980), which are used for selecting the next move $m_k \in M_d$ of all available moves at dialogue time step $k$. To allow for adaptation in real-time, we again utilize reinforcement learning.

To enable the learning of an emotional policy, the markov game formalism is modified as follows:

**Definition 31** (Modified State Space $S_{emo}$)**.** *Let $e^p_t \in [0,1]$ be the prediction of the persuasive effectiveness of the player to move $p \in I$ at time $t$ and with respect to component $\varphi_0$. In addition, let arousal and valence be discrete values $\in [-1,1]$, then the sub state of agent $p$ at time $t$ is defined as*

$$s^p_t := (arousal, valence, e^p_t). \tag{8.8}$$

*The resulting overall state of the game is given as*

$$s_t := \times_{p \in I} s^p_t\,. \tag{8.9}$$

Table 8.2: Artificial dialogue between the agents Alice and Bob with affect state (-0.5, 0.5). Content corresponding to a component in the utilized argument structure is shown in italic font.

| Player | Utterance | Speech Acts |
|---|---|---|
| Alice | *You should visit this hotel.* | claim($\varphi_0$) |
| Bob | I think *the facilities are bad.* | argue($\varphi_1 \Rightarrow \neg\varphi_0$) |
| Alice | I concede. | concede($\varphi_0$) |

It can be seen that the persuasive effectiveness of the players is now also included in the state which therefore includes information regarding the opponent's performance. In contrast to the previous version, the state is independent of argument components in the game moves. As the emotion is now represented by valence-arousal values in the state, the corresponding action space is defined as follows:

**Definition 32** (Modified Action Space $A_{emo}$)**.** *The action space $A_p$ for player $p \in I$ consists of an* INCREASE *and* DECREASE *action both for valence and arousal and an action* NONE *that leaves the state $s_t$ unchanged.*

The reward for this setup is analogous to Definition 30. Finally, the transition function updates the state according to the selected action and the resulting changes in the persuasive effectiveness.

For learning the emotional policy, we again use a linear function approximation along with a Fourier basis transformation (Konidaris et al., 2011). At every learning step $t$, the agent $p \in I$ selects one of the available actions $a_t^p \in A_p$ according to the current state $s_t \in S$ and policy $\pi_{emo}^p$ ($\varepsilon-$greedy with $\varepsilon = 0.05$), modifies its current emotional state $s_{t+1}$ and selects the next move(s) as follows:

Let $\mathbf{f} : M \to [-1,1] \times [-1,1]$ be the function that maps any move $m_k \in M$ into the 2D valence-arousal (VA) space. For that, we employ DEVA (Islam and Zibran, 2018), a text analysis tool designed for mapping any given sentence into the VA space (precision 82%, recall 78%) on the natural language representation of the corresponding move. For any move $m_k$ that includes an *argue_extend($\Phi_i$)* speech act, we define the corresponding mapping as

$$\mathbf{f}(m_k) := \frac{1}{2}(\mathbf{f}(m_i) + \mathbf{f}(m_j)), \tag{8.10}$$

where $m_i$ includes *argue($\Phi_i$)*, $m_j$ includes *argue($\Phi_j$)* and $\Phi_j$ extends $\Phi_i$. Further, let $\mathbf{g} : S \to [-1,1] \times [-1,1]$ be the respective function for the state space. The agent uses the emotional state $s_{t+1}$ to select the next move $m_k$ that is closest to the agent's state using the $L_2$ norm:

$$m_k = \min_{m \in M_d} \|\mathbf{f}(m) - \mathbf{g}(s_{t+1})\|_2 \,. \tag{8.11}$$

The obtained feedback signal $f$ is used to compute the current effectiveness $e_t$ and the corresponding reward signal according to Definition 30. However, relying on the distance metric only leads

**Algorithm 5:** Emotional Dialogue Generation

Init: $t = 0, a_t^p, s_{t+1}, \forall p \in I$
**foreach** $k = 1, \ldots, n$ **do**
  $p \leftarrow$ active player
  $s_{t+1} \leftarrow$ observe $s_{t+1} \in S$
  $m_k \leftarrow \pi_{ratio}^p$
  **if** $m_k$ $isType(argue(\_extend))$ **then**
    $m_k \leftarrow \min_{m \in M_d, argue} \|\mathbf{f}(m) - \mathbf{g}(s_{t+1})\|_2$
  apply $m_k$
  **if** $f_p$ **then**
    $\pi_{emo}^p \leftarrow$ update policy using $f$
    $a_{t+1}^p \leftarrow$ select next action.
    $s_{t+2} \leftarrow$ modify emotional state.
    $t \leftarrow t + 1$

to several issues, one of which is sketched in Table 8.2. Because an affect state of (-0.5, 0.5) is used, Alice concedes immediately after Bob's first argument. This seems odd at first glance, but can be explained with respect to the utilized argument structure:

- The second quadrant of the VA space only contains two arguments and only one argument ($\Phi_1 = \varphi_1 \Rightarrow \neg\varphi_0$) is allowed to be played by Bob.
- Since its representation in the 2D valence-arousal space is the closest one to (-0.5, 0.5), it is selected by Bob following the distance metric. However, Alice does not have any argument within quadrant two and the only arguments that she can make use of are within quadrant one.
- Computing the distance between all available arguments and the *concede* move, inevitably leads to the *concede* move as the closest one.

Conceding right away seems irrational or is at least not competitive and indicates that rationality should not be completely excluded from the emotional policy but should support the emotional policy with regard to some general rational decisions during argumentation, such as: When is the right time...

- ...to *concede* or *retract* an argument?
- ...to introduce a new argument?
- ...to request additional information?

To address this issue, we propose a new policy that considers both, the pre-trained rational policy and the adapted emotion the agents convey in the decision making. To this end, the rational policy $\pi_{ratio}$ is utilized to decide which move type comes next, and whenever an *argue(_extend)* move is chosen, the agent follows the metrics of the emotional policy according to Equation 8.11. The

Table 8.3: Artificial dialogue between the agents Alice and Bob with affect state (-0.5, -0.5) using Algorithm 5. Content corresponding to a component in the utilized argument structure is shown in italic font.

| Player | Utterance | Speech Acts |
|---|---|---|
| Alice | *You should visit this hotel.* | claim($\varphi_0$) |
| Bob | In my opinion *the rooms are bad.* | argue($\varphi_7 \Rightarrow \neg\varphi_0$) |
| Alice | I am not sure I understand what you are getting at. | why($\varphi_7$) |
| Bob | I think that's enough for the moment. I would rather focus on another aspect of the topic. It seems to me that *the facilities are bad.* | argue_ex($\varphi_1 \Rightarrow \neg\varphi_0$) |
| | *An elevator was broken during our last stay and it was most annoying, but did not greatly impact the overall experience.* | argue($\varphi_9 \Rightarrow \varphi_1$) |
| Alice | I think *the Restaurant was great.* | argue($\varphi_2 \Rightarrow \neg\varphi_1$) |

complete procedure is summarized in Algorithm 5 and an example dialogue generated with it is shown in Table 8.3.

It should be noted that (in contrast to the previous version) between any state transition $s_t \rightarrow s_{t+1}$, multiple moves $m_k, m_{k+1}, ..., m_{k+m}$ can be selected before the agent's emotional state changes again. This is a direct consequence of the herein employed framework and the communication language $L_c$ since user feedback regarding the overall perceived persuasiveness is only requested for speech acts of type *argue(_extend)*.

For an evaluation of the adaptive feasibility of the proposed approach, two scenarios were considered:

- An adaptation of one agent only to verify that it is able to increase its performance within the RL task (**S1**).
- An adaptation of both agents to verify that they are able to optimize their policy with respect to each other (**S2**).

The evaluation was done for 150 simulated users (see Fig. 8.6) with randomly assigned affective states $(x, y) \in$ VA space, i.e., the affective state that the agents had to learn. In addition, multiple dialogues were run with an overall minimum length of 40 RL time steps for each agent. For every time step $t$ the simulated user feedback $f_i$ is defined as the normalized distance between the agent's affective state and the user's affective state:

$$f_i = 1 - \frac{\|(x, y) - \mathbf{g}(s_{t+1})\|_2}{2\sqrt{2}}. \tag{8.12}$$

Figure 8.6 shows initial results of the simulation. For the case of **S1**, it can be seen that Agent 1 is able to increase its performance with respect to its opponent as expected. In addition, the results for **S2** show that both are able to keep the balance of their individual persuasive effectiveness and hence are capable of reacting appropriately to the strategy of their respective opponent.

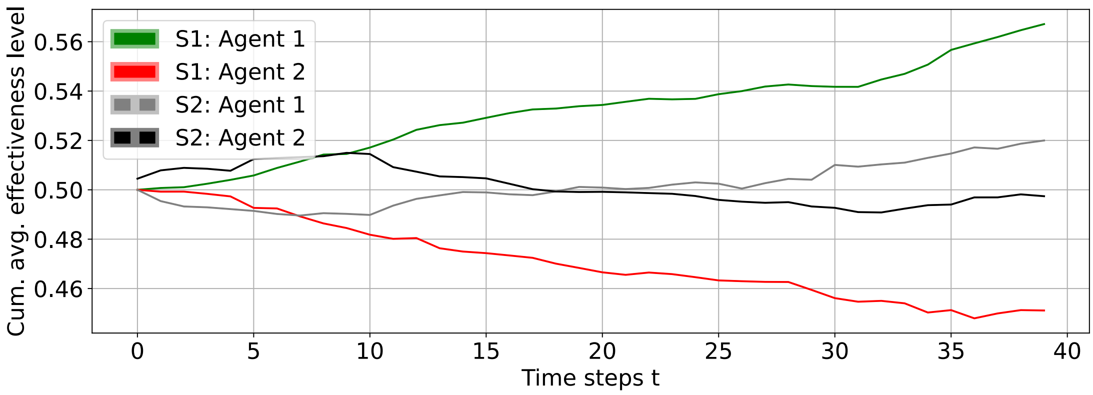


Figure 8.6: Initial simulation results of **S1** and **S2**.

### 8.3.4 Discussion

We have introduced a fully integrated version of our persuasive multi-agent system EVA 2.0 in which two agents engage in a multimodal discussion. The system utilizes argument structures extracted from reviews and a dialogue game for argumentation to structure the interaction. Besides, we have discussed two approaches to policy optimization within the dialogue game framework. The rational strategy is optimized on the objective winning criterion of the dialogue game and before the interaction by means of deep actor-critic RL. The emotional policy on the other hand utilizes a mapping of the available arguments into the valence-arousal space to select arguments that are close to the current emotion the system is supposed to convey. This emotion is adapted to user feedback to enable an individual and adaptive strategy. The adaptation is again approached by means of RL, although in this case the learning is done during interaction and in real-time. Both approaches were formally introduced and we discussed first results that indicate the feasibility of the proposed techniques.

Although the emotional policy was introduced in the context of a specific setup, it can be applied in other systems as well if the required user feedback is provided. Moreover, the approach also allows for switching between the emotional and the rational strategy during the ongoing interaction, if required. For example, in setups with implicit feedback that is prone to recognition errors (as discussed in Section 7.3), the rational policy can be used as a fallback strategy. The decision of which strategy to use is then made by the system based on the information that is available in the corresponding setup.

In comparison to the previous version, the system allows for much more complex dialogues between the virtual agents due to the use of the modified dialogue game that increases the freedom of choice for both players and the different dialogue policies. In the context of the present thesis, the system combines methods from all three investigated aspects of flexible argumentation, namely the modified dialogue game (including modified winning criterion) and multi-agent reinforcement learning for policy optimization in Chapter 5, argument structures retrieved with methods from Chapter 6 as well as the preference model and the adaptation approach discussed in Chapter 7. Besides the combination of all these techniques, the scaling problem of policy optimization discussed in Chapter 5 was addressed through the use of deep reinforcement learning and the adaptation approach was extended to utterance selection.

An approach that is not included in the system is the recognition of argument quality aspects from social signals introduced in Section 7.3. Since the user assumes the passive role of a judging audience in the current setup, explicit feedback is feasible and also more robust. However, the system functionalities do not depend on this explicit feedback and can be combined with implicit feedback as well. The two main tasks for including the proposed recognition model into systems of the herein discussed kind are the comparatively low estimation performance reported in Section 7.3 and the unknown effect of multiple avatars on this performance. Whereas the first point can be addressed through an improvement in the recognition model, the effect of multiple avatars on the overall recognition performance has yet to be investigated. In general, however, the approach for recognizing subjective argument quality aspects outputs the same information (convincingness) that is used in the discussed system for adaptation. It is therefore concluded that the system is conceptually compatible with the recognition approach proposed in Section 7.3.

## 8.4 Conclusion

Three implementations of argumentative dialogue systems that utilize methods proposed or investigated in the scope of this thesis have been introduced. The first one is concerned with persuasive dialogues between a human user and the system and combines the general setup of Chapter 4 with policy optimization based on multi-agent reinforcement learning in Chapter 5 and a virtual avatar. The system utilizes a drop-down menu in the interface for user input and extensions to enable natural language input in the future were proposed and discussed. In the second introduced system, the user assumes the role of a passive audience and the persuasive dialogue includes two virtual agents. The system again includes the general setup of Chapter 4 and a reinforcement learning-based strategy. In addition, the adaptation approach introduced in Chapter 7 was adapted to the multi-agent case and included in the system. The third system also considers agent-agent dialogues and a human user as a judging audience. In comparison to its predecessors, it utilizes the modified dialogue game introduced in Section 5.2, a rational strategy optimized by means of deep reinforcement learning and an emotional policy that extends the adaptation approach to an emotion-based utterance selection.

Overall, it can be seen that with one exception all approaches proposed in the present thesis were included into one or more systems or (in the case of the discussed topics) can be directly applied without further modifications on the system side. The approach that is not considered in the application systems is the automatic recognition of subjective argument quality aspects due to the current limitation in the estimation performance and the unknown effect of multiple avatars in the multi-agent case. On a conceptual level, however, the approach is also compatible with each of the three systems. Consequently, the general applicability of the proposed methods (alone and in combination) and hence their flexibility with respect to applications is concluded from the present chapter.

In addition, two issues identified throughout earlier chapters were addressed: First, the performance issues of reinforcement learning algorithms that rely on linear function approximation in the case of larger argument structures were addressed through the use of a sample efficient deep actor-critic method. In addition, the original reward defined in Chapter 5 relies solely on the formal winning criterion in the dialogue game and does not consider the subjective view of individual users. This issue was addressed through an extension that enables the combination of a pre-trained rational policy with the user-adaptive emotional one. Both extensions were assessed in a proof-of-principle evaluation to confirm their functionality. It is hence concluded that both contributions provide an application-oriented completion of the corresponding chapters.

In the context of the present thesis, the introduced systems were investigated as application instances for the proposed methods to enable flexible argumentation. Consequently, the discussion was focused on the formal and practical combination of these methods to confirm their functionality and flexibility. However, each of the systems addresses challenging argumentative tasks and an overall evaluation of the systems in interaction with real users is hence required in future work.

# 9 Conclusions and Future Work

Throughout this thesis, the task of *enabling conversational agents to argue about various topics in a way that is perceived as both natural and challenging by humans* was addressed. Following the approach of composite AI, the task was divided into three sub-tasks, namely challenging agent strategies, topic flexibility and user-adaptive argumentation. Due to the interdisciplinary nature of argumentation, the proposed approaches build on background and technologies from multiple research fields. This includes formal models from computational argumentation, various techniques from supervised and reinforcement learning, recognition techniques from affective computing as well as dialogue systems technology. As the thesis was conducted in the scope of a joint PhD program, the corresponding work was distributed over two research institutes, namely the Dialogue Systems Group at Ulm University and the Ubiquitous Computing Systems Lab at NAIST, Japan. In addition, cooperations with researchers from two different research groups as well as two bachelor's theses and one master's thesis supervised in the scope of this dissertation enabled the proposed solutions for the individual sub-tasks. The respective works, including the corresponding cooperations, are summarized in the following. The explicit thesis contributions are listed subsequently in terms of *theoretical*, *experimental* and *practical* contributions.

The strategy optimization based on markov games for dialogue games, the modification of the utilized dialogue game instantiation and the general experimental setup were developed at Ulm University and comprised one bachelor's thesis on argument representation (Langhammer, 2017) and one master's thesis on deep reinforcement learning in dialogue games for argumentation (Yang, 2019).

The use of argument search engines for enabling topic flexibility resulted from a cooperation with TU Darmstadt. The corresponding work was conducted at both NAIST and Ulm University: Whereas the evaluation experiment to assess the suitability of arguments retrieved by argument search engines for dialogue systems was done in Japan, the mapping of the retrieved arguments into a tree structure and the corresponding evaluation were done at Ulm University and included a bachelor's thesis on the automatic retrieval of argument structures from argument search results (Schindler, 2020).

The works on multimodal argumentation and adaptation approaches, including all system implementations of EVA were conducted in close cooperation with Augsburg University in the scope of the DFG Ratio project *EVA: Empowering Virtual Agents to improve their persuasiveness* at Ulm University. The second part on user-adaptive argumentation, namely the recognition of subjective argument quality aspects, was conducted at NAIST and constitutes the main contribution for the ubiquitous computing part of the thesis.

## 9.1 Contributions

This thesis has made contributions to different fields of research. They are listed in the following based on the division of *theoretical*, *experimental* and *practical* contributions. In addition to a summary of the proposed approaches, the reference of the corresponding publication is included.

### 9.1.1 Theoretical

Most of the theoretical work is related to the development of agent strategies. The first contribution in this regard is the reformulation of dialogue games for argumentation as a markov game (Rach et al., 2017b; Rach et al., 2018b) which allows for an optimization of the game strategy by means of reinforcement learning and without topic-specific training data or pre-defined opponent strategies. As the reformulation is based on the most general definition of dialogue games for argumentation provided in the referenced work, the approach can be applied to any instantiation of the formal framework. Since the development of specific strategies is based on the individual dialogue model, drawbacks in the one utilized for testing and implementation were addressed through a modified game protocol that allows for chaining multiple arguments in a single turn and hence increases the freedom of choice for the players (Rach et al., 2020b). In addition, a more advanced winning criterion was introduced to enable learning of more natural and advanced strategies (Rach et al., 2021c).

An additional theoretical contribution is the definition of meaningful evaluation categories for the assessment of argument quality aspects in the context of argumentative dialogue systems. The four categories are based on theoretical work from the fields of argument quality assessment and dialogue systems evaluation. They were utilized to assess the suitability of arguments retrieved with argument search engines for argumentative dialogue systems (Rach et al., 2020a) as well as (in part) for the automatic assessment of subjective argument quality aspects from non-verbal user cues (Rach et al., 2021a).

The third theoretical contribution is the development of a formal model to capture user preferences regarding arguments. It takes their relations and the effect of specific preferences on previous arguments based on these relations into account. The original model was proposed for use in an opinion-building system (Aicher et al., 2021) and subsequently transferred to the domain of agent-agent persuasion (Rach et al., 2019b).

Finally, the markov game framework was again utilized to combine the previously discussed (rational) argumentation strategy with a user-adaptive emotional policy that adjusts the emotion conveyed by the system with respect to user feedback. Both policies were formally combined into a new policy that utilizes the rational policy to select the next type of move and the specific content of the move based on the emotional one (Rach et al., 2021c).

### 9.1.2 Experimental

On the experimental side, the contributions are divided over all three investigated sub-tasks. For the development of agent strategies, numerical proof-of-principle experiments were conducted that confirmed the feasibility of the approach (Rach et al., 2018b; Rach et al., 2021c). In addition, two user studies assessed transcripts of agent-agent dialogues. In the first instance, the aim of the study was to reveal pending issues in the general evaluation setup and to confirm the compatibility of the utilized models. The second study was conducted to confirm that the modifications in the utilized dialogue game to address the previously identified issues have the intended effect.

For the sub-task of providing topic flexibility, an evaluation system was introduced that enables the assessment of arguments retrieved by argument search engines with respect to their applicability in dialogue systems. For a comparison of two state-of-the-art search engines in this setup, a baseline argument search system was proposed and an extensive user study was conducted (Rach et al., 2020a). The results revealed different strengths of the utilized search engines and both of them were able to beat the baseline in (at least) one category. Based on these results, a classifier was utilized to estimate relations between the retrieved arguments for a subsequent mapping into an argument (tree) structure and consequently to enable an automatic retrieval of complete structures. An extensive crowd-sourcing survey was then conducted to assess the coherence of artificial agent-agent dialogues that were generated with these automatically retrieved structures (Rach et al., 2021b). The results showed a dependency of the retrieval approach on the available data as well as the expected room for improvement in comparison to annotated structures. Nevertheless, the results were also in some instances surprisingly close to the ones achieved with an annotated structure which indicates the feasibility of the proposed approach.

As for user-adaptive argumentation, the main experimental contribution was the automatic estimation of subjective argument quality aspects from non-verbal cues (Rach et al., 2021a). In this work, social signal data recorded during the assessment of argument search engines was utilized to predict the corresponding labels in the categories *interesting* and *convincing*. Due to the novelty of the task, a human annotation was also conducted to compare the automatic estimation with human performance. The results indicated the difficulty of the task for both humans and machine learning model but also showed that, for the investigated data, both perform similarly. Based on these findings and a comparison to literature results on similar tasks, the general feasibility was concluded.

### 9.1.3 Practical

The main practical contributions are included in three implementations of the multimodal dialogue system EVA that utilize the approaches investigated in this thesis to different extents. The first version of the system (Rach et al., 2018c) allows for a direct interaction between the system and human users through a drop-down menu. The system architecture is based on the general experimental setup (argument structure, dialogue game and NLG) and utilizes a reinforcement learning-based strategy. Although not fully implemented, yet, possible extensions that allow for a natural language understanding of the user input were discussed (Rach et al., 2018a).

The second version of the system considers agent-agent argumentation to display different persuasion strategies (Weber et al., 2020b). In this setup, users assume the role of a (judging) audience and provide explicit feedback. Again, the architecture is based on the original (unmodified) dialogue game and utilizes a reinforcement learning-based argumentation strategy. In the final version of the system (Rach et al., 2021c), the original dialogue game was replaced with the modified one and a corresponding strategy was trained again with reinforcement learning. In addition, a merged policy that considers both, the pre-trained rational policy and the user-adaptive emotional policy for selecting the next utterance was included. All versions of EVA can directly apply argument structures retrieved with the herein introduced retrieval approaches.

In addition, a semi-automatic approach to extract argument structures from reviews based on sentiment analysis labels was introduced (Weber et al., 2020a). It is based on findings from the field of argument mining and operates on the annotations of the *SemEval-2015 Task 12* data set. The approach was motivated by the requirements of the above-discussed systems, i.e. high-quality structures on non-opinion based topics, and hence listed under practical contributions.

## 9.2 Conclusion and Future Steps

In conclusion, all herein proposed approaches extend the state of the art and contribute to the modelling of the cognitive capacity to argue in conversational agents. The overall task was divided into the sub-tasks of *challenging agent strategies*, *topic flexibility* and *user adaptation* and flexible methods applicable in different scenarios and systems were proposed for all three. A valuable insight gained in the process of this thesis is that the development of a complex human-like capacity like argumentation cannot be approached by choosing one general technique (for example data-driven or formal model) but instead by combining appropriate techniques in a way that enables them to work together efficiently. This requires knowledge about multiple different technologies, tools and research areas. An interdisciplinary approach that includes cooperation and discussion with researchers from other fields is hence not just beneficial but, in fact, necessary. In the following, an individual conclusion alongside an outlook on future steps for each of the addressed sub-tasks is provided.

The approach proposed for the development of a challenging agent strategy does not depend on a training corpus with conversational data or pre-defined opponent strategies. In addition, the general reformulation of dialogue games for argumentation as markov games enables the application of the proposed approach in multiple scenarios. It is hence concluded that the goal of providing flexibility in this regard was accomplished. A reasonable next step is the investigation of meaningful rewards in the context of different application setups. In addition, many dialogue games for argumentation are still too restrictive for a free natural language interaction with human users. As could be shown for the case of the herein utilized dialogue game, careful modifications of existing protocols can significantly influence the perception of the resulting dialogues. The task for future work is hence to enable the maximum amount of freedom in a given framework while preserving the required formal properties. In this regard, a detailed analysis of the differences between human argumentation and argumentation generated with the corresponding framework will be crucial for identifying the limiting restrictions that need to be adapted.

For the sub-task of topic flexibility, the proposed combination of argument search engines with relation classification enables a corresponding system to discuss arbitrary topics on which the search engine can find suitable arguments. The dependency on state-of-the-art technology thereby ensures that the approach directly benefits from future developments made in this area. It is hence concluded that the envisioned flexibility in this regard could be provided. However, a side-effect of the large scale argument mining approach is that the retrieved argument components include formulations that are in some cases difficult to combine in a single utterance. In addition, spelling errors and inappropriate language can also lead to confusion in the dialogue. Future work should hence focus on paraphrasing and corrective approaches that can provide a 'cleaned' version of the retrieved components. In addition, the combination with generative approaches to fill gaps in the retrieved structure is also a possible extension that allows for a dynamic adaptation of the argument database during the interaction.

The approaches proposed for user-adaptive argumentation are flexible in the sense that neither the introduced preference model nor the recognition of subjective argument quality aspects depends on a specific domain. In the case of the subjective argument quality aspects, it does not even depend on a specific argument structure. However, the application of the proposed recognition techniques in actual systems requires a larger training corpus to ensure a robust and high-quality estimate of the user opinion. In addition, the use of prosodic features in scenarios with a more active user role is likely to improve the performance. Also, the proposed real-time adaptation approach is tailored to the specific task and extending approaches can be investigated.

An additional aspect of human-computer argumentation that was only discussed in the context of specific applications is the recognition of user input, i.e. the natural language understanding of arguments. However, the retrieval approaches that were herein utilized to acquire arguments from external sources can in future work also be utilized to interpret the user input in the context of an existing argument structure. Nevertheless, flexible NLU approaches that can be applied in multiple different scenarios are yet to be investigated.

In general, there are several promising directions for further developing argumentation with conversational agents. The rapid improvements in the fields of dialogue systems technology, computational argumentation and natural language processing will enable additional applications with increasing complexity. In particular, hybrid systems that combine argumentation with other conversational and cognitive capacities as for example question answering (Sakai et al., 2018a) are a challenging but promising next step. Applications that are of particular interest in this regard are virtual assistants, for example in smart homes, and recommendation systems. In both cases, the use of argumentation will allow the corresponding system to better react to individual user needs by discussing different aspects of the investigated issue or request directly with the corresponding user. An additional direction with high potential is the further development of multimodal argumentation. As indicated by Wachsmuth et al. (2017b), there are several aspects that contribute to high-quality argumentation and are yet to be integrated into argumentative systems. In this regard, the further combination of different modalities (and corresponding modules) will be of particular interest for the task of achieving human-like argumentation with conversational agents. It is hence fair to assume that argumentation in dialogue systems will continue to be both promising and challenging in the future development of AI.

# Bibliography


Abbott, R., Ecker, B., Anand, P., & Walker, M. (2016). Internet argument corpus 2.0: An sql schema for dialogic social media and the corpora to go with it. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 4445–4452.

Ævarsson, K. (2006). *Human-computer debating system evaluation and development* (Doctoral dissertation). University of Akureyri, Iceland.

Aharoni, E., Polnarov, A., Lavee, T., Hershcovich, D., Levy, R., Rinott, R., Gutfreund, D., & Slonim, N. (2014). A benchmark dataset for automatic detection of claims and evidence in the context of controversial topics. *Proceedings of the First Workshop on Argumentation Mining*, 64–68.

Ahmed, T., Singh, R., Pandey, A. K., & Singh, S. K. (2020). A cognitive model to predict human interest in smart environments. *Computer Communications*, *161*, 1–9.

Aicher, A., Rach, N., Minker, W., & Ultes, S. (2021). Opinion building based on the argumentative dialogue system bea. *Increasing Naturalness and Flexibility in Spoken Dialogue Interaction: 10th International Workshop on Spoken Dialogue Systems*, 307–318. `https://doi.org/10.1007/978-981-15-9323-9_27`

Ajjour, Y., Wachsmuth, H., Kiesel, J., Potthast, M., Hagen, M., & Stein, B. (2019). Data acquisition for argument search: The args. me corpus. *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*, 48–59.

Alahmari, S., Yuan, T., & Kudenko, D. (2019). Reinforcement learning for dialogue game based argumentation. *Proceedings of the 19th Workshop on Computational Models of Natural Argument*, 29–37.

Amgoud, L., & Ben-Naim, J. (2018). Weighted bipolar argumentation graphs: Axioms and semantics. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, 5194–5198.

Asai, S., Yoshino, K., Shinagawa, S., Sakti, S., & Nakamura, S. (2020). Emotional speech corpus for persuasive dialogue system. *Proceedings of The 12th Language Resources and Evaluation Conference*, 491–497.

Atkinson, K., Baroni, P., Giacomin, M., Hunter, A., Prakken, H., Reed, C., Simari, G., Thimm, M., & Villata, S. (2017). Towards artificial argumentation. *AI magazine*, *38*(3), 25–36.

Baltrusaitis, T., Zadeh, A., Lim, Y. C., & Morency, L.-P. (2018). Openface 2.0: Facial behavior analysis toolkit. *13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, 59–66.

Bar-Haim, R., Bhattacharya, I., Dinuzzo, F., Saha, A., & Slonim, N. (2017). Stance classification of context-dependent claims. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 251–261.

Bar-Haim, R., Eden, L., Friedman, R., Kantor, Y., Lahav, D., & Slonim, N. (2020). From arguments to key points: Towards automatic argument summarization. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4029–4039.

Barlier, M., Perolat, J., Laroche, R., & Pietquin, O. (2015). Human-machine dialogue as a stochastic game. *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2–11.

Baroni, P., Caminada, M., & Giacomin, M. (2011). An introduction to argumentation semantics. *The knowledge engineering review*, *26*(4), 365–410.

Bechhofer, S. (2009). Owl: Web ontology language. *Encyclopedia of database systems* (pp. 2008–2009). Springer.

Bench-Capon, T. J. (1998). Specification and implementation of toulmin dialogue game. *Proceedings of JURIX*, 5–19.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Bentahar, J., Moulin, B., & Bélanger, M. (2010). A taxonomy of argumentation models used for knowledge representation. *Artificial Intelligence Review*, *33*(3), 211–259.

Besnard, P., & Hunter, A. (2008). *Elements of argumentation*. MIT press Cambridge.

Bevendorff, J., Stein, B., Hagen, M., & Potthast, M. (2018). Elastic chatnoir: Search engine for the clueweb and the common crawl. *European Conference on Information Retrieval*, 820–824.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python: Analyzing text with the natural language toolkit*. O'Reilly Media, Inc.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory*, 144–152.

Bousmalis, K., Mehu, M., & Pantic, M. (2009). Spotting agreement and disagreement: A survey of nonverbal audiovisual cues and tools. *3rd international conference on affective computing and intelligent interaction and workshops*, 1–9.

Bousmalis, K., Mehu, M., & Pantic, M. (2013). Towards the automatic detection of spontaneous agreement and disagreement based on nonverbal behaviour: A survey of related cues, databases, and tools. *Image and vision computing*, *31*(2), 203–221.

Bousmalis, K., Morency, L.-P., & Pantic, M. (2011). Modeling hidden dynamics of multimodal cues for spontaneous agreement and disagreement recognition. *IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 746–752.

Bowling, M., & Veloso, M. (2001). Rational and convergent learning in stochastic games. *International joint conference on artificial intelligence*, *17*(1), 1021–1026.

Braun, D., Mendez, A. H., Matthes, F., & Langen, M. (2017). Evaluating natural language understanding services for conversational question answering systems. *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, 174–185.

Bunk, T., Varshneya, D., Vlasov, V., & Nichol, A. (2020). Diet: Lightweight language understanding for dialogue systems. *arXiv preprint arXiv:2004.09936*.

Carstens, L., & Toni, F. (2015). Towards relation based argumentation mining. *Proceedings of the 2nd Workshop on Argumentation Mining*, 29–34.

Casanueva, I., Hain, T., Christensen, H., Marxer, R., & Green, P. (2015). Knowledge transfer between speakers for personalised dialogue management. *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 12–21.

Cayrol, C., & Lagasquie-Schiex, M.-C. (2005). On the acceptability of arguments in bipolar argumentation frameworks. *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, 378–389.

Cayrol, C., & Lagasquie-Schiex, M.-C. (2009). Bipolar abstract argumentation systems. *Argumentation in artificial intelligence* (pp. 65–84). Springer.

Chalaguine, L. A., & Hunter, A. (2020). A persuasive chatbot using a crowd-sourced argument graph and concerns. *Computational Models of Argument: Proceedings of COMMA 2020*, *326*, 9–20.

Chalaguine, L. A., Hunter, A., Hamilton, F. L., & Potts, H. W. W. (2019). Impact of argument type and concerns in argumentation with a chatbot. *31st IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, 1557–1562.

Chalaguine, L. A., & Hunter, A. (2019). Knowledge acquisition and corpus for argumentation-based chatbots. *Proceedings of the 3rd Workshop on Advances In Argumentation In Artificial Intelligence*, 1–14.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321–357.

Chen, S., Khashabi, D., Callison-Burch, C., & Roth, D. (2019). PerspectroScope: A window to the world of diverse perspectives. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 129–134.

Chernodub, A., Oliynyk, O., Heidenreich, P., Bondarenko, A., Hagen, M., Biemann, C., & Panchenko, A. (2019). Targer: Neural argument mining at your fingertips. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 195–200.

Choi, E., He, H., Iyyer, M., Yatskar, M., Yih, W.-t., Choi, Y., Liang, P., & Zettlemoyer, L. (2018). QuAC: Question answering in context. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2174–2184.

Cocarascu, O., & Toni, F. (2016). Detecting deceptive reviews using argumentation. *Proceedings of the 1st International Workshop on AI for Privacy and Security*, 1–8.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, *20*(1), 37–46.

Degris, T., White, M., & Sutton, R. (2012). Off-policy actor-critic. *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, 457–464.

Deriu, J., Rodrigo, A., Otegi, A., Echegoyen, G., Rosset, S., Agirre, E., & Cieliebak, M. (2020). Survey on evaluation methods for dialogue systems. *Artificial Intelligence Review*, *54*, 755–810.

DeSteno, D., Petty, R. E., Rucker, D. D., Wegener, D. T., & Braverman, J. (2004). Discrete emotions and persuasion: The role of emotion-induced expectancies. *Journal of personality and social psychology*, *86*(1), 43–56.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87.

Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, *77*(2), 321–357.

Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, *6*(3-4), 169–200.

Engel, Y., Mannor, S., & Meir, R. (2005). Reinforcement learning with gaussian processes. *Proceedings of the 22nd international conference on Machine learning*, 201–208.

Engelbrecht, K.-P., Gödde, F., Hartard, F., Ketabdar, H., & Möller, S. (2009). Modeling user satisfaction with hidden markov models. *Proceedings of the SIGDIAL 2009 Conference*, 170–177.

Filar, J., & Vrieze, K. (2012). *Competitive markov decision processes*. Springer Science & Business Media.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological bulletin*, *76*(5), 378.

Fogg, B. (2007). *Mobile persuasion: 20 perspectives on the future of behavior change*. Stanford Captology Media.

Galitsky, B. (2019). Enabling a bot with understanding argumentation and providing arguments. *Developing enterprise chatbots* (pp. 465–532). Springer.

Georgila, K., & Traum, D. (2011). Reinforcement learning of argumentation dialogue policies in negotiation. *Twelfth Annual Conference of the International Speech Communication Association*, 2073–2076.

Gleize, M., Shnarch, E., Choshen, L., Dankin, L., Moshkowich, G., Aharonov, R., & Slonim, N. (2019). Are you convinced? choosing the more convincing evidence with a Siamese network. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 967–976.

Gold, D., Kovatchev, V., & Zesch, T. (2019). Annotating and analyzing the interactions between meaning relations. *Proceedings of the 13th Linguistic Annotation Workshop*, 26–36.

Gretz, S., Friedman, R., Cohen-Karlik, E., Toledo, A., Lahav, D., Aharonov, R., & Slonim, N. (2020). A large-scale dataset for argument quality ranking: Construction and analysis. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(05), 7805–7813.

Habernal, I., Eckle-Kohler, J., & Gurevych, I. (2014). Argumentation mining on the web from information seeking perspective. *Proceedings of the Workshop on Frontiers and Connections between Argumentation Theory and Natural Language Processing*.

Habernal, I., & Gurevych, I. (2016a). What makes a convincing argument? empirical analysis and detecting attributes of convincingness in web argumentation. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 1214–1223.

Habernal, I., & Gurevych, I. (2016b). Which argument is more convincing? analyzing and predicting convincingness of web arguments using bidirectional lstm. *Proceedings of the*

*54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, *1*, 1589–1599.

Hadoux, E., Beynier, A., Maudet, N., Weng, P., & Hunter, A. (2015). Optimization of probabilistic argumentation with markov decision models. *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2004–2010.

Hadoux, E., & Hunter, A. (2017). Strategic sequences of arguments for persuasion using decision trees. *Thirty-First AAAI Conference on Artificial Intelligence*, 1128–1134.

Hadoux, E., & Hunter, A. (2019). Comfort or safety? gathering and using the concerns of a participant for better persuasion. *Argument & Computation*, *10*(2), 113–147.

Hara, S., Kitaoka, N., & Takeda, K. (2010). Estimation method of user satisfaction using n-gram-based dialog history model for spoken dialog system. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 78–83.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, *1*(1), 77–89.

Higashinaka, R., Minami, Y., Dohsaka, K., & Meguro, T. (2010). Modeling user satisfaction transitions in dialogues from overall ratings. *Proceedings of the SIGDIAL 2010 Conference*, 18–27.

Higashinaka, R., Sakai, K., Sugiyama, H., Narimatsu, H., Arimoto, T., Fukutomi, T., Matsui, K., Ijima, Y., Ito, H., Araki, S. et al. (2017). Argumentative dialogue system based on argumentation structures. *Proceedings of the 21st Workshop on the Semantics and Pragmatics of Dialogue*.

Hirayama, T., Dodane, J.-B., Kawashima, H., & Matsuyama, T. (2010). Estimates of user interest using timing structures between proactive content-display updates and eye movements. *IEICE Transaction on Information and Systems*, *93*(6), 1470–1478.

Hunter, A. (2014). Probabilistic strategies in dialogical argumentation. *International Conference on Scalable Uncertainty Management*, 190–202.

Hunter, A., Chalaguine, L., Czernuszenko, T., Hadoux, E., & Polberg, S. (2019). Towards computational persuasion via natural language argumentation dialogues. *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*, 18–33.

Hutto, C., & Gilbert, E. (2015). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 216–225.

Islam, M. R., & Zibran, M. F. (2018). DEVA: Sensing emotions in the valence arousal space in software engineering text. *Proceedings of the 33rd Annual ACM Symposium on Applied Computing - SAC '18*, 1536–1543.

Janarthanam, S., & Lemon, O. (2010). Learning to adapt to unknown users: Referring expression generation in spoken dialogue systems. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 69–78.

Jeon, J. H., Xia, R., & Liu, Y. (2014). Level of interest sensing in spoken dialog using decision-level fusion of acoustic and lexical evidence. *Computer Speech & Language*, *28*(2), 420–433.

Kaptein, M., Lacroix, J., & Saini, P. (2010). Individual differences in persuadability in the health promotion domain. *International Conference on Persuasive Technology*, 94–105.

Keogh, E., & Mueen, A. (2011). Curse of dimensionality. *Encyclopedia of machine learning* (pp. 257–258). Springer.

Konda, V. R., & Tsitsiklis, J. N. (2000). Actor-critic algorithms. *Advances in neural information processing systems*, 1008–1014.

Konidaris, G., Osentoski, S., & Thomas, P. (2011). Value function approximation in reinforcement learning using the fourier basis. *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, *25*(1), 380–385.

Krapinger, G. (1999). Aristoteles: Rhetorik. *Übersetzt und herausgegeben von Gernot Krapinger. Stuttgart: Reclam.*

Kraus, M., Wagner, N., & Minker, W. (2020). Effects of proactive dialogue strategies on human-computer trust. *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, 107–116.

Kwon, O.-W., Lee, K., Roh, Y.-H., Huang, J.-X., Choi, S.-K., Kim, Y.-K., Jeon, H. B., Oh, Y. R., Lee, Y.-K., Kang, B. O. et al. (2015). Genietutor: A computer-assisted second-language learning system based on spoken language understanding. *Natural language dialog systems and intelligent assistants* (pp. 257–262). Springer.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Langhammer, S. (2017). *A debating ontology for argumentative dialogue systems* (Bachelor's thesis). Ulm University.

Lauscher, A., Ng, L., Napoles, C., & Tetreault, J. (2020). Rhetoric, logic, and dialectic: Advancing theory-based argument quality assessment in natural language processing. *Proceedings of the 28th International Conference on Computational Linguistics*, 4563–4574.

Lawrence, J., & Reed, C. (2020). Argument mining: A survey. *Computational Linguistics*, *45*(4), 765–818.

Le, D. T., Nguyen, C.-T., & Nguyen, K. A. (2018). Dave the debater: A retrieval-based and generative argumentative dialogue agent. *Proceedings of the 5th Workshop on Argument Mining*, 121–130.

Levy, R., Bogin, B., Gretz, S., Aharonov, R., & Slonim, N. (2018). Towards an argumentative content search engine using weak supervision. *Proceedings of the 27th International Conference on Computational Linguistics*, 2066–2081.

Li, X., Chen, Y.-N., Li, L., Gao, J., & Celikyilmaz, A. (2017). End-to-end task-completion neural dialogue systems. *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 733–743.

Li, Y., McLean, D., Bandar, Z. A., O'shea, J. D., & Crockett, K. (2006). Sentence similarity based on semantic nets and corpus statistics. *IEEE transactions on knowledge and data engineering*, *18*(8), 1138–1150.

Lin, L.-J. (1992). Self-improving reactive agents based on reinforcement learning, planning and teaching. *Machine learning*, *8*(3-4), 293–321.

Lippi, M., & Torroni, P. (2016). Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology (TOIT)*, *16*(2), 1–25.

Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. *Proceedings of the Eleventh International Conference on Machine Learning*, 157–163.

Mackenzie, J. (1979). Begging the question in non-cumulative systems. *Journal of Philosophical Logic*, *8*, 117–133.

Mahmood, A. R., & Sutton, R. S. (2015). Off-policy learning based on weighted importance sampling with linear computational complexity. *UAI'15: Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*, 552–561.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60.

Mann, W. C. (1988). Dialogue games: Conventions of human interaction. *Argumentation*, *2*(4), 511–532.

Matsuda, Y., Fedotov, D., Takahashi, Y., Arakawa, Y., Yasumoto, K., & Minker, W. (2018). Emotour: Estimating emotion and satisfaction of users based on behavioral cues and audiovisual data. *Sensors*, *18*(11).

Miehle, J., Minker, W., & Ultes, S. (2018). Exploring the impact of elaborateness and indirectness on user satisfaction in a spoken dialogue system. *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization (UMAP)*, 165–172.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to wordnet: An on-line lexical database. *International journal of lexicography*, *3*(4), 235–244.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Mo, K., Zhang, Y., Li, S., Li, J., & Yang, Q. (2018). Personalizing a dialogue system with transfer reinforcement learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *32*(1), 5317–5324.

Munos, R., Stepleton, T., Harutyunyan, A., & Bellemare, M. G. (2016). Safe and efficient off-policy reinforcement learning. *NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems*, 1054–1062.

O'Keefe, D. J., & Jackson, S. (1995). Argument quality and persuasive effects: A review of current approaches. *Argumentation and values: Proceedings of the 9th Alta conference on argumentation*, 88–92.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Communication and persuasion* (pp. 1–24). Springer.

Poggi, I., D'Errico, F., & Vincze, L. (2011). Agreement and its multimodal communication in debates: A qualitative analysis. *Cognitive Computation*, *3*(3), 466–479.

Pollock, J. L. (1987). Defeasible reasoning. *Cognitive science*, *11*(4), 481–518.

Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., & Androutsopoulos, I. (2015). Semeval-2015 task 12: Aspect based sentiment analysis. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 486–495.

Potthast, M., Gienapp, L., Euchner, F., Heilenkötter, N., Weidmann, N., Wachsmuth, H., Stein, B., & Hagen, M. (2019). Argument search: Assessing argument relevance. *Proceedings of the 42Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1117–1120.

Prakken, H. (2000). On dialogue systems with speech acts, arguments, and counterarguments. *Proceedings of the 7th European Workshop on Logic for Artificial Intelligence (JELIA' 2000)*, 224–238.

Prakken, H. (2005). Coherence and flexibility in dialogue games for argumentation. *Journal of logic and computation*, *15*(6), 1009–1040.

Prakken, H. (2006). Formal systems for persuasion dialogue. *The knowledge engineering review*, *21*(2), 163–188.

Prakken, H. (2018). Historical overview of formal argumentation. *Handbook of formal argumentation* (pp. 73–141). College Publications.

Prakken, H., & Vreeswijk, G. (2001). Logics for defeasible argumentation. *Handbook of philosophical logic* (pp. 219–318). Springer.

Rach, N., Langhammer, S., Minker, W., & Ultes, S. (2019a). Utilizing argument mining techniques for argumentative dialogue systems. *9th International Workshop on Spoken Dialogue System Technology*, 131–142. https://doi.org/10.1007/978-981-13-9443-0_12

Rach, N., Matsuda, Y., Daxenberger, J., Ultes, S., Yasumoto, K., & Minker, W. (2020a). Evaluation of argument search approaches in the context of argumentative dialogue systems. *Proceedings of the 12th Language Resources and Evaluation Conference*, 513–522.

Rach, N., Matsuda, Y., Ultes, S., Minker, W., & Yasumoto, K. (2021a). Estimating subjective argument quality aspects from social signals in argumentative dialogue systems. *IEEE Access*, *9*, 11610–11621. https://doi.org/10.1109/ACCESS.2021.3051526

Rach, N., Minker, W., & Ultes, S. (2017a). Interaction quality estimation using long short-term memories. *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, 164–169. https://doi.org/10.18653/v1/W17-5520

Rach, N., Minker, W., & Ultes, S. (2017b). Towards an argumentative dialogue system. *Proceedings of the 17th Workshop on Computational Models of Natural Argument*, 27–29.

Rach, N., Minker, W., & Ultes, S. (2018a). Handling unknown user arguments in argumentative dialogue systems. *Proceedings of the 32nd International BCS Human Computer Interaction Conference*, 1–4. https://doi.org/10.14236/ewic/HCI2018.67

Rach, N., Minker, W., & Ultes, S. (2018b). Markov games for persuasive dialogue. *Computational Models of Argument: Proceedings of COMMA 2018*, *305*, 213–220. https://doi.org/10.3233/978-1-61499-906-5-213

Rach, N., Minker, W., & Ultes, S. (2020b). Increasing the naturalness of an argumentative dialogue system through argument chains. *Computational Models of Argument: Proceedings of COMMA 2020*, *326*, 331–338. https://doi.org/10.3233/FAIA200516

Rach, N., Schindler, C., Feustel, I., Daxenberger, J., Minker, W., & Ultes, S. (2021b). From argument search to argumentative dialogue: A topic-independent approach to argument acquisition for dialogue systems. *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 368–379.

Rach, N., Weber, K., Aicher, A., Lingenfelser, F., André, E., & Minker, W. (2019b). Emotion recognition based preference modelling in argumentative dialogue systems. *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 838–843. https://doi.org/10.1109/PERCOMW.2019.8730650

Rach, N., Weber, K., Pragst, L., André, E., Minker, W., & Ultes, S. (2018c). Eva: A multimodal argumentative dialogue system. *Proceedings of the 2018 on International Conference on Multimodal Interaction*, 551–552. https://doi.org/10.1145/3242969.3266292

Rach, N., Weber, K., Yang, Y., Ultes, S., André, E., & Minker, W. (2021c). Eva 2.0: Emotional and rational multimodal argumentation between virtual agents. *it-Information Technology*, *63*(1), 17–30. https://doi.org/10.1515/itit-2020-0050

Rakshit, G., Bowden, K. K., Reed, L., Misra, A., & Walker, M. (2019). Debbie, the debate bot of the future. *Advanced social interaction with agents : 8th international workshop on spoken dialog systems* (pp. 45–52). Springer International Publishing.

Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian processes for machine learning* (Vol. 1). MIT press Cambridge.

Reed, C., & Norman, T. (2003). *Argumentation machines: New frontiers in argument and computation* (Vol. 9). Springer Science & Business Media.

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *The 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP 2019)*, 3973–3983.

Reimers, N., Schiller, B., Beck, T., Daxenberger, J., Stab, C., & Gurevych, I. (2019). Classification and clustering of arguments with contextualized word embeddings. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 567–578.

Rieser, V., & Lemon, O. (2011). *Reinforcement learning for adaptive dialogue systems: A data-driven methodology for dialogue management and natural language generation*. Springer Science & Business Media.

Ritschel, H., Baur, T., & André, E. (2017). Adapting a robot's linguistic style based on socially-aware reinforcement learning. *26th International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 378–384.

Rodgers, J. L., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59–66.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, *65*(6), 386.

Rosenfeld, A., & Kraus, S. (2016). Strategical argumentative agent for human persuasion. *Proceedings of the Twenty-second European Conference on Artificial Intelligence*, 320–328.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Sakai, K., Higashinaka, R., Yoshikawa, Y., Ishiguro, H., & Tomita, J. (2018a). Introduction method for argumentative dialogue using paired question-answering interchange about personality. *Proceedings of the 19th annual SIGDIAL Meeting on discourse and dialogue*, 70–79.

Sakai, K., Higashinaka, R., Yoshikawa, Y., Ishiguro, H., & Tomita, J. (2020). Hierarchical argumentation structure for persuasive argumentative dialogue generation. *IEICE Transactions on Information and Systems*, *103*(2), 424–434.

Sakai, K., Inago, A., Higashinaka, R., Yoshikawa, Y., Ishiguro, H., & Tomita, J. (2018b). Creating large-scale argumentation structures for dialogue systems. *Proceedings of the Eleventh*

*International Conference on Language Resources and Evaluation (LREC 2018)*, 3975–3980.

Sasaka, Y., Ogawa, T., & Haseyama, M. (2016). Multimodal interest level estimation via variational bayesian mixture of robust cca. *Proceedings of the 24th ACM International Conference on Multimedia*, 387–391.

Schatzmann, J., Georgila, K., & Young, S. (2005). Quantitative evaluation of user simulation techniques for spoken dialogue systems. *Proceedings of the 6th SIGdial Workshop on Discourse and Dialogue*, 45–54.

Schatzmann, J., Thomson, B., Weilhammer, K., Ye, H., & Young, S. (2007). Agenda-based user simulation for bootstrapping a pomdp dialogue system. *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*, 149–152.

Schiller, B., Daxenberger, J., & Gurevych, I. (2021). Aspect-controlled neural argument generation. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 380–396.

Schindler, C. (2020). *Argumentative relation classification for argumentative dialogue systems* (Bachelor's thesis). Ulm University.

Schmitt, A., & Ultes, S. (2015). Interaction quality: Assessing the quality of ongoing spoken dialog interaction by experts—and how it relates to user satisfaction. *Speech Communication*, *74*, 12–36.

Schuller, B., Müller, R., Hörnler, B., Höthker, A., Konosu, H., & Rigoll, G. (2007). Audiovisual recognition of spontaneous interest within conversations. *Proceedings of the 9th international conference on Multimodal interfaces*, 30–37.

Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P. (2015). Trust region policy optimization. *International conference on machine learning*, 1889–1897.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive Robotics: The influence of robot gender on human behavior. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2563–2568.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. et al. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489.

Slonim, N., Bilu, Y., Alzate, C., Bar-Haim, R., Bogin, B., Bonin, F., Choshen, L., Cohen-Karlik, E., Dankin, L., Edelstein, L. et al. (2021). An autonomous debating system. *Nature*, *591*(7850), 379–384.

Sprent, P. (2011). Fisher exact test. *International encyclopedia of statistical science*, 524–525.

Stab, C., Daxenberger, J., Stahlhut, C., Miller, T., Schiller, B., Tauchmann, C., Eger, S., & Gurevych, I. (2018a). Argumentext: Searching for arguments in heterogeneous sources. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: System Demonstrations*, 21–25.

Stab, C., & Gurevych, I. (2014a). Annotating argument components and relations in persuasive essays. *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 1501–1510.

Stab, C., & Gurevych, I. (2014b). Identifying argumentative discourse structures in persuasive essays. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 46–56.

Stab, C., & Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, *43*(3), 619–659.

Stab, C., Miller, T., Schiller, B., Rai, P., & Gurevych, I. (2018b). Cross-topic argument mining from heterogeneous sources. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, *Long Papers*, 3664–3674.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). MIT press Cambridge.

Sutton, R. S., McAllester, D. A., Singh, S. P., Mansour, Y. et al. (1999). Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems*, *99*, 1057–1063.

Thakur, N., Reimers, N., Daxenberger, J., & Gurevych, I. (2021). Augmented SBERT: Data Augmentation Method for Improving Bi-Encoders for Pairwise Sentence Scoring Tasks. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 296–310.

Toledo, A., Gretz, S., Cohen-Karlik, E., Friedman, R., Venezian, E., Lahav, D., Jacovi, M., Aharonov, R., & Slonim, N. (2019). Automatic argument quality assessment - new datasets and methods. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 5624–5634.

Tomimasu, S., & Araki, M. (2016). Assessment of users' interests in multimodal dialog based on exchange unit. *Proceedings of the Workshop on Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction*, 33–37.

Toulmin, S. E. (2003). *The uses of argument* (2nd ed.). Cambridge university press.

Trautmann, D., Daxenberger, J., Stab, C., Schütze, H., & Gurevych, I. (2020). Fine-grained argument unit recognition and classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(05), 9048–9056.

Turing, A. M. (2009). Computing machinery and intelligence. *Parsing the turing test* (pp. 23–65). Springer.

Ultes, S., ElChab, R., & Minker, W. (2014). Application and evaluation of a conditioned hidden markov model for estimating interaction quality of spoken dialogue systems. *Natural interaction with robots, knowbots and smartphones* (pp. 303–312). Springer.

Ultes, S., Miehle, J., & Minker, W. (2019). On the applicability of a user satisfaction-based reward for dialogue policy learning. *Advanced social interaction with agents* (pp. 211–217). Springer.

Ultes, S., Rojas-Barahona, L. M., Su, P.-H., Vandyke, D., Kim, D., Casanueva, I., Budzianowski, P., Mrkšić, N., Wen, T.-H., Gašić, M., & Young, S. (2017). PyDial: A multi-domain statistical dialogue system toolkit. *Proceedings of ACL 2017, System Demonstrations*, 73–78.

Ultes, S., Schmitt, A., & Minker, W. (2013). On quality ratings for spoken dialogue systems–experts vs. users. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 569–578.

Van Kleef, G. A., van den Berg, H., & Heerdink, M. W. (2015). The persuasive power of emotions: Effects of emotional expressions on attitude formation and change. *Journal of Applied Psychology*, *100*(4), 1124.

van Kleef, G. (2014). Emotions as Agents of Social Influence. *The Oxford Handbook of Social Influence*. Oxford University Press.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*, 5998–6008.

Venkatesh, A., Khatri, C., Ram, A., Guo, F., Gabriel, R., Nagar, A., Prasad, R., Cheng, M., Hedayatnia, B., Metallinou, A. et al. (2017). On evaluating and comparing conversational agents. *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)*.

Verheij, B. (2009). The toulmin argument model in artificial intelligence. *Argumentation in artificial intelligence* (pp. 219–238). Springer.

Vinyals, O., & Le, Q. (2015). A neural conversational model. *arXiv preprint arXiv:1506.05869*.

Wachsmuth, H., Naderi, N., Habernal, I., Hou, Y., Hirst, G., Gurevych, I., & Stein, B. (2017a). Argumentation quality assessment: Theory vs. practice. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 250–255.

Wachsmuth, H., Naderi, N., Hou, Y., Bilu, Y., Prabhakaran, V., Thijm, T. A., Hirst, G., & Stein, B. (2017b). Computational argumentation quality assessment in natural language. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 176–187.

Wachsmuth, H., Potthast, M., Al Khatib, K., Ajjour, Y., Puschmann, J., Qu, J., Dorsch, J., Morari, V., Bevendorff, J., & Stein, B. (2017c). Building an argument search engine for the web. *Proceedings of the 4th Workshop on Argument Mining (ArgMining 2017) at EMNLP*, 49–59.

Walker, M. A., Litman, D. J., Kamm, C. A., & Abella, A. (1997). PARADISE: A framework for evaluating spoken dialogue agents. *35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics*, 271–280.

Walton, D. N. (1996). *Argumentation schemes for presumptive reasoning* (1st ed.). Routledge.

Wang, Z., Bapst, V., Heess, N., Mnih, V., Munos, R., Kavukcuoglu, K., & de Freitas, N. (2016). Sample efficient actor-critic with experience replay. *arXiv preprint arXiv:1611.01224*.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine learning*, *8*(3-4), 279–292.

Weber, K., Janowski, K., Rach, N., Weitz, K., Minker, W., Ultes, S., & Andrè, E. (2020a). Predicting persuasive effectiveness for multimodal behavior adaptation using bipolar weighted argument graphs. *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems*, 1476–1484.

Weber, K., Rach, N., Minker, W., & André, E. (2020b). How to win arguments. *Datenbank-Spektrum*, *20*(2), 161–169. `https://doi.org/10.1007/s13222-020-00345-9`

Weber, K., Ritschel, H., Aslan, I., Lingenfelser, F., & André, E. (2018). How to Shape the Humor of a Robot - Social Behavior Adaptation Based on Reinforcement Learning. *Proceedings of the International Conference on Multimodal Interaction*, 154–162.

Weisbuch, M., Ambady, N., Clarke, A. L., Achor, S., & Weele, J. V.-V. (2010). On being consistent: The role of verbal–nonverbal consistency in first impressions. *Basic and Applied Social Psychology*, *32*(3), 261–268.

Wen, T.-H., Gašić, M., Kim, D., Mrkšić, N., Su, P.-H., Vandyke, D., & Young, S. (2015). Stochastic language generation in dialogue using recurrent neural networks with convolutional sentence reranking. *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 275–284.

Wen, T.-H., Vandyke, D., Mrkšić, N., Gašić, M., Rojas-Barahona, L. M., Su, P.-H., Ultes, S., & Young, S. (2017). A network-based end-to-end trainable task-oriented dialogue system. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 438–449.

Wilcock, G., & Jokinen, K. (2021). Towards increasing naturalness and flexibility in human-robot dialogue systems. *Increasing Naturalness and Flexibility in Spoken Dialogue Interaction: 10th International Workshop on Spoken Dialogue Systems*, 109–114.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K. et al. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.

Yang, X., Song, Q., & Wang, Y. (2007). A weighted support vector machine for data classification. *International Journal of Pattern Recognition and Artificial Intelligence*, *21*(05), 961–976.

Yang, Y. (2019). *Multi-agent actor-critic reinforcement learning for argumentative dialogue systems* (Master's thesis). Ulm University.

Yeasin, M., Bullot, B., & Sharma, R. (2006). Recognition of facial expressions and measurement of levels of interest from video. *IEEE Transactions on Multimedia*, *8*(3), 500–508.

Young, S. (2007). *CUED Standard Dialogue Acts* (tech. rep.). Cambridge University Engineering Dept.

Young, S., Gašić, M., Thomson, B., & Williams, J. D. (2013). Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, *101*(5), 1160–1179.

Yuan, T. (2004). *Human-computer debate, a computational dialectics approach* (Doctoral dissertation). Leeds Metropolitan University.

Yuan, T., Moore, D., & Grierson, A. (2008). A human-computer dialogue system for educational debate: A computational dialectics approach. *International Journal of Artificial Intelligence in Education*, *18*(1), 3–26.

Zhang, G. P. (2000). Neural networks for classification: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *30*(4), 451–462.

# Contributing Publications


Aicher, A., Rach, N., Minker, W., & Ultes, S. (2021). Opinion building based on the argumentative dialogue system bea. *Increasing Naturalness and Flexibility in Spoken Dialogue Interaction: 10th International Workshop on Spoken Dialogue Systems*, 307–318. `https://doi.org/10.1007/978-981-15-9323-9_27`

Rach, N., Langhammer, S., Minker, W., & Ultes, S. (2019a). Utilizing argument mining techniques for argumentative dialogue systems. *9th International Workshop on Spoken Dialogue System Technology*, 131–142. `https://doi.org/10.1007/978-981-13-9443-0_12`

Rach, N., Matsuda, Y., Daxenberger, J., Ultes, S., Yasumoto, K., & Minker, W. (2020a). Evaluation of argument search approaches in the context of argumentative dialogue systems. *Proceedings of the 12th Language Resources and Evaluation Conference*, 513–522.

Rach, N., Matsuda, Y., Ultes, S., Minker, W., & Yasumoto, K. (2021a). Estimating subjective argument quality aspects from social signals in argumentative dialogue systems. *IEEE Access*, *9*, 11610–11621. `https://doi.org/10.1109/ACCESS.2021.3051526`

Rach, N., Minker, W., & Ultes, S. (2017b). Towards an argumentative dialogue system. *Proceedings of the 17th Workshop on Computational Models of Natural Argument*, 27–29.

Rach, N., Minker, W., & Ultes, S. (2018a). Handling unknown user arguments in argumentative dialogue systems. *Proceedings of the 32nd International BCS Human Computer Interaction Conference*, 1–4. `https://doi.org/10.14236/ewic/HCI2018.67`

Rach, N., Minker, W., & Ultes, S. (2018b). Markov games for persuasive dialogue. *Computational Models of Argument: Proceedings of COMMA 2018*, *305*, 213–220. `https://doi.org/10.3233/978-1-61499-906-5-213`

Rach, N., Minker, W., & Ultes, S. (2020b). Increasing the naturalness of an argumentative dialogue system through argument chains. *Computational Models of Argument: Proceedings of COMMA 2020*, *326*, 331–338. `https://doi.org/10.3233/FAIA200516`

Rach, N., Schindler, C., Feustel, I., Daxenberger, J., Minker, W., & Ultes, S. (2021b). From argument search to argumentative dialogue: A topic-independent approach to argument acquisition for dialogue systems. *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 368–379.

Rach, N., Weber, K., Aicher, A., Lingenfelser, F., André, E., & Minker, W. (2019b). Emotion recognition based preference modelling in argumentative dialogue systems. *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 838–843. `https://doi.org/10.1109/PERCOMW.2019.8730650`

Rach, N., Weber, K., Pragst, L., André, E., Minker, W., & Ultes, S. (2018c). Eva: A multimodal argumentative dialogue system. *Proceedings of the 2018 on International Conference on Multimodal Interaction*, 551–552. `https://doi.org/10.1145/3242969.3266292`

Rach, N., Weber, K., Yang, Y., Ultes, S., André, E., & Minker, W. (2021c). Eva 2.0: Emotional and rational multimodal argumentation between virtual agents. *it-Information Technology*, *63*(1), 17–30. https://doi.org/10.1515/itit-2020-0050

Weber, K., Janowski, K., Rach, N., Weitz, K., Minker, W., Ultes, S., & Andrè, E. (2020a). Predicting persuasive effectiveness for multimodal behavior adaptation using bipolar weighted argument graphs. *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems*, 1476–1484.

Weber, K., Rach, N., Minker, W., & André, E. (2020b). How to win arguments. *Datenbank-Spektrum*, *20*(2), 161–169. https://doi.org/10.1007/s13222-020-00345-9

# Additional Publications


Abro, W. A., Aicher, A., Rach, N., Ultes, S., Minker, W., & Qi, G. (2021). Natural language understanding for argumentative dialogue systems in the opinion building domain. *arXiv preprint arXiv:2103.02691*.

Pragst, L., Rach, N., Minker, W., & Ultes, S. (2018). On the vector representation of utterances in dialogue context. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 776–780.

Rach, N., Minker, W., & Ultes, S. (2017a). Interaction quality estimation using long short-term memories. *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, 164–169. https://doi.org/10.18653/v1/W17-5520

Wagner, N., Kraus, M., Rach, N., & Minker, W. (2021). How to address humans: System barge-in in multi-user hri. *Increasing Naturalness and Flexibility in Spoken Dialogue Interaction: 10th International Workshop on Spoken Dialogue Systems*, 147–152. https://doi.org/10.1007/978-981-15-9323-9_13