Towards Integration of Cognitive Models in Dialogue Management: Designing the Virtual Negotiation Coach Application[*]

Andrei Malchanau[1], Volha Petukhova[1] and Harry Bunt[2]
[1]Spoken Language Systems Group, Saarland University, Germany
[2]Tilburg Center for Communication and Cognition, Tilburg University, The Netherlands
{andrei.malchanau,v.petukhova}@lsv.uni-saarland.de; harry.bunt@uvt.nl

October 6, 2019

Abstract

This paper presents an approach to flexible and adaptive dialogue management driven by cognitive modelling of human dialogue behaviour. Artificial intelligent agents, based on the ACT-R cognitive architecture, together with human actors are participating in a (meta)cognitive skills training within a negotiation scenario. The agent employs instance-based learning to decide about its own actions and to reflect on the behaviour of the opponent. We show that task-related actions can be handled by a cognitive agent who is a plausible dialogue partner. Separating task-related and dialogue control actions enables the application of sophisticated models along with a flexible architecture in which various alternative modelling methods can be combined. We evaluated the proposed approach with users assessing the relative contribution of various factors to the overall usability of a dialogue system. Subjective perception of effectiveness, efficiency and satisfaction were correlated with various objective performance metrics, e.g. number of (in)appropriate system responses, recovery strategies, and interaction pace. It was observed that the dialogue system usability is determined most by the quality of agreements reached in terms of estimated Pareto optimality, by the user's negotiation strategies selected, and by the quality of system recognition, interpretation and responses. We compared human-human and human-agent performance with respect to the number and quality of agreements reached, estimated cooperativeness level, and frequency of accepted negative outcomes. Evaluation experiments showed promising, consistently positive results throughout the range of the relevant scales.

1 Introduction

The increasing complexity of human-computer systems and interfaces results in an increasing demand for intelligent interaction that is natural to users and that exploits the full potential of spoken and multimodal communication. Much of the research in human-computer system design has been conducted in the area of task-oriented systems, especially for information-seeking dialogues concerning well-defined tasks in restricted domains – see Table 1 for the main paradigms used for dialogue modelling in domains of varying complexity.

Many existing systems represent a set of possible dialogue state transitions for a given dialogue task. Dialogue states are typically defined in terms of dialogue actions, e.g. question, reply, inform, and slot-filling goals. States in a finite state transition network are often used to represent the dialogue states (Bilange, 1991; Dahlbäck and Jönsson, 1998). Some flexibility has been achieved when applying statistical machine learning methods to dialogue state tracking (Williams et al., 2013). Statistical dialogue managers were initially based on Markov Decision Processes (Young, 2000) where given a number of observed dialogue events (often dialogue acts), the next event is predicted from the probability distribution of the events which have followed these observed events in the past. Partially Observable Markov Decision Processes (Williams and Young, 2007) model unknown user goals by an unknown probabilistic distribution over the user states. The POMDP approach is considered as the state-of-the-art in task-oriented spoken dialogue systems, see Young et al. (2013). However,

---

Technique	Example task	Dialogue phenomena handled
Finite state script	Long-distance calling	User answers questions
Frame based	Getting train arrival and departure information	User asks questions, simple clarifications by the system
Information State Update	Travel booking agent	Flexible shifts between pre-determined topics/tasks
		Refined grounding mechanisms
Plan based	Kitchen design consultant	Dynamically generated topic structures, e.g. negotiation dialogues
Agent based	Disaster relief management	Different modalities, e.g. planned world and actual world
		Collaborative planning and acting
Probabilistic approaches	Various information-seeking tasks, negotiation games	Dialogue policies design, i.e. learning combined with the most approaches mentioned above
Chat-oriented; interactive pattern matching /template-based	Retail 'chat commerce' Psychotherapies, personal assistant	Question-answering skills Social interactive aspects

Table 1: State-of-the-art techniques for task-oriented dialogue system.

when dealing with real users, the defined global optimisation function poses important computational difficulties. Recently, deep neural networks have gained a lot of attention (Henderson et al, 2013; 2014). Hierarchical recurrent neural networks have also been proposed to generate open domain dialogues and build end-to-end dialogue systems trained on large amounts of data without any detailed specification of information states (Serban et a., 2016). The real challenge for end-to-end frameworks is however the decision-taking problem related to the dialogue management for goal-oriented dialogues. Statistical and end-to-end approaches require really large amounts of data, while offering a rather limited set of dialogue actions (Kim et al., 2015). While such dialogue systems may perform well on simple information-transfer tasks, they are mostly unable to handle real-life communication in complex settings like, for example, multi-party conversations, tutoring sessions and debates. More conversationally plausible dialogue models are based on rich representations of dialogue context for flexible dialogue management, e.g. information-state updates (ISU, Traum et al., 1999; Bunt, 1999; Bos et al., 2003; Keizer et al., 2011). Other approaches to dialogue processing and management are built as full models of rational agency accounting for planning and plan recognition (Cohen and Perrault, 1979; Carberry, 1990; Sadek, 1991). Plan construction and inference are activities that can however easily get very complex and become computationally intractable. Alternatively, dialogue plans and strategies can be learned and adapted through reinforcement learning (Sutton and Barto, 1998). However, this seems to require even greater amounts of data, Henderson et al. (2008).

The research community is currently targeting more flexible, adaptable, open-domain multimodal dialogue systems. Advances are made in modelling and managing multi-party interactions, e.g. for meetings or multi-player games, where approaches developed for two-party dialogue have to be extended in order to model phenomena specific to multi-party interactions. Nevertheless, simple command/control and query/reply systems prevail. Some dialogue systems developed for research purposes allow for more natural conversations, but they are often limited to a narrow manually crafted domain and to rather restricted communication behaviour models, e.g. often modelled on information retrieval tasks. In some cases, these restrictions are imposed deliberately by the researchers to be able to investigate a limited set of dialogue phenomena without having to deal with unrelated details. However, this reduces the practical realism of the dialogue system.

Expectations of the users of today are rather high, requiring a real-time engagement with highly relevant personalized content that mimics human natural behaviour and is able to adapt to changing user needs and goals. Nowadays, there is a growing interest in Artificial Intelligence (AI)-powered conversational systems that are able to learn and reason, to facilitate realistic interactive scenarios with realistic assets and lifelike, believable characters and interactions. AI models may represent rather complex research objects. Despite their acknowledged potential, generating plausible AI models from scratch is challenging. For instance, cognitive models were successfully integrated into intelligent tutoring and intelligent narrative systems, see Paiva et al. (2004); Riedl and Stern (2006); Vanlehn (2006); Ritter et al. (2007); Lim et al. (2012). Since such models produce detailed simulations of human performance encompassing many domains such as learning, multitasking, decision making, and problem solving, they are also perfectly capable to play the role of a believable human-like agent in various human-agent settings. Although the abilities of cognitive agents continue to improve, human-agent interaction is often awkward and unnatural. The agents most of the time cannot deliver human-like interactive behaviour, but deal well with task actions thanks to the use of well-defined computational cognitive task models.

This paper presents an approach to the incorporation of cognitive task models into Information State Update (ISU) based dialogue management in multimodal dialogue systems. Such integration has important advantages. The ISU methodology has been applied successfully to a large variety of interactive tasks, e.g. information seek-

ing (Keizer et al., 2011), human-robot communication (Peltason and Wrede, 2011), instruction giving (Lauria et al., 2001), and controlling smart home environments (Bos et al, 2003). Several ISU development environments are available, such as TrindiKit (Larsson and Traum, 2000), Dipper (Bos et al., 2003) and FLoRes (Morbini et al., 2014). The ISU approach provides a flexible computational model for understanding and generation of dialogue contributions in term of effects on the information states of the dialogue participants. ISU models account for the creation of (shared) beliefs and mechanisms for their transfer, and have well-defined machinery for tracking, understanding and generation of natural human dialogue behaviour. Cognitive modelling of human intelligent behaviour, on the other hand, enables deep understanding of complex mental task processes related to human comprehension, prediction, learning and decision making. Threaded cognition (Salvucci and Taatgen, 2008) and Instance-Based Learning (Gonzalez and Lebiere, 2005) models developed within the ACT-R cognitive architecture (Anderson, 2007) are used to design a cognitive agent that can respond and adapt to new situations, in particular to a communicative partner changing task goals and strategies. The agent is equipped with Theory of Mind skills (Premack and Woodruff, 1978) and is able to use its task knowledge not only to determine its own actions, but also to interpret the human partner's actions, and to adjust its behaviour to whom it interacts with. In this way, we expect to achieve flexible adaptive dialogue system behaviour in dynamic non-sequential interactions. The integrated cognitive agent does not only compute the most plausible task action(-s) given its understanding of the partner's actions and strategies, provides alternatives and plans possible outcomes, but it also knows why it selects a certain action and can explain why its choices lead to the specific outcome. This enables the agent to act as a cognitive tutor, supporting the development of the (meta)cognitive skills of a human learner. Finally, the agent can be built with rather limited real or simulated dialogue data: it is supplied with initial state-action templates encoding domain knowledge and the agent's preferences, and the agent further learns from the collected interactive experiences.

The present study investigates the core properties of cognitive models that underlie human task-related and interactive dialogue behaviour, shows how such models provide a basis for dialogue management and can be integrated into a dialogue system, and assesses the resulting system usability. As the use and evaluation case, our simulated agents and human actors participate in (meta)cognitive skills training within negotiation based scenarios.

This paper is structured as follows. Section 2 discusses cognitive modelling, with a focus on human interactive multitasking, learning and adaptive behaviour. We briefly discuss the ACT-R architecture and provide details on an instance-based cognitive model that we used as a basis for designing an agent's decisions-making processes and generation of task-related actions. Section 3 describes an interactive learning scenario for the development of metacognitive skills in a multi-issue bargaining setting. We provide an overview of existing approaches and systems for cognitive tutoring tasks, as well as dialogue systems used in negotiation domains. We specify tasks and actions performed by negotiators, negotiation structures procedures and negotiation strategies. The data collection scenario is outlined and the semantic annotations of the data are discussed. Section 4 specifies a multi-agent dialogue manager architecture that makes use of a dynamic multidimensional context model and incorporates a cognitive task agent plus various interaction control agents trained on the annotated data. Section 5 presents the Virtual Negotiation Coach, outlining the system architecture and providing important details for key modules.[1] Section 6 reports on the system evaluation, where users' subjective perception of effectiveness, efficiency and satisfaction were correlated with various objective performance metrics. Evaluation results are also provided with respect to the number and quality of agreements reached, estimated level of cooperativeness, and acceptance of negative outcomes, as well as the subjective assessment of the skill training effects. Section 7 summarises our findings and outlines future research.

2 Cognitive modelling

Cognitive models have been used for decades to explain and model human intelligent behaviour, and have been successful in capturing a wide variety of phenomena across multiple domains such as decision making (Marewski and Link, 2014), memory (Nijboer et al., 2016), problem solving (Lee et al., 2015), task switching

---

[1]Note that details on the integrated system are provided to enable the evaluation of the designed dialogue manager. Since it is difficult to evaluate the dialogue manager as a separate module, its evaluation is performed as part of the user-based evaluation of the integrated dialogue system with negotiation, tutoring and interactive capabilities. The detailed description of the full dialogue system functionality is thus out of scope of this study.

(Altmann and Gray, 2008), user models in tutoring applications (Ritter et al., 2007), and neuroimaging data interpretation (Borst and Anderson, 2015).

One of the most widely researched cognitive architecture is ACT-R, see Anderson (2007), a theory and platform for building models of human cognition, which accounts for hundreds of empirical results obtained in the field of experimental psychology. ACT-R proposes a hybrid architecture that combines a production system to capture the sequential, symbolic structure of cognition, with a sub-symbolic, statistical layer to capture the adaptive nature of cognition.

Since available cognitive models produce detailed simulations of human (multi-)task performance, they are also of interest for playing a role in a multi-agent setting. This application is exploited in this study. It is of chief importance that our artificial agents exhibit plausible human behaviour, notably a human-like way of learning and interacting. This means that such an agent makes decisions and takes actions that humans might also make and take, but also that the agent is influenced by its experiences and builds representations of the people it interacts with. Thus, the agent should be able to (1) learn by collecting a variety of experiences, through instruction and feedback, and through monitoring and reasoning about its own behaviour and that of others; (2) adapt its interactive behaviour to a human dialogue partner's knowledge, intentions, preferences and competences; and (3) process and perform several actions related to the interactive tasks and the roles it should play, e.g. as a partner or as a tutor.

2.1   Models of human learning: Reinforcement and Instance-Based Learning

Human learning involves acquiring and modifying knowledge, skills, strategies, beliefs, attitudes, and behaviors. Learning may involve synthesizing different types of information (Schunk, 2012). Learning is a relatively permanent change in behavior as a result of experience (Gross, 2016). People learn from their successes and failures, from observing situations around them, and from imitating the behaviour of others (Bandura, 2012). Two widely used learning models are Reinforcement Learning (RL) and Instance-Based Learning (IBL).

Reinforcement learning is a formal model of action selection where the utility of different actions is learned by attending to the reward structure of the environment. Generally speaking, RL works in a trial-and-error fashion attempting various actions and recording the reward gained by those actions, see Sutton and Barto (1998). One of the limitations of RL as a model of human decision making becomes apparent in environments where goals change. This may happen, for example, due to changes in the environment or to newly obtained knowledge of the environment, e.g. you need to mail a letter, you searched online for the closest post office, but on your way to it you see a street mailbox, so you drop the letter in there. Initial goal changes may occur due to the understanding and evaluation of partner behaviour. This often happens in negotiations where a negotiator may revise his initial offers and make concessions dependent on the interpretation of partner behaviour concerning these goals. RL models make decisions based solely on the learned state-action utilities. Rewards are set a priori, are fixed and never revisited. If the goal changes, the utilities representing the reward structure from the initial goal become irrelevant at best, and subversive at worst (Veksler et al., 2012). Recently, serious efforts have been undertaken to solve this issue combining concurrent learning (co-learning) of the system policy training and the policy trained against simulated users. For instance, Georgila et al. (2014) and Xiao and Georgila (2018) showed that in negotiation setting agents using multi-agent RL techniques are able to adapt to the human users, also in situations which were not observed during training.

Humans, by contrast, employ their knowledge of the environment and their interactive partners to make decisions for achieving new goals, e.g. acting from experience or by association. Our memories are retrieved based on their recency and frequency of use (Anderson and Schooler, 1991) and strategies are adapted with increasing task experience (Siegler and Stern, 1998).

Human learning often occurs as a result of experience. Decisions are made by finding a prior experience (an instance) that is similar to the current situation and/or most recently, frequently used under comparable conditions, see Logan (1988); Gonzalez and Lebiere (2005). An instance consists of a representation of the current state of the world (what do I know, what do I know about others, what am I asked, what can I do, what has happened before), and an action to be taken in that situation (give information, run tests, examine something, reason about others, change attitude, etc.). Information is encoded in an instance as a state-actions template specifying decisions about which activity to engage in and how to move from one activity to the other. Initial templates can be designed (pre-programmed) by experts and/or modelled as the result of dialogue corpus analysis. An IBL agent can start an interaction with an (almost) empty template, request information from

the partner and add it to the memory as the interaction proceeds. Newly created (partially) filled instances are stored in a human-like memory that models forgetting, similarity and blending of experiences. The most active instance is retrieved. Activation is based on history (e.g. frequency and recency) and on similarity (e.g. how similar the instance is, given the context), see Section 4.2.1 for the specification of instances and activation functions for our interactive settings. An agent can be also trained by giving it a set of instances (learning-by-instruction), which it can refine and/or augment in actual interaction (learning-by-doing and learning-by-feedback).

RL is a useful paradigm where the possible strategies are relatively clear. If the underlying interaction structure is very flexible, unclear or absent (i.e. hard to derive on the basis of the system's behaviour), IBL based models have advantages, see also Arslan et al. (2017). For instance, whenever a new goal is given, the IBL model will employ its stored knowledge (instances) to make informed goal-directed decisions. It does not need to learn the reward structure through trial-and-error; rather, the decision what action will be performed is based on the computed activation level, e.g. similarity between a past experience and the given current goal. Moreover, feedback can be used in IBL to create an instance that contains the correct solution, i.e. the model will add an instance of another strategy, whereas the RL model will punish the strategies that lead to a wrong solution. Strategy selection, which is implicit in RL, is explicit in the IBL model which makes it particularly suitable for tutoring applications. IBL is moreover robust to missing information due to the partial matching component in the ACT-R activation function, e.g. when the agent does not have access to the same information as his partner. We applied the instance-based learning approach to create flexible cognitive agents, also because it requires far less experience than machine learning methods that learn bottom-up, and the agent's decision-taking behaviour incrementally improves as its set of instances increases in size. Instance-based learning takes the middle ground between expert systems, in which knowledge typically lacks flexibility, and bottom-up machine learning, which requires extensive training data, and in which decisions are reached in an opaque manner.

2.2 Adaptive interactive behaviour

Interactive systems and interfaces tailored towards specific users have been demonstrated to outperform traditional systems in usability. Nass et al. (2005) present an in-car user study with a "virtual passenger". Experimental results indicate that subjective and objective criteria, such as driving quality, improve when the system adapts its voice characteristics to the driver's emotion. Nass and Li (2000) confirm in the study of spoken dialogues in a book shop that similarity attraction is important for personality expression: matching the users' degree of extroversion strongly influenced trust and attributed intelligence.

These observations have triggered the development of interactive systems that model and react to the users' traits and states, for example by adapting the interaction based on language generation techniques (Mairesse and Walker, 2005). In Gnjatovic and Rösner (2008) a gaming interface is based on emotional states computed from the interaction history and actual user command. Nasoz and Lisetti (2007) describe a user modelling approach for an intelligent driving assistant, which derives the best system action in terms of driving safety, given estimated driver states.

The above approaches adapt locally, i.e. the adaptation decision is made at turn level with very limited context and thus with no or very limited foresight. Reinforcement Learning has emerged as a promising approach for long-term considerations. While early studies (Walker et al., 1998; Singh et al., 2002) used RL to build strategies for simple systems, more complex paradigms are represented by statistical models, see Frampton and Lemon (2009). However, when users with different personalities in different states are systematically confronted with a learning system, most studies resort to user simulation: Janarthanam and Lemon (2009) simulate users of different levels of expertise, López-Cózar et al. (2009) simulate users with different levels of cooperativeness, and Georgila et al. (2010) simulate interactions of old and young users.

These studies demonstrate that the simulation of different user types is expected to lead to strategies which adapt to each user type. However, adaptivity has been not achieved at the level of dynamically changing goals within one dialogue. Rewards that are used in dialogue policy learning and optimizations are fixed a priori. Human learning however does not only involve strengthening of existing knowledge, compilation of new rules, collection of episodic experiences to improve future decisions, etc., but often requires more explicit reasoning, assessing why a particular solution worked or not, and manipulating the task representation accordingly - this process is called 'metacognition'. In this study, metacognition plays two major roles: (1) it guides and regulates

system task behaviour; and (2) it improves a participant's learning by triggering reasoning about one's own and partner behaviour.

Metacognitive processes concern reasoning about other people's intentions and knowledge. Mastering metacognitive skills is important in language use (Van Rij et al., 2010) and in playing knowledge games (Meijering et al., 2012). A more elaborate form of these reasoning skills is important in collaboration, negotiation and other social and interpersonal skills. People with well-developed metacognitive skills are more concerned that their interactions will go well, and are able to flexibly modify their actions during interaction in order to better adapt to the dynamics of the situation, typically by using other people's behaviour as a guide to their own (Ickes et al., 2006). They are also better able to accomplish their goals, which appears to be the result of their superior planning skills (Jordan and Roloff, 1997).

Metacognitive skills can be trained by humans and learned by a system. When learning, humans also observe their partners' behaviour. In addition to using experiences to determine its own decisions, an interactive agent can use them to interpret and reason about the behaviour of others (i.e. humans). The ability to understand that other people have mental states, with desires, beliefs and intentions, which can be different from one's own, is called Theory of Mind (ToM; Premack and Woodruff, 1978). In our application, the ToM methodology has been used to design agents that can infer, explain, predict and correct a partners' negotiation behaviour and negotiation strategies.

2.3 Multitasking in human-computer interaction

A dialogue system has at least three core tasks: (1) to monitor user dialogue behaviour; (2) to understand user dialogue contributions; and (3) to react adequately. Participation in a dialogue is thus a complex activity. Participants do not only need to exchange certain information, instruct another participant, negotiate an agreement, discuss results or plan future actions, etc., but among other things dialogue participants also share information about the processing of each others messages, elicit feedback, manage the use of time, take turns, and monitor contact (Allwood, 2000). They often use linguistic and nonverbal elements to address several interactive and task-related aspects at the same time.

During interaction, a dialogue system is usually in the role of "speaker" (or "sender')' or in the role of "addressee" (also called "hearer" or "recipient"). The system may also play the role of a side-participant who witnesses a dialogue without participating in it, see Clark (1996).

A dialogue system's tasks depend also on the application domain in relation to the role(-s) it plays, e.g. as a full-fledged interactive partner with equal responsibilities as a human one, as an assistant, adviser or mediator, as a passive observer, as a tutor or coach, and so on. For our Virtual Negotiation Coach application we identified the following key roles:

- *Observer*: system observes dialogue sessions between two or more humans and keeps track of human-human dialogue without actively participating in it;

- *Experiencer*: system actively plays the role of one of the interaction participants, i.e. sender and addressee;

- *Mirror*: system re-plays the user's performance in a human-system dialogue in real time. The user observes his own performance and has the opportunity to terminate, re-enter and re-play the dialogue session from any point;

- *Tutor* or *Coach*: system provides feedback from ongoing formative or summative assessment of the user performance in one or more tutoring sessions (Mory, 2004).

The system may play multiple roles simultaneously and/or interchangeably.

In most existing approaches to dialogue management the Dialogue Manager (DM) is able to handle one particular dialogue task at a time. Most human activities however are essentially multitasking. For example, driving a car consists of two main processes: one that keeps the car in the middle of the driveway by looking at the road ahead of the car while operating the steering wheel and the gas and brake pedals, and a second process that monitors the traffic environment (e.g., is there a car behind you). Thus, human cognition can be conceptualized as a set of parallel cognitive modules (e.g. vision, declarative memory, working memory, procedural memory, manual control, vocal control, etc.). As long as multiple tasks do not need the same resources at the same time, these tasks can be carried out in parallel without interference. In the case of the

driving example, if the driver is given an additional task, for example to operate a cell phone, he will abandon the monitoring task due to lack of resources.

Threaded cognition, as the theory of parallel execution of tasks, has been proposed to explain human multi-tasking behaviour: why and when certain tasks may be performed together with ease, and which combinations pose a difficulty, what types of multitasking are disruptive, and when are they most disruptive. Threaded cognition models have been used in a wide spectrum of multitasking experiments (Salvucci and Taatgen 2008; 2010). This theory has been built on top of the ACT-R cognitive architecture. We designed a multi-threaded Dialogue Manager with integrated multitasking cognitive agent which, along with being an active dialogue participant with monitoring, understanding and reacting tasks, is capable of providing feedback on partner performance and which can reason about its own and a partner's behaviour, and suggest alternative actions.

3 Interactive training of metacognitive skills

3.1 Interactive learning and tutoring

Cognitive Tutoring Systems aim to support the development of metacognitive skills. Examples of such systems are described in Bunt and Conati (2003); Azevedo et al. (2002); Gama (2004); Aleven et al. (2006) and Baker et al. (2006). These systems rely on artificial intelligence and cognitive science as a theoretical basis for analysing how people learn (Roll. et a., 2007).

Research by Chi et al. (2001) revealed that the interactivity of human tutoring drives its effectiveness. *Interactive learning* is a modern pedagogical approach that has devolved out of the hyper-growth in the use of digital technology and virtual communication. Interactive learning is a promising and powerful way to develop metacognitive skills. In this study, the interactivity of a tutoring system is achieved through the use of multimodal dialogue. While many intelligent tutoring dialogue systems have been developed in the past (Litman and Silliman, 2004; Riedl and Stern, 2006; Core et al., 2014; Moore et al., 2005; Paiva et al., 2004), to the best of our knowledge no existing cognitive tutoring system makes use of natural spoken and multimodal dialogue.

Metacognitive skills are domain-independent and should be applicable in any learning domain and in a variety of different learning environments, but despite their transversal nature, metacognitive skills training can only be practiced within certain domains and activity types. Some systems have been developed successfully for the domains of mathematics, physics, geometry, biology and computer programming (MetaTutor, Azevedo et al., 2009; Rus et al., 2009: Harley et al., 2013). For negotiation, metacognition has been empirically proven to be important since it significantly improves decision-making processes (Aquilar and Galluccio, 2007).

For many existing human-computer negotiation systems, interactions are typically modelled as a sequence of competitive offers where partners claim a bigger share for themselves. Valuable work has been done on well-structured negotiations where a few parties interact with fixed interests and alternatives, see e.g. Traum et al. (2008), Georgila and Traum (2011), Guhe and Lascarides (2014), Efstathiou and Lemon (2015). In many real-life negotiations, parties negotiate not over one but over multiple issues, see e.g. Cadilhac et al. (2013), where they have interests in reaching agreements about several issues, and their preferences concerning these issues are not completely identical (Raiffa et al., 2002a). Negotiators may have partially competitive and partially cooperative goals, and may make trade-offs across issues in order for both sides to be satisfied with the outcome. Parties can delay making a complete agreement on the first discussed issue, e.g. they postpone making an agreement or make a partial agreement, until an agreement is reached on the second one. They can revise their past offers, accept or decline any standing offer, make counter-offers, etc. We consider such complex strategic negotiations as multi-issue integrative bargaining dialogues, see Petukhova et al. (2016 and 2017). We aim at modelling these interactions with the main goal to train metacognitive skills. Comparable work has been performed on modelling so-called semi-cooperative multi-issue bargaining dialogues, see (Lewis et al., 2017), who proposed an approach to end-to-end training of negotiation agents using a dataset of human-human negotiation dialogues, and applying reinforcement learning. Their study presents a new form of planning ahead where possible complete dialogue continuations are simulated - *dialogue rollout*. Our approach also allows to compute the best alternative move at each negotiation stage and plan ahead the complete negotiation. We compute about 420 outcomes per scenario, for 9 scenarios in total, each featuring different participant preference profiles. Additionally, for tutoring purposes the model provides an explanation for all alternative choices and how they lead to what outcomes. The two approaches differ with respect to

the amount of data/resources used (our 50 vs 5808 dialogues); scenario complexity (4 issues, 16 values and 9 different preference profiles in our scenario vs 3 types of items and 6 objects in Lewis et al., 2017); and modalities modelled (multimodal vs typed conversations). In our study, we explicitly model various negotiation strategies, while in Lewis et al. (2017), evidence of such strategies is observed, e.g. compromising or deceiving, and are implicitly learned but not considered by design.

3.2 Models of multi-issue bargaining

Three main types of negotiations can be distinguished: *distributive*, *joint problem-solving* and *integrative*[2]. Distributive negotiation means that any gain of one party is made at the expense of the other and vice versa; any agreement divides a fixed pie of value between the parties, see e.g. Walton and McKersie (1965). The goal of joint problem-solving negotiations is, by contrast, to work together on an equitable and reasonable solution: negotiators will listen more and discuss the situation longer before exploring options and finally proposing solutions. The relationship is important for joint problem solving, mostly in that it helps trust and working together on a solution, see Beach and Connolly (2005). In integrative bargaining, parties bargain over several goods and attributes, search for an integrative potential (interest-based bargaining or win-win bargaining, see Fisher an Ury, 1981). This increases the opportunities for cooperative strategies that rely on maximizing the total value of the negotiated agreement (enlarging the pie) in addition to maximizing one's own value at the expense of the partner (dividing the pie).
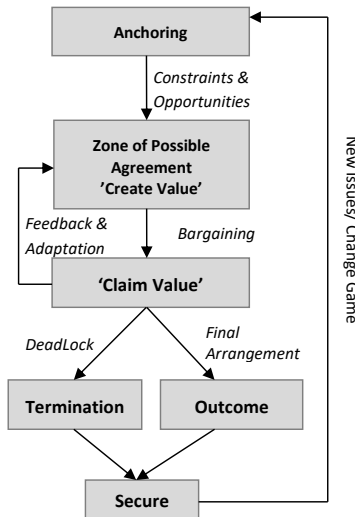

Figure 1: Negotiation phases associated with negotiation structure, based on Watkins (2003); Sebenius (2007).

The different types of negotiation are manifest mainly in how parties create and claim values. Negotiation starts with the **Anchoring** phase, in which participants introduce negotiation issues and options. They also obtain and provide information about preferences, establishing jointly possible values contributing to the **Zone of Possible Agreement** (ZOPA, Sebenius, 2007). Participants may bring up early (tentative) offers, typically in the form of suggestions, and refer to the least desirable events - 'Create Value'. The actual bargaining occurs in the **'Claim Value'** phase, potentially leading to adaptation, adjustment or cancelling the originally established ZOPA actions. Patterns of concessions, threats, warnings, and early tentative commitments are observed here. Distributive negotiations are more 'claiming values', while joint problem-solving negotiations are more 'value creating' interactions, and integrative negotiations are a mix of 'creating and claiming values' negotiations (Watkins, 2003a). In distributive negotiations the size of the ZOPA is mostly determined by the 'bottom lines' of the opposite parties, which are formed by their respective *best alternatives to a negotiated agreement* (BATNA), see Fisher and Ury (1981). In integrative bargaining the ZOPA is mainly determined by the number of possible Pareto optimal outcomes. Pareto optimality reflects a state of affairs when there is no alternative state that would make any partner better off without making anyone worse off.

After establishing the ZOPA, negotiators may still cancel previously made agreements and negotiations might be terminated. **Negotiation Outcome** is the phase associated with the "walk-away" positions for each partner. Finally, negotiators can move to the **Secure** phase summing up and restating negotiated agreements or termination outcomes. At this stage, strong commitments are expressed and weak beliefs concerning previously made commitments and agreements are strengthened. Participants take decisions to move on with another issue or re-start the discussion. Figure 1 depicts the general negotiation structure as described in Watkins (2003) and Sebenius (2007), and observed in our data described in the next section.

The negotiation outcome depends on the setting, but also on the agenda and the strategy used by each partner (Tinsley et al., 2002). The most common strategy of novice negotiators observed is issue-by-issue

---

[2]A fourth type of negotiations is *bad faith*, where parties only pretend to negotiate, but actually have no intention to compromise. Such negotiations often take place in a political context, see Cox (1958)

bargaining (see data collection below). Parties may start with what they think are the 'toughest' issues, where they expect the most sharply conflicting preferences and goals, or they may start to discuss the 'easiest', most compatible options. Sometimes, however, negotiators bring all their preferences on the table from the very beginning. This increases the chance to reach a Pareto efficient outcome, since a participant can explore the negotiation space more effectively, being able to reason about each others' goals, see e.g. Stevens et al. (2016b). Defensive behaviour, i.e. not revealing preferences, but also being misleading or deceptive, i.e. not revealing true preferences, results in missed opportunities for value creation, see e.g. Watkins (2003); Lax and Sebenius (1992). It has also been observed that as a rule it is easier for a negotiator to bargain down, i.e. to start with his highest preference and if this is not accepted by the partner, go down and discuss sub-optimal options, than it is to bargain in, i.e. to reveal his minimum goal and go up, offering preferences that are not necessarily shared by the partner.

All the aspects mentioned above may influence negotiators' strategies. Traum et al. (2008), who also consider a multi-issue bargaining setting, but viewed as a multi-party problem-solving task, define strategies as objectives rather than the orientations that lead to them. They distinguish seven different strategies: find issue, avoid, attack, negotiate, advocate, success and failure. Other researchers define negotiation strategies closely related to conflict management styles, i.e. the overall approach for conducting a negotiation. Five main strategies are observed: competing (adversarial), collaborating, compromising, avoiding (passive aggressive), and accommodating (submissive), see Raiffa et al. (2002a); Tinsley et al. (2002). As in integrative negotiation, where the negotiators strive to achieve a delicate balance between cooperation and competition (Lax and Sebenius, 1992), we define two basic negotiation strategies: cooperative and non-cooperative.

Cooperative negotiators share information about their preferences with their opponents, are engaged in problem-solving behaviours and attempt to find mutually beneficial agreements (De Dreu et al., 2000). A cooperative negotiator prefers the options that have the highest collective value. If not enough information is available to make this determination, a cooperative negotiator will elicit this information from his opponent concerning. A cooperative negotiator will not engage in positional bargaining[3] tactics, instead, he will attempt to find issues where a trade-off is possible.

Non-cooperative negotiators prefer to withhold their preferences in fear of weakening their power by sharing too much, or they may not reveal true preferences deceiving and misleading the partner. These negotiators focus on asserting their own preferred positions rather than exploring the space of possible agreements (Fisher and Ury, 1981). A negotiator agent using this strategy will rarely ask an opponent for preferences, and will often ignore a partner's interests and requests for information. Instead, a non-cooperative negotiator will find his own ideal offer, state it, and insist upon it in the hope of making the opponent concede. He will threaten to end the negotiation or will make very small concessions. The non-cooperative negotiator will accept an offer only if he can gain a lot from it.

We also model a *neutral* (or cautious) strategy. Neutral actions describe behaviours that are not indicative of either strategy above.

To sum up, our approach is based on the cognitive negotiation model of integrative multi-issue bargaining, which incorporates potentially different beliefs and preferences of negotiation partners, learns to reason about these beliefs and preferences, and accounts for changes in participants' goals and strategies.

3.3 Collection and annotation of negotiation data

For adequate modelling of human-like multi-issue bargaining behaviour, a systematic analysis of collected and semantically annotated human-human dialogue data was performed. The collected and analysed data also served for the IBL instance template definition as well as for the training agent's negotiation behaviour, e.g. various classifiers were built using this data, see Section 5. The specific setting considered in this study involved a real-life scenario about anti-smoking legislation in the city of Athens passed in 2015-2016. After a new law was enacted, many cases of civil disobedience were reported. Different stakeholders came together to (re-)negotiate and improve the legislation. The main negotiation partner was the Department of Public Affairs of the City Council who negotiated with representatives of small businesses, police, insurances, and others.

The anti-smoking regulations were concerned with four main *issues*: (1) smoke-free public areas (scope); (2) tobacco tax increase (taxation); (3) anti-smoking program promotion (campaign); and (4) enforcement

---

[3]Positional bargaining involves holding on to a fixed set of preferences regardless of the interests of others.

SCOPE
- ○ All outdoor smoking allowed
- ○ No smoking in public transportation
- ○ No smoking in public transportation and parks
- ○ No smoking in public transportation, parks and open air events

TAXATION
- ○ No change in tobacco taxes
- ○ 5% increase in tobacco taxes
- ○ 10% increase in tobacco taxes
- ○ 15% increase in tobacco taxes
- ○ 25% increase in tobacco taxes

CAMPAIGN
- ○ Flyer and billboard campaign in shopping district
- ○ Anti-smoking posters at all tobacco sales points
- ○ Anti-smoking television advertisements
- ○ Anti-smoking advertisements across all traditional mass media

ENFORCEMENT
- ○ Police fines for minors in possession of tobacco products
- ○ Ban on tobacco vending machines
- ○ Police fines for selling tobacco products to minors
- ○ Identification required for all tobacco purchases
- ○ Government issued tobacco card for tobacco purchases

Figure 2: Preference card: example of values in four negotiated issues presented in colours: brighter orange colours indicated increasingly negative options and brighter blue colours increasingly positive options. When incorporated into the graphical interface, partners' offers were visualized with red arrow (system) and green one (user).

Dialogue Act		Relative frequency (in %)	Dialogue Act		Relative frequency (in %)
Communicative function	Modality/ Qualifier		Communicative function	Modality/ Qualifier	
propositionalQuestion		2.0	suggest		10.0
checkQuestion		2.2	addressSuggest		1.4
setQuestion		10.3	acceptSuggest		2.0
choiceQuestion		0.6	declineSuggest		1.7
inform $->$		30.3	offer $->$		16.7
...	non-modalized	41.3	...	conditional	28.3
...	prefer	30.4	...	tentative	35.0
...	disprefer	3.1	...	final	36.7
...	acquiesce	3.0	addressOffer		0.6
...	need	2.0	acceptOffer $->$		5.8
...	able	19.0	...	tentative	47.6
...	unable	1.2	...	final	52.4
agreement		10.3	declineOffer	tentative	2.0
disagreement		4.1			

Table 2: Distribution of task-related dialogue acts in the analysed multi-issue bargaining dialogues.

policy and police involvement (enforcement), see Figure 2. Each of these issues involves four to five most important negotiation *values* with preferences representing negotiation positions, i.e. preference profiles. Nine cases with different preference profiles were designed. The strength of preferences was communicated to the negotiators through colours. Brighter orange colours indicated increasingly negative options; brighter blue colours increasingly positive options.

In the data collection experiments, each participant received the background story and a preference profile. Their task was to negotiate an agreement which assigns exactly one value to each issue, exchanging and eliciting offers concerning ⟨*ISSUE*; *VALUE*⟩ options. Participants were randomly assigned their roles. They were not allowed to show their preference cards to each other. No further rules on the negotiation process, order of discussion of issues, or time constraints were imposed. They were allowed to withdraw or re-negotiate previously made agreements within a session, or terminate a negotiation.

16 subjects (young professionals aged between 19 and 25 years) participated in the experiments. The resulting data collection consists of 50 dialogues of a total duration of about 8 hours, comprising approximately 4.000 speaking turns (about 22.000 tokens).

The recorded speech was transcribed, segmented and annotated with ISO 24617-2 dialogue act information. The ISO 24617-2 taxonomy (ISO, 2012; see also Bunt et al., 2010) distinguishes 9 dimensions, addressing information about a certain *Task*; the processing of utterances by the speaker (*Auto-feedback*) or by the addressee (*Allo-feedback*); the management of difficulties in the speaker's contributions (*Own-Communication Management*) or that of the addressee (*Partner Communication Management*); the speaker's need for time to continue the dialogue (*Time Management*); the allocation of the speaker role (*Turn Management*); the structuring of the dialogue (*Dialogue Structuring*); and the management of social obligations (*Social Obligations Management*).

Negotiation Move	Relative frequency (in %)
Offer	75.0
CounterOffer	12.4
Exchange	6.6
Concession	1.2
BargainIn	0.4
BargainDown	1.2
Deal	2.4
Terminate	0.8

Table 3: Negotiation moves and their relative frequencies in the annotated multi-issue bargaining corpus.

Additionally, to capture the negotiation task structure, *Task Management* acts are introduced. These dialogue acts explicitly address the negotiation process and procedure. This includes utterances for coordinating the negotiators' activities (e.g., "Let's go issue by issue") or asking about the status of the process (e.g., "Are we done with the agenda?"). Task Management acts are specific for a particular task and are often similar in form but different in meaning from Discourse Structuring acts, which address the management of the interaction, e.g. "To sum up ...", "Let's move to a next round".

At the negotiation task level, human-computer negotiation dialogue is often modelled as a sequence of offers. The offers represent participants' commitments to a certain negotiation outcome. In human negotiation, however, offers as binding commitments are rare and a larger variety of negotiation actions is observed, see Raiffa et al. (2002b). Participant actions are focused mainly on obtaining and providing preference information. A negotiator often states his preferences without expressing (strong) commitments to accept an offer that includes a positively evaluated option, or to reject an offer that includes a negatively evaluated option. To capture these variations, we distinguished five levels of commitment using the ISO 24617-2 dialogue act taxonomy[4] and its superset DIT^{++}[5]: (1) zero commitment for offer elicitations and preference information requests, e.g. by questions; (2) the lowest non-zero level of commitment for informing about preferences, abilities and necessities, e.g. in the form of modalized answers and informs; (3) an interest and consideration to offer a certain value, i.e. suggestions; (4) weak (tentative) or conditional commitment to offer a certain value; and (5) strong (final) commitment to offer a certain value, see Petukhova et al., 2017.

To model negotiation behaviour with respect to preferences, abilities, necessity and acquiescence, and to compute negotiation strategies as accurately as possible, we define several *modal relations* between the modality 'holder' (typically the speaker of the utterance) and the target which consists of the negotiation move (and its arguments), see Lapina and Petukhova (2017). Additionally, to facilitate structuring the interaction and enable participants to interpret partner intentions, dynamically changing goals and strategies efficiently, we defined a set of *qualifiers* attached to offer acceptances or rejections and agreements, tentative or final.

Semantically, dialogue acts correspond to update operations on the information states of the dialogue participants. They have two main components: (1) the *communicative function*, that specifies how to update an information state, e.g. Inform, Question, and Request, and (2) the *semantic content*, i.e. the objects, events, situations, relations, properties, etc. involved in the update, see Bunt (2000), Bunt (2014a). Negotiations are commonly analysed in terms of certain actions, such as offers, counter-offers, and concessions, see Watkins (2003), Hindriks et al. (2007). We considered two possible ways of using such actions, also referred to as 'negotiation moves', to compute the update semantics in negotiation dialogues. One is to treat negotiation moves as task-specific dialogue acts. Due to its domain-independent character, the ISO 24617-2 standard does not define any communicative functions that are specific for a particular kind of task or domain, but the standard invites the addition of such functions, and includes guidelines for how to do so. For example, a negotiation-specific kind of *Offer$_N$* function could be introduced for the expression of commitments concerning a negotiation value.[6] Another possibility is to use negotiation moves as the semantic content of general-purpose dialogue acts. For example, a negotiator's statements concerning his preference for a certain option can be represented as $Inform(A,B,\diamond offer(X;Y))$. We chose the latter possibility and specified 8 basic negotiation moves, whose distribution in the analysed data is shown in Table 3.

To sum up, the designed negotiation dialogue model accounts for several types of action performed by

---

[4]For more information see Bunt (2009); visit also `http://dit.uvt.nl/\#iso_24617-2`

[5]`http://dit.uvt.nl/`

[6]Negotiation 'Offers' may have a more domain-specific name, e.g. *Bid* for selling-buying bargaining.

ISO 24617-2 dimension	Relative frequency (in %)
Task	47.6
Task Management	10.3
AutoFeedback	18.7
AlloFeedback	2.3
Turn Management	6.6
Time Management	6.6
Discourse Structuring	4.6
Own Communication Management	2.1
Partner Communication Management	na
Social Obligation Management	1.2

Table 4: Distribution of dialogue acts per ISO 24617-2 dimension in the multi-issue bargaining corpus.

negotiators: (1) task-related dialogue acts expressing negotiation preferences and commitments; (2) qualified ('modalized') actions expressing participants' negotiation strategies, see Table 2; (3) negotiation moves specifying events and their arguments, see Table 3; and (4) communicative actions to control the interaction, see Table 4. A detailed specification of negotiation update semantics can be found in Petukhova et al. (2017).

Semantic annotations were performed by three trained annotators who reached a good inter-annotator agreement in terms of Cohen's kappa of 0.71 on average, when performing segmentation and annotation simultaneously. In total, the corpus data contains more than 18.000 annotated entities. Annotations were delivered in ISO DiAML format (ISO 24617-2, 2012), `.diaml` files consisting of primary data in TEI-compliant representation, with 24617-2 dialogue act annotations. The collected data and annotations is part of the Metalogue Multi-Issue Bargaining (MIB) corpus (Petukhova et al., 2016) which is released through LDC.[7].

4 Multi-Agent Dialogue Manager: functional design and technical integration

As ACT-R based computational cognitive models of threaded cognition and IBL can be used to design cognitive agents that simulate task-related behaviour showing close to human decision-making performance. If such agents have Theory of Mind (ToM) skills they can exhibit metacognitive capabilities that are beneficial for better understanding and adequate modelling of adaptive and proactive task behaviour. They cannot yet deliver natural human-like interactive performance, but combining them with interactive agents based on advanced computational dialogue models opens new possibilities. Inspired by the distinction that can be made between *task control* actions and *dialogue control* actions (Bunt, 1994), we explored these possibilities by integrating a cognitive task agent into the ISU-based dialogue manager as part of a dialogue system.

In the dialogue system design community, involving both theorists and practitioners, a clean separation into two layers is observed. One layer deals with the task at hand, and the other with the communicative performance itself, see e.g. Lemon et al. (2003). To design task managers (agents), detailed *task analysis*, originally proposed by Annett et al. (1971), is often performed. The method, in which a task is described in terms of a hierarchy of operations and plans, has been used successfully to simulate human decision-making processes. In dialogue management, it has also been deployed in the form of *hierarchical task decomposition* and *expectation agenda generation* within the RavenClaw framework (Bohus and Rudnicky, 2003) and tested successfully in several systems. Examples include the use of a tree-of-handlers in the Agenda Communicator (Xu and Rudnicky, 2000), of activity trees in WITAS (Lemon et al., 2001), and of recipes in Collagen (Rich et al., 1998). However, models based on task hierarchies, agendas, recipes and trees are rather static and are difficult to apply for non-linear (multi-branching) or non-sequential interactions, like multi-issue barganing dialogues.

A more flexible approach is the *plan-based* approach. For instance, in the TRIPS system (Allen et al., 2001) a Task Manager is implemented that relies on planning and plan recognition, and coordinates actions with a Conversational Manager. Plan construction and inference are activities that can easily get very complex, however, and become computationally intractable.

Multi-agent architectures have been proposed for adaptive and flexible human-computer interaction, e.g. in the JASPIS speech application (Turunen et al., 2005), in the Open Agent Architecture (Martin et al., 1999), and

---

[7]Please visit `https://catalog.ldc.upenn.edu/LDC2017S11`

in Galaxy-II (Seneff et al., 1998).

An ISU-based approach to dialogue management has been used to handle multiple aspects ('dimensions') simultaneously, see Keizer et al. (2011); Petukhova (2011); Malchanau et al. (2015), separating task control acts and various classes of dialogue control acts. The dialogue manager tracks updates in multiple dimensions of the participants' information states, as the effect of processing incoming dialogue acts, and generates multiple task control acts and dialogue control acts in response.

In order to capture the dynamics related to frequently changing participants' interactive and strategic goals, we propose a flexible *adaptive* form of multidimensional dialogue management inspired by cognitive models of multitasking, learning and cognitive skills transfer. To this end, we designed a Cognitive Task Agent and integrated it as part of an ISU-based multidimensional Dialogue Manager (DM). The DM receives data in the form of the recognized dialogue acts, updates the information state, and generates output.

4.1 Information State: the multidimensional context model

According to the ISU approach, dialogue behaviour, when understood by a dialogue participant, evokes certain changes in the participants' information state or 'context model'. Since we deal with several different interactive, task-related and tutoring aspects, an articulate context model should contain all the information considered relevant for interpreting such rich dialogue behaviour in order to enable the system to generate an adequate reaction playing the role of a Negotiator or that of a Tutor. An articulate dialogue model and context model have been proposed by Bunt (1999). Complexities of natural human dialogue are handled by analysing dialogue behaviour as having communicative functions in several dimensions, as discussed above.

$$
FunctionalSegment(FS) : \begin{bmatrix}
\begin{bmatrix}
start : \langle token_{index} | time\_point \rangle \\
end : \langle token_{index} | time\_point \rangle \\
verbatim : \langle token_{index} | time\_points = `token1', \ldots \rangle \\
prosody : \langle duration, pitch, energy, \ldots \rangle \\
nonverbal : \langle \begin{bmatrix} head\langle element_{index} | time\_points = expression1, \ldots \rangle \\ hands\langle element_{index} | time\_points = expression1, \ldots \rangle \\ face\langle element_{index} | time\_points = expression1, \ldots \rangle \\ posture\langle element_{index} | time\_points = expression1, \ldots \rangle \end{bmatrix} \rangle \\
sender : \langle participant \rangle \\
dial\_acts(DAs) : \{ \langle \begin{bmatrix} dimension(D) : \langle dim \rangle \\ comm\_function(CF) : \langle cf \rangle \\ sem\_content(SC) : \langle content \rangle \\ sender/speaker : \langle participant \rangle \\ addressee(-s) : \{ \langle participant \rangle \} \\ func\_dependency : [antecedent : \{ \langle DA \rangle \}] \\ fb\_dependency : [antecedent : \{ \langle FS \rangle \}] \\ rhetorical\_relation : \begin{bmatrix} antecedent : \{ \langle DA \rangle \} \\ type : \langle elaborate | \ldots \rangle \end{bmatrix} \end{bmatrix} \rangle \}
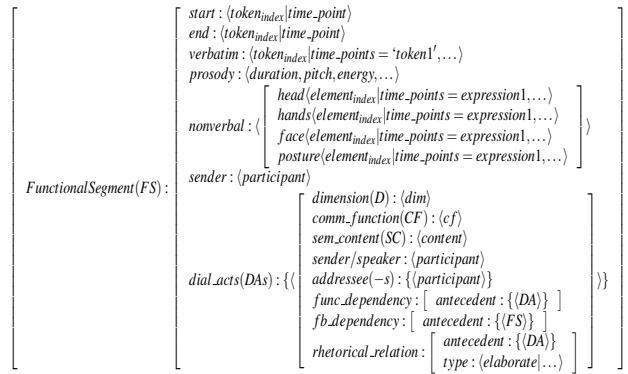\end{bmatrix}
\end{bmatrix}
$$

Figure 3: Feature structure representation of a functional segment. Adopted from Petukhova, 2011.

The proposed context model has five components: (1) **Linguistic Context (LC)** with information about (a) 'dialogue history'; (b) 'latest segment' in the form of functional segment to which one or multiple dialogue acts are assigned (see Fig. 3), and (c) 'dialogue future' (or 'planned state'); (2) **Semantic Context (SemC)** containing information about the task/domain; (3) **Cognitive Context (CC)** representing information about the current and expected participants' processing states; (4) **Perceptual/Physical Context (PC)** having information about the perceptible aspects of the communication process and the task/domain; (5) **Social Context (SocC)** containing information about current speaker's beliefs about his own and his partner's social obligations and rights.

Each of these five components contains the representation of three parts: (1) the speaker's beliefs about the task, about the processing of previous utterances, or about certain aspects of the interactive situation; (2) the addressee's beliefs of the same kind, according to the speaker; and (3) the beliefs of the same kind which the speaker assumes to be shared (or 'grounded') with the addressee. A context model for multi-party dialogues is more complex, containing representations of the speaker's beliefs about contexts of more than one addressee and possibly also of other participants (e.g. of the audience in a debate). Figure 4 shows the context model with its component structure.

Each of the model parts can be updated independently while other parts remain unaffected. For instance, the Linguistic Context is updated when dealing with linguistic/multimodal behavioural aspects and some interactive aspects, such as turn management; in the Cognitive Context participants' processing states are modelled, as well as aspects related to time and own communication management (e.g. speech production errors). The semantic

LingContext :
  speaker :
    dialogue_history : { ⟨previous_segments [Fig. 3]⟩ }
    latest_segment : [FS : [Fig. 3]
                       state = opening|body|closing]
    dialogue_future : plan : [candidates : ⟨list_DAs⟩
                               order : ⟨ordered_list_DAs⟩]
  partner : ⟨partner_linguistic_context⟩ (according to speaker)
  shared : ⟨shared_linguistic_context⟩

SemContext :
  speaker_task_model : ⟨beliefs⟩
  partner_task_model : ⟨beliefs⟩ (according to speaker)
  shared_task_model : ⟨mutual_beliefs⟩

CogContext :
  speaker_own_proc_state : [proc_problem : yes|no
                             problem_input : FS
                             time_need : negligible|small|substantial]
  partner_proc_state : ⟨partner_cognitive_context⟩ (according to speaker)
  shared : ⟨shared_cognitive_context⟩

PercContext :
  speaker : [own_presence : positive|negative
              own_readiness : positive|negative]
  partner : ⟨partner_perceptual_context⟩ (according to speaker)
  shared : ⟨shared_perceptual_context⟩

SocContext :
  speaker : [interactive_pressure : none|greet|apology|thanking|...
              reactive_pressure : {⟨dialogue_acts⟩}]
  partner : ⟨partner_social_context⟩ (according to speaker)
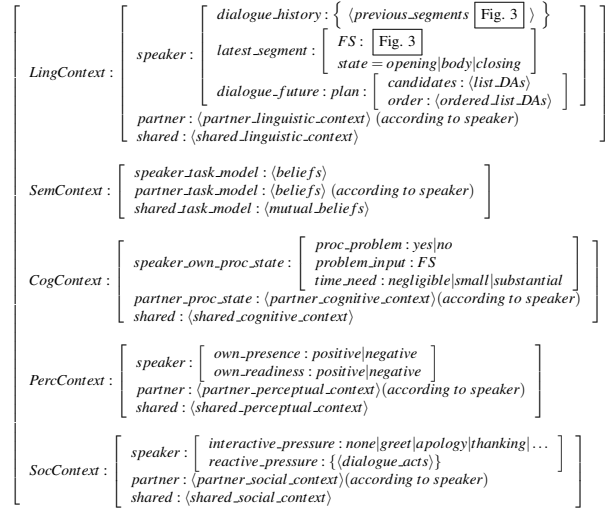 shared : ⟨shared_social_context⟩

Figure 4: Feature structure representation of the context model. Adopted from Malchanau et al., 2015.

context contains representations of task-related actions, in our scenario a participant's negotiation moves and their arguments, partners' negotiation strategies, and the system's tutoring goals and expectations on a trainee's learning progress.

4.2 Cognitive Task Agent

The Cognitive Task Agent (CTA) operates on a structured dynamic Semantic Context as described above, identifies the partner's task-related goals, and uses a strategy to compute its next negotiation move. It interprets and produces negotiation actions based on the estimation of partner's preferences and goals. The Agent adjusts its strategy according to the perceived level of the opponent's cooperativeness. Currently, the Agent distinguishes three strategies: cooperative, non-cooperative and neutral. The agent starts neutrally, requesting the partner's preferences. If the Agent believes the opponent is behaving cooperatively, it will react with a cooperative negotiation move. For instance, it will reveal its preferences when asked for, it will accept the opponent's offers, and propose concessions or cross-issues trade-offs. It will use modality triggers of liking and ability. If the Agent experiences the opponent as non-cooperative, it will switch to non-cooperative mode. It will stick to its preferences and insist on acceptance by the opponent. It will repeatedly reject the opponent's offers using modal expressions of inability, dislike and necessity. It will rarely make concessions. It will threaten to withdraw reached agreements and/or terminate negotiation. Such meta-strategies for strategy adjustment are observed in human negotiation and coordination games, see Kelley and Stahelski (1970), Smith et al. (1982). We explain in some detail how this is implemented.

4.2.1 Instance design: creation, activation and retrieval

The Agent's negotiation moves and their arguments are encoded as 'instances', represented as a set of slot-value pairs corresponding to the Agent's preference profile. Information encoded in an instance concerns beliefs about Agent's and partner's preferences (state of the negotiation and conditions), and Agent's and estimated partner's goals (actions), see Table 5. The Agent assumes that the partner's preferences are comparable to his, but values may differ. At the beginning of the interaction, the Agent may have no or weak assumptions about the partner's preferences. As the interaction proceeds the Agent builds up more knowledge about the partner's negotiation options. The Agent achieves this by taking the perspective of its partner and using its own knowledge to evaluate the partner's strategy, i.e. apply ToM skills. The Agent's memory holds three sets of preference values: the Agent's own preferences (zero ToM), the Agent's beliefs about the user's preferences (first-order ToM), and the Agent's beliefs about the user's beliefs about the Agent's preference values (second-order ToM).

When a negotiation move and its arguments are recognized, the information is passed to the CTA. The Agent constructs a retrieval instance and fills in as many slots as it can with the received details and the current context. Subsequently, the CTA updates its own representation of the negotiation state by retrieving the most *active* instance from its declarative memory. An instance i that is used most recently and most frequently gets the highest activation value, which is derived from the following equation, see Bothell (2004):

Information type	Explanation	Source
Strategy	The strategy associated with the instance	negotiationMove, modality
My-bid-value-me	The number of points the agent's bid is worth to the agent	
My-bid-value-opp	The number of points that the agent believes its bid is worth to the user	
Opp-bid-value-me	The number of points the user's bid is worth to the agent	
Opp-bid-greater	*true* if the user's bid is at least as much as the agent's current bid, *false* otherwise	Preference profile
Next-bid-value-me	The number of points that the next best option is worth	
	The next best option is defined as the option closest in value to the current one	
	(Not including those that are worth more than the current option.)	
Overall-value	The total value of all options that have been agreed upon so far.	
	This is a measure of how the negotiation is going.	History
	If it is negative, negotiation is likely to result in an unacceptable outcome.	
My-move	The move that the agent should take in this context.	Planned future

Table 5: Structure of an instance in the Cognitive Task Agent, adopted from Stevens et al. (2016a).

$$A_i = ln(\sum_{j=1}^{n} t_j^{-d}) + Logistic(0,s)$$

where n is the number of times an instance i has been retrieved in the past; t represents the amount of time that has passed since the j_{th} presentation or creation of the instance, and d is the rate of activation decay.[8] The rightmost term of the equation represents noise added to the activation level, where s controls the noise in the activation levels and is typically set at about 0.25, consistent with the value used in Lebiere et al. (2000). Thus, the equation effectively describes both the effects of recency - more recent memory traces are more likely to be retrieved, and frequency - if a memory trace has been created or retrieved more often in the past it has a higher likelihood of being retrieved.

An instance does not have to be a perfect match to a retrieval request to be activated. ACT-R can reduce its activation according to the following formula used to compute partial matching P_i, see Bothell (2004):

$$P_i = \sum_{l} PM_{li}$$

where M_{li} indicates the similarity value between the relevant slot value in the retrieval request (l) and the corresponding slot instance i summed over all slot values in the retrieval request. P denotes the mismatch penalty and reflects the amount of weighting given to the matching, i.e. when P is higher, activation is more strongly affected by similarity. We set the constant P high at 5, consistent with the value used in Lebiere et al. (2000).[9] The Agent will thus be able to retrieve past instances for reasoning even when a particular situation has not been encountered before. Partial matching, combined with activation noise, allows for flexibility in the Agent's behaviour. The Agent will not rigidly make the exact same moves every time.

For example, suppose the CTA retrieves the following instance:

```
instance-a
strategy            cooperative    the opponent's strategy is cooperative
my-bid-value-me     4              the agent's current offer is worth 4 points to him
opp-bid-value-me    1              the opponent's offer is worth 1 point to the agent
opp-bid-greater     true           the opponent's offer is equal or greater than agent's current bid
next-bid-value-me   2              the next best option for the agent is worth 2 points
opp-move            concede        opponent changed its offer to one that was less valuable to him
my-move             concede        the agent repays the opponent by also selecting a less valuable option
```

Two pieces of information will be extracted from these instances: the strategy of the user (cooperative) and an estimate of the user's preference for the options mentioned in the move (1,true). If there are other good options available, a cooperative negotiator will explore those options first before insisting on his current position, so from this behaviour the Agent infers that it is dealing with a cooperative negotiator with positive preferences on at least two issues. Now the Agent uses its own context to choose an appropriate response to the user. Depending on how the user has acted, and what the Agent knows (guesses) about the user's preferences, the Agent chooses to respond cooperatively, i.e. to concede.

---

[8]In the ACT-R community, 0.5 has emerged as the default value for the parameter d over a large range of applications, Anderson et al. (2004).

[9]To disable partial matching P can be set at 0.

4.2.2 Multitasking behaviour

The CTA can reason about the overall state of the negotiation task, and attempts to identify the best negotiation move for the next action. It computes: (1) the Agent's counter-move, and (2) feedback sharing the Agent's beliefs about the user's preferences and the user's negotiation strategy. The Agent may propose a strategically better alternative move that the user could have taken and explain 'why'. As the result, the system is able to play simultaneously or interchangeably the four roles specified in Section 2.1: Observer, Negotiator, Mirror and Tutor.

In the Observer mode, the Agent monitors and keeps track of all performed own and partner's actions and logs them. The created log files are used to evaluate the participants' performance and for system improvement (see Section 6).

As a Mirror, the Agent's monitoring and interpretation results are immediately displayed to the user. These displays include a transcript of the Agent's and user's utterances (as recognized by the system), the Agent's perceived cooperativeness level and the recognized partner's preferences. The Agent's and partner's most recent offers and estimated partner's preferences are also flagged in the dynamically updated preference card (Fig. 2). The latter can have certain tutoring effects as well, since it may activate a user's monitoring, reflection and regulating strategies, but also trigger a user's corrective actions in case of Agent processing failures.

As a Negotiator, the Agent takes into account the recognized partner negotiation strategy, the Agent's preferences, and the estimation of those of the partner, and computes the most appropriate next negotiation move. This leads to relevant updates in the Semantic Context that give rise to goals to perform a certain dialogue act, e.g. tentative Agreement. Other contexts may be also updated in parallel and goals are created to perform, for example, turn-taking (Linguistic Context) and feedback (Cognitive Context) actions, see next section. The Dialogue Manager passes dialogue act list for generation, $< DA_1 = turnTake, DA_2 = positiveAutoFeedback, DA_3 = Task; Agreement >$, where DA_1 is decided to be generated implicitly, DA_2 - non-verbally by a smiling and nodding avatar and verbally by 'okay', and DA_3 is generated by the utterance 'I can live with it'.

As a Tutor, the Agent shares its beliefs about the current negotiation state and its planned continuation, e.g. may offer strategically better user negotiation moves leading to higher quality negotiation outcomes in terms of Pareto efficiency. After each action, the Agent is also able to provide an explanation why decisions are made to perform certain actions. At the end of each negotiation session summative feedback is generated in terms of estimated Pareto optimality, degree of cooperativeness, and acceptance of negative outcomes. This type of feedback accumulates across multiple consecutive negotiation rounds.

The execution of these shared and varied tasks is expected to have positive effects both on user and system performance, enabling activation and improvement of metacognitive processes. Moreover, since these processes do not require additional resources (memory, processing and control), but are model-inherent belief creation and transfer processes and characteristics (instance slots), multiple tasks related to various roles can be executed by the DM in parallel without interference.

4.2.3 Dialogue Manager state update: example

Table 6 provides an example of a dialogue between an agent A playing the role of the Business Representative and a human negotiator C in the role of the City Councilor. The CTA starts neutrally. A elicits an offer from C on the first issue and does this in the form of a Set Question. The understanding that a certain dialogue act is performed leads to corresponding context model updates.[10] If the partner reacts to the agent's elicitation by sharing his preferences in C_1, he is evaluated by the agent as being cooperative. The agent's preferences are not identical but not fully conflicting either: it is possible for the agent to agree with the opponent's preferences accepting his offer in $A_{2.2}$, where A believes that the offer made in C_1 is not the most preferred one but still acceptable/possible for A.[11] The CTA stays in the cooperative mode. If the negotiator's preferences differ from the options proposed by the partner, he may refuse to accept the partner offer as in $C_{2.1}$ and may offer another value which is more preferable for him, i.e. perform a counter-offer move ($C_{2.2}$ repeated in C_3 after the agent signaled that his processing was unsuccessful. The CTA interprets the partner's strategy as being non-cooperative and switches his strategy to neutral, proposing to exchange offers (in $A_{4.2}$) that still aim at the

---

[10]A detailed specification of dialogue act update semantics is provided in Bunt (2014b) and Petukhova (2011).

[11] We provide here a simplified representation of the participants' information states as tracked and updated by the DM. The full specification of participants' information states and their updates can be found in Petukhova et al. (2017).

	Utterance	DM input			DM Information State		DM output
ID	(wording)	DA ID	D;CF [dependence]	SC=NM($I;V$)/ modality	CTA state /decision	DM update[11]	DA for generation
A_1	what do you want for scope	da1		p_1=offer(1;?v)	neutral/elicit	$Wants(A, Know(A, p_1))$	task;setQuestion
C_1	i think it would be fine					$Bel(C, Wants(A, Know(A, p_1)))$	
	if we stop smoking			p_2=offer(1;b)/		$Bel(C, \Box p_2)$	
	in public transportation	da2	task;answer[da1]	prefer	cooperative	$Wants(C, Know(A, \Box p_2))$	
$A_{2.1}$	okay					$Bel(A, Wants(C, Know(A, \Box p_2)))$	
$A_{2.2}$	i would go for that point				cooperative/	$Bel(A, \neg\Box p_2))$	
		da3		p_2=offer(1;b)	agree(A,1b)	$Wants(A, Know(C, \Diamond p_2))$	task;agreement
$A_{2.3}$	i prefer that we do not				cooperative/		
	increase taxes	da4		p_3=offer(2;a)	offer(A,2a)	$Wants(A, Know(C, \Box p_3))$	task;inform
$C_{2.1}$	this is not possible	da5	task;	p_3=offer(2;a)	non-cooperative/	$Bel(C, Wants(A, Know(C, \Box p_3)))$	
	for us		disagreement[da4]	unable	reject(C,2a)	$Wants(C, Know(A, \neg\Diamond p_3))$	
$C_{2.2}$	i would like 5% tax						
	increase at least	da6	task;inform	-	-	-	
A_3	sorry i was distracted						apology &
	could you repeat	da7				$Bel(A, \neg Interpreted(A, C_{2.2}))$	autoNegative
C_3	sure	da8	autoPositive[da7]			$Bel(C, Interpreted(C, A_3))$	
	I prefer 5% tax increase	da9	task;inform	p_4=offer(2;b)	non-cooperative/	$Bel(C, Wants(A, Know(C, \neg\Diamond p_3)))$	
				prefer	offer(C,2b)	$Wants(C, Know(A, \Box p_4))$	
$A_{4.1}$	okay	da10	autoPositive[da9]	p_4=offer(2;b)		$Bel(A, Interpreted(A, C_3))$	autoPositive
$A_{4.3}$	i will give you 5% if you			p_4=offer(2;b)		$Bel(A, Wants(C, Know(A, \Box p_4)))$	
						$Bel(A, \Diamond p_4)$	
	agree to ban all tobacco						
	vending machines			p_5=offer(4;b)	neutral/	$Bel(A, \Box p_5)$	
		da11		p_6=offer(	exchange(A,	$Bel(A, \Diamond p_6)$	task;offer
				(2;b) $\rightarrow$ (4;b))	(2b $\wedge$ 4b))	$Wants(A, Know(C, \Diamond p_6))$	
C_4	i think i can live with that	da12	task;	p_6=offer(	cooperative/agree(C,	$Bel(C, Wants(A, Know(C, \Diamond p_6)))$	
			agreement[da11]	(2;b) $\rightarrow$ (4;b))	(2b $\rightarrow$ 4b))	$Bel(C, \Diamond p_6)$	

Table 6: Example of a negotiation dialogue with processing and generation by the Dialogue Manager. (A = agent (Business Representative); C = human negotiator (City Councilor); DA = dialogue act; D = dimension; CF = communicative function; SC= semantic content; NM = negotiation move; I = issue; V=value; Bel = believes; $\Diamond$ = possible; $\Box$ = preferable)

better deal for himself. If this is again rejected, the agent will apply the non-cooperative strategy and insist on his previous proposal expressed in $A_{2.2}$, otherwise he will either elicit an offer for the next issue or propose an offer himself.

The agent computes the partner's negotiation strategy using the linguistic modality expressed in the partner's utterance and the type of the dialogue act performed. The collected data was used to train classifiers in the supervised setting to make such predictions, see Section 5. To assess the minimal amount of data required to detect the partner's negotiation strategy reliably, a series of learnability experiments was performed. To achieve an accuracy higher than 75%, about 1300 training instances are used. It was noticed the classifier performance further benefits from adding more training data. An accuracy of 83% was achieved on a training set comprising 3800 instances, so twice as many as in the first iteration and consuming almost the entire human-human MIB corpus. The system showing this performance was evaluated, see Section 6. Follow-up experiments indicated that adding more data (e.g. evaluation and simulated data) further improves the classification performance, although not significantly, gaining 1% in accuracy when adding additional 1000 instances.

4.3 Dialogue control

Task actions account for less than half of all actions in our negotiation data, see Table 4. Other frequently occurring acts are concerned with Task Management, Discourse Structuring, Feedback and Social Obligations. Along with moving towards a final set of agreements, negotiators need to take care how to optimally structure and manage the negotiation and the interaction. In multi-issue bargaining, negotiators have a variety of task management strategies. They may discuss issues sequentially or bargain simultaneously about multiple options, making trade-offs across issues. They may withdraw and re-negotiate previously reached agreements. All these decisions require explicit communicative actions. The Task Management acts are recognized and generated by the system, and are modelled as part of the system's Semantic Context containing, along with the information about the speaker's beliefs about the negotiation domain, information concerning task progress and success. A Task Planner as part of the Task Manager (see Fig. 3) takes care of updates and generation processes of this type.

Acts related to negotiators' perception of the partner's physical presence and readiness to start, continue or terminate the interaction as well as participants' beliefs concerning the availability and properties of communicative and perceptual channels are modelled as part of the Perceptual Context. Dialogue behaviour addressing these aspects is important, in particular, these actions are considered for generation, since the system's mul-

Processing level	Latest dialogue act		Previous dialogue act	Planned dialogue act
	Communicative Function	Negotiation Move		
Perception	unknown	unknown	any	Request Repeat and/or AutoNegative
Interpretation	unknown	offer(x)	any	Accept(offer(x)) or Reject(offer(x))
	unknown	offer(unknown)	any	Question(offer(?)) and/or AutoNegative
	any	unknown	any	Request Repeat or Rephrase and/or AutoNegative
	unknown	unknown	Accept(offer(x)) or Reject(offer(x))	Question(offer(?)) or Inform(offer(y))
	unknown	unknown	offer(x)	Question(offer(y))
	Question	unknown	Accept(offer(x)) or Reject(offer(x))	Inform(offer(y))

Table 7: Decision-making support for the system's feedback strategies concerning perception and interpretation of task-related actions, and expected dialogue continuation. Note: $x \neq y$.

timodal behaviour related to Contact Management is embodied by a virtual character (full body avatar). The Contact Manager takes care of updates and the generation of these acts. A participant's beliefs concerning the interaction structure (i.e. history, present and future states) and beliefs concerning topic shifts are modeled as a part of the Linguistic Context; the Discourse Structuring module takes care of the updates and generation specific for the interaction management and monitoring.

4.4 Validity checking, repair and clarification strategies

For an interactive system it is important to know that its contributions are understood and accepted by the user, as well as to signal the system's processing of the same kind. Conversation is a **bilateral process** - that is, a joint activity, and speaking and listening are not autonomous processes - conversational partners monitor their own processing of the exchanged utterances as well as the processing done by the others, see Clark and Krych (2004) for discussion. Given the bilateral nature of conversation, interlocutors can construct and provide feedback on both their own processing (*auto-feedback*) as and on that by the other (*allo-feedback*).

Feedback is crucial for successful communication. Feedback can be provided at different **levels** of processing the communicative behaviour of interlocutors. Allwood et al. (1993) and Clark (1996) notice that interlocutors need to establish *contact* and gain or pay *attention* to each others behaviour, in order be involved in conversation. A speaker's behaviour needs to be *perceived* (i.e. heard, seen) or *identified* (Clark, 1996). Perceived behaviour should be *interpreted*, i.e. interlocutors should be able to extract the meaning of each other's behaviour. The constructed interpretation needs to be *evaluated* against one's information state: if it is consistent with the current information state it can be incorporated into that state; if it is inconsistent, this can be reported as negative feedback. The incorporation of new information, and the performance of other mental and physical actions in response to communicative behaviour is called the *execution* or *application* (Bunt, 2000). A speaker may provide feedback (*feedback giving*) or elicit feedback (*feedback eliciting*).

As for positive feedback acts, explicitly signalled acceptances are generated, either verbally or non-verbally. We also consider generation of multimodal expressions of implied and entailed positive feedback (see Bunt 2007; 2012) for strategic reasons, e.g. to provide more certainty due to potentially erroneous automatic speech recognition output.

Detected difficulties and inconsistencies in recognition, interpretation, evaluation and execution need to be resolved immediately if these problems are serious enough to impede further task performance; such problems are reported accordingly. Problems due to deficient recognition and interpretation are frequent in spoken human-computer dialogue systems, but rarely observed in the collected human-human dialogue data. Good news however is that humans generally exhibit certain re-occurring behavioural patterns when their processing fails. For our scenario and dialogue setting we incorporated observations and analyses of other available dialogue resources such as the human-human AMI and HCRC MapTask corpora (Carletta, 2006; Anderson et al., 1991), and human-human and human-computer DBox quiz game data (Petukhova et al., 2014; 2015).

Observations from human-human and human-computer dialogues resulted in the definition of feedback strategies at the level of perception (recognition) and interpretation mostly comprising corrections and requests to repeat or rephrase (Table 7), at the level of evaluation reporting inconsistencies/(in)validity due to certain

Processing level	Latest dialogue act		Previous dialogue act	Validity	Planned dialogue act
	Communicative Function	Negotiation Move			
Evaluation	Inform	terminate	any	valid	stop negotiation
	Accept	offer(x)	final Offer(x)	valid	Inform(deal(x))
	any other than Accept	offer(x)	final Offer(x)	invalid	Auto/AlloFeedback: Question(?offer(x))
	Accept	deal(x))	Inform(deal(x))	valid	DiscourseStructuring: TopicShift TaskManagement:Suggest(next_issue) DiscourseStructuring: Closing
	Reject or Accept	offer(x)	Suggest(offer(x))	valid	Inform(offer(y)); Question(offer(y))
	Reject or Accept	offer(x)	Inform(offer(x))	valid	Inform(offer(y)); Question(offer(y))
	Question	offer(x)	Inform(deal(x))	valid	Inform(deal(x))
	Reject	offer(x)	Reject(offer(x))	valid	Reject(offer(x))
	final Offer or Accept	offer(x)	Accept(offer(x))	valid	Accept(offer(x))
	Inform Suggest Offer	offer(x)	Accept(offer(y))	valid if x= ¬y	Accept(offer(?x)) or Reject(offer(x))
	Inform	offer(y)	Inform(deal(x))	invalid	interpret as Accept(deal(x)) if x=y otherwise Reject(offer(x))
	Accept	offer(x)	any other than Suggest or Inform(offer(x))	invalid	interpret as Inform(offer(x)) or Suggest(offer(x)) if x=y otherwise generate Auto- or AlloNegative
	Reject	offer(x)	any other than Suggest or Inform(offer(x))	invalid	interpret as Inform(offer(y)) or Suggest(offer(y)) if x¬y otherwise generate Auto- or AlloNegative
	Accept	offer(x)	Inform(terminate)	invalid	interpret as Accept(terminate) and generate Auto- or AlloNegative
	Reject	offer(x)	Inform(terminate)	invalid	interpret as Accept(terminate) and/or generate Auto- or AlloNegative
	Inform	deal(x)	Inform(terminate)	invalid	Question(offer(?)) and/or Auto- or AlloNegative
	Inform Suggest Offer	offer(x)	Accept(offer(y))	invalid if x=y	Question(offer(?x)); Accept(offer(y)) and/or Auto- or AlloNegative
	Inform Suggest Offer	offer(x)	Reject(offer(y))	invalid if x=y	Question(offer(?x)); Reject(offer(y))and/or Auto- or AlloNegative
	Inform	deal(x)	Reject(offer(x))	invalid	Reject(offer(x)); Question(offer(?x)) and/or Auto- or AlloNegative

Table 8: Decision-making support for the system's recovery and clarification strategies concerning evaluation of task-related actions, and expected dialogue continuation. In this table, *valid* stands for the state that can be recovered from the available information, otherwise *invalid* - state that cannot be automatically recovered and requires activation of the clarification strategy. Note: $x \neq y$.

Processing level	Latest dialogue act		Previous dialogue act	Preferences	Validity	Planned dialogue act
	Communicative Function	Negotiation Move				
Execution	Inform Suggest Offer	offer(x)	any	negative	valid	Reject(offer(x)) and/or Inform(offer(y)) and/or AutoNegative
	Inform Suggest Offer	offer(x)	any	positive	valid	Accept(offer(x)) and/or Inform(offer(y)) and/or AutoNegative
	Inform Suggest Offer	offer(x)	any	neutral	valid	Accept(offer(x)) and/or Inform(offer(y))
	Inform	deal(x)	no Accept(offer(x))	any	invalid	Reject(deal(x)); Question(offer(?x)) and/or AutoNegative
	Inform	deal(x)	Reject(offer(x))	any	invalid	Reject(deal(x)); Question(offer(?x)) and/or AutoNegative
	inform	terminate	no final offer(x)	any	invalid	Question(offer(?x)) and/or AutoNegative

Table 9: Decision-making support for the system's feedback strategies concerning execution of task-related actions. In this table, *valid* stands for the state that can be recovered from the available information, otherwise *invalid* - state that cannot be automatically recovered and requires activation of the clarification strategy. Note: $x \neq y$

logical constraints, given the grounded negotiation history (Table 8), and at the level of execution reporting inability to accept an offer or to reach an agreement due to the negotiator's preference profile (Table 9). Certain system processing flaws can be recovered from the information available to the system, some problems are too severe to continue the dialogue successfully and trigger feedback acts (clarification requests). In total, about 30 clarification and recovery strategies have been defined and evaluated (see also Section 6).

Information concerning successes and failures in the processing of a partners' dialogue contributions are modelled as part of the Cognitive Context (see Fig. 3).

Thus, dialogue control acts present an important part for any interaction. In a shared cultural and linguistic context, choices concerning the frequency of such actions and the variety of expressions are rather limited. Conventional forms are mostly used to greet each other, to apologize, to manage the turns and the use of time, to deal with speaking errors, and to provide or elicit feedback. Models of dialogue control behaviour once designed can therefore be applied in a wide range of communicative situations. The use of task-related dialogue acts, by contrast, is more application-specific. The separation between task-related and dialogue control actions is therefore not only a cost-effective solution, but also allows designing flexible architectures and combinations of different modelling approaches and techniques, resulting in more robust and rich system behaviour.

4.5 Dialogue Manager architecture

The above considerations have resulted in a Dialogue Manager consisting of multiple Agents corresponding currently to six ISO 24617-2 or DIT^{++} dimensions[12]: the Task Manager with the integrated CTA and Task Planner for task control, the Auto/Allo Feedback Agent, the Turn Manager, the Discourse Structuring Manager, the Contact Manager, and the Social Obligations Manager.

The Dialogue Manager (DM) is designed as a set of processes ('threads') that receive data, update the information state, and generate output. Additionally, consistency checking and conflict resolution is performed to avoid that the context model would be updated with inconsistent or conflicting information and incompatible dialogue acts are generated, see also Petukhova (2011). Figure 5 presents the overall DM architecture. First, data are received from the Fusion/Interpretation module. Next, the information state ('context model') is updated based on the received input. The Process Manager decides what parts of the context model to update. Following receiving and updating, the output based on the analysis of the information state is generated. The output presents the ordered list of dialogue acts which is sent to the Fission module, see next Section for complete dialogue system architecture.

---

[12]The set of Agents may in future be extended to include all nine ISO 24617-2 dimensions and possibly other additional dimensions.

Figure 5: Cognitive Task Agent (grey box) incorporated into the Dialogue Manager architecture: fused dialogue act information is passed to the Dialogue Manager from the Interpretation Manager for context model update and next action(-s) generation which are 'fissed' in different output modalities; both processes are regulated by the Process Manager.

5 The Virtual Negotiation Coach: design and evaluation

As a proof of concept, and for assessing the potential value of the integration of a cognitive agent into a dialogue manager, we designed the Virtual Negotiation Coach (VNC), an interactive system with the functionality described in the scenario for data collection (Section 3.2). The VNC gets a speech signal, recognizes and interprets it, identifies relevant actions and generates multimodal actions, i.e. speech and gestures of a virtual negotiator and positive and negative visual feedback for tutoring. Figure 6 shows the VNC architecture and processing workflows.

Speech signals are recorded from multiple sources, such as wearable microphones, headsets for each dialogue participant, and an all-around microphone placed between participants. The speech signals serve as input for two types of further processing: Automatic Speech Recognition (ASR), leading to lexical, syntactic, and semantic analysis, and prosodic analysis concerned with voice quality, fluency, stress and intonation of speech. The Kaldi-based ASR component incorporates acoustic and language models developed using various available data sources: the Wall Street Journal WSJ0 corpus[13], HUB4 News Broadcast data[14], the VoxForge corpus[15], the LibriSpeech corpus[16] and AMI project data[17]. In total, about 759 hours of data has been used to train an acoustic model. The collected in-domain negotiation data is used as language model adaptation. The background language model is based on a combination of different corpora, like the approach taken to train the acoustic model. The ASR performance is measured at 34.4% Word Error Rate (WER), see Singh et al. (2017)[18]. The ASR outputs a single best word sequence without any scores. Prosodic properties were computed automatically using PRAAT (Boersma and Weenink, 2009) such as minimum, maximum, mean, and standard deviation of *pitch*, *energy*, *voicing* and speaking rate.[19]

The ASR output is used by the negotiation moves and dialogue act classifiers. Negotiation moves specify events and their arguments represented as *NegotiationMove(ISSUE; VALUE)*. Conditional Random Field

---

[13]https://catalog.ldc.upenn.edu/ldc93s6a

[14]https://catalog.ldc.upenn.edu/ldc98s71

[15]http://www.voxforge.org/

[16]http://www.openslr.org/12/

[17]http://groups.inf.ed.ac.uk/ami/corpus/

[18]It should be noticed that the ASR performance has been measured when interacting with non-native English speakers, who significantly varied in language skills level and speech fluency, some having a rather strong Greek accent.

[19]We computed both raw and normalized versions of these features. Speaker-normalized features were obtained by computing z-scores (z = (X-mean)/standard deviation) for the feature, where mean and standard deviation were calculated from all functional segments produced by the same speaker in the debate session. We also used normalizations by the first speaker turn and by prior speaker turn.
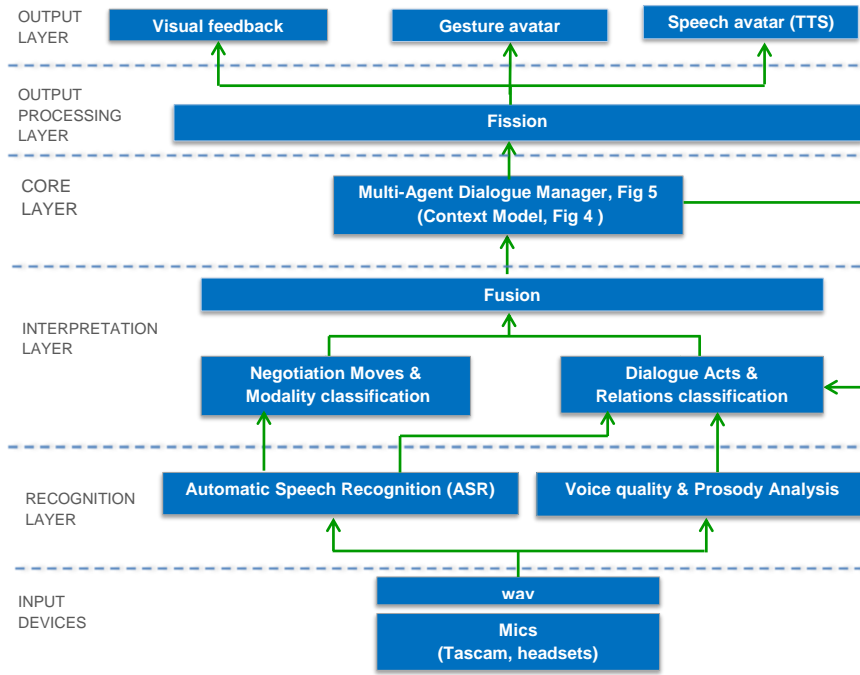
Figure 6: Architecture of the Virtual Negotiation Coach system. From bottom to top, signals are received through input devices, further recognized by tailored processing modules. After interpretation concerned with Negotiation Moves, Modality and Dialogue Act classification, semantic representations from different modalities and modules are fused as Dialogue acts. Fused dialogue act information is passed to the Dialogue Manager for context model update and next action generation. The generated system response is rendered or 'fissed' in different output modalities. Adopted with extensions and adjustments from van Helvert et al., 2016

.

models for sequence learning (CRF, Lafferty et al. (2001)) are trained to predict three types of classes (move, issue and value) and their boundaries in ASR n-best strings: negotiation move, issue, preference value. A tenfold cross-validation using 5000 words of transcribed speech from the negotiation domain yielded an F-score of 0.7 on average.

For the recognition of the intentions encoded in participants' utterances various machine learning techniques have been applied, such as Support Vector Machine (SVM, Boser et al., 1992), Logistic Regression (Yu et al., 2011), AdaBoost (Zhu et al., 2009), and the Linear Support Vector Classifier (Vapnik, 2013). F-scores ranging between 0.83 and 0.86 were obtained, which corresponds to state-of-the-art performance, see Amanova et al. (2016). The incremental token- and chunk-based dialogue act CRF-classifiers showed a performance of .80 F-scores on average, see Ebhotemhen et al. (2017). After extensive testing, a non-incremental SVM-based classifier has been integrated into the VNC system. The SVM-based modality classifiers show accuracies in the range between 73.3 and 82.6% Lapina and Petukhova (2017). Finally, information from the Linguistic Context related to the dialogue history has been used to ensure context-dependent interpretation of dialogue acts. Additionally, the trainee has a choice to select options using a graphical interface as depicted in Figure 2. As task progress support, partner offers and possible agreements are visualized with red (system) and green arrows (user).

The system's Fusion module currently fuses interpretations from two modules obtaining full semantic representations of user speech contributions. In the future, we will extend the system to other non-verbal modalities by integrating modern sensing technology at the input level. Given the dialogue acts provided by the Dialogue Manager, the Fission module generates responses splitting their parts into different modalities, such as Avatar[20] and Voice (TTS[21]) for negotiation actions, and visual feedback for tutoring actions. The latter includes a representation of the negotiators' current cooperativeness, visualized by happy and sad face emoticons.

At the end of each negotiation session, summative feedback is generated specifying the number of points

---

[20]Commercial software of Charamel GmbH has been used, see Reinecke (2003)

[21]Vocalizer of Nuance, http://www.nuance.com/for-business/text-to-speech/vocalizer/index.htm, was integrated.

gained or lost for each partner, the number of negative agreements, and the Pareto optimality of the reached agreements. All messages exchanged between modules are in the standard TEI and ISO DiAML formats.

6 Evaluation

It is generally not a trivial task to evaluate the performance of a Dialogue Manager as a single module due to its dependency on the quality of its potentially erroneous inputs.

The performance of a DM is often evaluated as a part of the integrated dialogue system in a user-based fashion, by letting end users assess their interaction with the system. Such assessment is typically based on the satisfaction of the users with the completion of the task. For example, PARADISE, one of the most widely-used evaluation models (Walker et al., 1997), predicts user global satisfaction given a set of parameters related to task success and dialogue costs. Satisfaction is calculated as the arithmetic mean of nine judgements on different quality aspects rated on 5-point Likert scales. Subsequently, the relation between task success and dialogue costs parameters and the mean human judgement is estimated carrying out a multivariate linear regression analysis. Another way to evaluate a dialogue system is on the basis of interaction with computer agents that substitute human users and emulate user behaviour, see e.g. López-Cózar et al. (2006). The various types of users and system factors can be systematically manipulated, e.g. interactive, dialogue task and error recovery strategies.

Several sets of parameters have been recommended for spoken dialogue system evaluation, ranging from a single BLEU score metric for end-to-end system evaluation (Wen et al., 2017), to seven parameters related to the entire dialogue (duration, response delay, number of turns) defined in Fraser (1998) and 52 parameters in Möller (2004) to meta-communication strategies (number of help requests, correction turns), to the system's cooperativity (contextual appropriateness of system utterances), to the task which can be carried out with the help of the system (task success, solution quality), as well as to the speech input performance of the system (word error rate, understanding error rate).

As for measuring satisfaction, various questionnaires have been proposed: nine satisfaction questions defined within PARADISE (Walker et al., 2000); 44 evaluative statements of the Subjective Assessment of Speech System Interfaces (SASSI) questionnaire (Hone and Graham, 2001); 53 evaluative statements in REVU (Report on the Enjoyment, Value, and Usability, Dzikovska et al., 2011); 24 bipolar adjective pairs defined in the Godspeed questionnaire (Bartneck et al., 2009); 122 evaluative statements in the Questionnaire for User Interface Satisfaction (QUIS version 7.0, Chin et al., 1988). The absence of standard performance metric sets and questionnaires for dialogue system evaluation makes it difficult to compare the results from different studies, and the various existing dialogue system evaluation results exhibit great differences.

One of the common practices is to evaluate an interactive system or user interface by measuring usability, using well-defined observable and quantifiable metrics (see ISO 9241-11 and ISO/IEC 9126-4 standards for usability metrics for effectiveness, efficiency and satisfaction). For this purpose, the usability perception questionnaire was constructed assessing eight main factors: task completion and quality, robustness, learnability, flexibility, likeability, ease of use and usefulness of the application,. The questionnaire has sufficient internal consistency reliability (Cronbach's alpha of 0.87) and comprises 32 evaluative statements[22].

Using this questionnaire, we collected human judgements concerning the system performance in 28 evaluation sessions, with 28 participants aged 25-45, all professional politicians or governmental workers. Nine negotiation scenarios were used, based on different negotiator preference profiles, see Petukhova et al. (2016). Participants were assigned a Councilor role and a random scenario. The questionnaire allows human judgements to be linked to the performance of certain modules (or module combinations), see Table 11. User judgements were presented in 5-point Likert scales.

The usability of the VNC system was measured in terms of effectiveness, efficiency and satisfaction. Previous research suggests that there are differences in perceived and actual performance (Nielsen, 2012): performance and perception scores are correlated, but they are different usability metrics and both need to be considered when conducting quantitative usability studies. In our design, subjective perception of effectiveness, efficiency and satisfaction were correlated with various performance metrics and interaction parameters to assess their impact on the qualitative usability properties. We computed bi-variate correlations to determine

---

[22]The usability questionnaire is available at `https://docs.google.com/forms/d/e/ 1FAIpQLSf1H110UOflMGAqtT0HacBD7T0nIHqLOKoi1qS7o28wLzizSw/viewform`

Evaluation criteria	Human-human	Human-computer
Number of dialogues	25 (5808)	185 (NA)
Mean dialogue duration (in turns per dialogue)	23 (6.6)	40 (NA)
Agreements (%)	78 (80.1)	66 (57.2)
Pareto optimal (%)	61 (76.9)	60 (82.4)
Negative deal (%)	21 (NA)	16 (NA)
Cooperativeness rate (%)	39 (NA)	51 (NA)

Table 10: Comparison of human-human and human-agent negotiation behaviour. Adopted from Petukhova et al. (2017). In brackets the best results reported by Lewis et al. (2017) for comparison. *NA* stands for not applicable, i.e. not measured.

possible factors impacting user perception of the system usability and the performance metrics and interaction parameters derived from logged and annotated evaluation sessions.

As performance metrics, system and user performance related to task completion rate[23] and its quality[24] were computed. We also compared system negotiation performance with human performance on the number of agreements reached, the ability to find Pareto optimal outcomes, the degree of cooperativeness, and the number of negative outcomes[25]. It was found that participants reached a lower number of agreements when negotiating with the system than when negotiating with each other, 66% vs 78%. Participants made a similar number of Pareto optimal agreements (about 60%). Human participants show a higher level of cooperativity when interacting with the system, i.e. 51% of the actions are perceived as cooperative. This may mean that humans were more competitive when interacting with each other. A lower number of negative deals was observed for human-agent pairs, 21% vs 16%. Users perceived their interaction with the system as effective when they managed to complete their tasks successfully reaching Pareto optimal agreements by performing cooperative actions but avoiding excessive concessions. Our results differ from those reported in Lewis et al. (2017) for both the human-human and the human-agent setting, see Table 10. However, as noticed above, due to differences in tasks, scenario and interactive setting it is hard to draw clear comparative conclusions. Nevertheless, we can conclude that the implemented CTA is capable of making decisions and performing actions similar to those of humans. No significant differences in this respect were observed between human-human and human-system interactions.

As for efficiency, we assessed temporal and duration dialogue parameters, e.g. time elapsed and number of system and/or user turns to complete the task (or a sub-task) and the interaction as a whole. We also measured the system response time, the silence duration after the user completed his utterance and before the system responded. Weak negative correlation effects have been found between user perceived efficiency and system response delay, meaning users generally found the system reaction and the interaction pace too slow. Dialogue quality is often assessed measuring word and sentence error rates (Walker et al., 1997; López-Cózar et al., 2006) and turn correction ratio (Danielli and Gerbino, 1995). Many designers have noticed, however, that it is not so much how many errors the system makes that contributes to its quality, but rather the system's ability to recognize errors and recover from them. This contributes to the perceived system robustness and is appreciated by users. Users also value if they can easily identify and recover from their own mistakes. All system's processing results were visualized to the user in a separate window, which contributes to the system observability. The repair and recovery strategies used by the system and the user were evaluated by two expert annotators, whose agreement was measured in terms of kappa. Repairs were estimated as the number of corrected segments, recoveries as the number of regained utterances which were partially failed at recognition and understanding, see also Danieli and Gerbino (1995). While annotators agreed that repair strategies were applied adequately, longer dialogue sessions due to frequent clarifications are undesirable.

The VNC is evaluated to be relatively easy to interact with (4.2 Likert points). However, users found an instruction round with a human tutor prior to the interaction useful. Most users were confident enough to interact

---

[23] We consider the overall negotiation task as completed if parties agreed on all four issues or parties came to the conclusion that it is impossible to reach any agreement.

[24] Overall task quality was computed in terms of number of *reward points* the trainee gets at the end of each negotiation round and summing up over multiple repeated rounds; and *Pareto optimality* (coefficient from 0 to 1) which reflects a state of affairs when there is no alternative state that would make any partner better off without making anyone worse off.

[25] We considered negative deals as flawed negotiation action, i.e. the sum of all reached agreements resulted in an overall negative value meaning that the trainee made too many concessions and selected mostly dispreferred bright 'orange' options (see Figure 2).

Usability metric	Perception	Performance		R
	Assessment	Metric/parameter	Value	
effectiveness (task completeness)	mean rating score effectiveness 4.08	Task completion rate[23]; in %	66.0	.86*
		Reward points[24]; mean, max.10	5.2	.19
effectiveness (task quality)		User's Action Error Rate (UAER, in %)[25]	16.0	.27*
		Pareto optimality[24]; mean, between 0 and 1	0.86	.28*
		Cooperativeness rate; mean, in %	51.0	.39*
efficiency (overall)	mean rating score efficiency 4.28	System Response Delay (SRD); mean, in ms	243	-.16
		Interaction pace; utterance/min	9.98	-.18
		Dialogue duration; in min	9:37	-.21
		Dialogue duration; average, in number of turns	56.2	-.35*
efficiency (learnability)	3.3 (mean)	User Response Delay (URD); mean, in ms	267	-.34*
efficiency (robustness)	3.2 (mean)	System Recovery Strategies (SRS); correctly activated (Cohen's κ)	0.89	.48*
		User Recovery Strategies (URS); correctly recognized (Cohen's κ)	0.87	.45*
efficiency (flexibility)	3.8 (mean)	Proportion spoken/on-screen actions; mean, in % per dialogue	4.3	.67*
satisfaction (overall)	aggregated per user ranging between 40 and 78	ASR Word Error rate; WER, in %	22.5	-.29*
		Negotiation moves recognition; accuracy, in %	65.3	.39*
		Dialogue Act Recognition; accuracy, in %	87.8	.44*
		Correct responses (CR)[29]; relative frequency, in %	57.6	.43*
		Appropriate responses (AR)[28]; relative frequency, in %	42.4	.29*

Table 11: Summary of evaluation metrics and obtained results in terms of correlations between subjective perceived system properties and actions, and objective performance metrics (R stands for Pearson coefficient; * = statistically significant ($p < .05$)

with the system on their own, some of them however found the system too complex and experienced difficulties in understanding certain concepts/actions. A performance metric which was found to negatively correlate with system learnability is user response delay, the silence duration after the system completed its utterance and the user proposed a relevant dialogue continuation. Nevertheless, the vast majority of users learned how to interact with the system and complete their tasks successfully in consecutive rounds. We observed a steady decline in user response delays from round to round.[26]

Users appreciated the system's flexibility. The system offered the option to select continuation task actions using a graphical interface on a tablet in case the system processing failed entirely. The use of concurrent multiple modalities was positively evaluated by the users. It was always possible for users to take initiative in starting, continuing and wrapping up the interaction, or leave these decisions to the system. At each point of interaction, both the user and the system were able to re-negotiate any previously made agreement.[27]

As overall satisfaction, the interaction was judged to be satisfying, rather reliable and useful, however, less natural (2.76 Likert points). The latter is largely attributed to rather tedious multimodal generation and avatar performance. System actions were judged by expert annotators as appropriate[28], correct[29] and easy to interpret. Other module-specific performance parameters reflect commonly used metrics derived using reference annotations such as various types of error rates, accuracy, and κ scores measuring agreement between the system performance and human annotations of the evaluation sessions. Recognition and interpretation mistakes turned out to have moderate negative effects on user satisfaction. Table 11 summarizes the results.

Session recordings, system recognition and processing results, as well as the generated feedback were logged and converted to `.anvil` format in order to be able to use the Anvil video analysis tool[30] to view, browse, search, replay and edit negotiation sessions. Anvil allows for automatic generation of some summative feedback about one or multiple sessions. Moreover, applied prediction models can be evaluated by the negotiators and tutors on the fly, and edited and corrected annotated data can be used to retrain the system.

With the satisfaction questionnaire we were also able to evaluate the system's tutoring performance. Participants indicated that system feedback was valuable and supportive. However, they expected more visual real-time feedback and more explicit summative feedback on their learning progress. Most respondents think

---

[26]For now, this is just a general observation; this metric will be taken into consideration in future test-retest experiments.

[27]Performance metrics related to initiative and task substituitivity aspects and their impact on the perceived usability will be an issue for the future research.

[28] A system action is appropriate given the context if it introduces or continues a repair strategy.

[29] A system action is considered as correct if it addresses the user's actions as intended and expected. These actions exclude recovery actions and error handling.

[30]`ww.anvil-software.de`

that the system presents an interesting form of skills training, and would use it as part of their training routine.

7   Limitations and future work

We have presented an approach to dialogue management that integrates a cognitive task agent able to reason about the goals and strategies of human partners, and to successfully engage in a negotiation task. This agent leverages established cognitive theories, namely ACT-R and instance-based learning, to generate plausible, flexible behaviour in this complex setting. We also argued that separate modelling of task related and dialogue control actions is beneficial for current and future dialogue system designs. The implementation introduced a theoretical novelty in instance-based learning for Theory of Mind skills and integrating this in the dialogue management of a tutoring system. The Cognitive Task Agent used instance knowledge not only to determine its own actions, but also to interpret the human user's actions, allowing it to adjust its behaviour to its mental image of the user. This work was successful: human participants who took part in evaluation experiments were not able to discern human users from simulated task agents (see also Stevens et al. (2016b)), and an agent using Theory of Mind prompted users to use that themselves. Our evaluation results suggest that the dialogue system with the integrated cognitive agent technology delivers plausible negotiation behaviour leading to reasonable user acceptance and satisfaction.

The work presented here has certain limitations. Instance templates in the instance-based learning model, slots, values and preferences for both partners were largely pre-programmed, which limits their general applicability. In the future, the agent will learn from real human-human dialogues, e.g. extract negotiation issues and values, and assess their importance. We will also enable the collaborative creation and real-time interactive correction, (re-)training and generation of agents by domain experts and target users. We aim to design authoring tools supporting agent learning and re-training across different situations.

Furthermore, we successfully integrated cognitive, interaction and learning models into a baseline proof-of-concept system. More research is needed on the connections between the cognitive models and the interaction and learning models, and overall mechanisms need to be further specified that underlie communication strategies depending on information about the current state of the task, participant (learning) goals, a participant's affected state, and the interactive situation/environment. Negotiation is more than the exchange of offers, decision making or problem solving; it involves a wide range of aspects related to feelings, emotions, social status, power, and interpersonal relations, context and situation awareness. For instance, tentative cooperative actions can engender a positive reaction and build trust over time, while social barriers can trigger interactive processes that often lead to bad communication, polarization and conflict escalation (Sebenius, 2007). Such dynamics may be observed in negotiations involving participants of different genders, races, or cultures (Nouri et al., 2017). Aspects related to social and interpersonal relations like dominance, power, politeness, emotions and attitudes deserve substantially more attention.

Finally, recent advances in digital technologies open new possibilities for us to interact with our environment, as well as for our environment to interact with us. Everyday artefacts which previously were not aware of the environment at all are turning into smart devices and smart toys with sensing, tracking or alerting capabilities. This offers many new ways for real-time interaction with highly relevant, social and context-aware agents in multimodal multisensory environments which, in turn, enables designing rich immersive interactive experiences. An immersive and highly personalised coaching experience can be achieved by elaborate analysis and effective use of interaction data, applying advanced affective signal processing techniques and rich domain knowledge. A dialogue model that includes a comprehensive account of the user's feelings, motivations, and engagement will form a foundation for a new generation of interactive tutoring systems. A direction that is not yet fully explored is to optimise in a system for the user's feelings, motivation and engagement, as opposed to optimise for pure functional efficiency.

8   Acknowledgments

References

Aleven, V., Mclaren, B., Roll, I., and Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help-seeking with a cognitive tutor. *International Journal of Artificial Intelligence in Education*, 16:101–128.

Allen, J., Ferguson, G., and Stent, A. (2001). An architecture for more realistic conversational systems. In *Proceedings of the 6th International Conference on Intelligent User Interfaces*, pages 1–8. ACM.

Allwood, J. (2000). An activity-based approach to pragmatics. *Abduction, Belief and Context in Dialogue*, pages 47–81.

Allwood, J., Nivre, J., and Ahlsén, E. (1992). On the semantics and pragmatics of linguistic feedback. *Journal of semantics*, 9(1):1–26.

Altmann, E. and Gray, W. (2008). An integrated model of cognitive control in task switching. *Psychological review*, 115(3):602.

Amanova, D., Petukhova, V., and Klakow, D. (2016). Creating annotated dialogue resources: Cross-domain dialogue act classification. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2016)*. ELRA, Paris.

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., et al. (1991). The HCRC Map Task corpus. *Language and speech*, 34(4):351–366.

Anderson, J. (2007). *How Can the Human Mind Occur in the Physical Universe?* New York, NY: Oxford University Press.

Anderson, J., Bothell, D., Byrne, M., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological review*, 111(4):1036.

Anderson, J. and Schooler, L. (1991). Reflections of the environment in memory. *Psychological science*, 2(6):396–408.

Annett, J. and Stanton, N. (2000). *Task analysis*. CRC Press.

Aquilar, F. and Galluccio, M. (2007). *Psychological processes in international negotiations: Theoretical and practical perspectives*. Springer Science & Business Media.

Arslan, B., Taatgen, N., and Verbrugge, R. (2017). Five-year-olds systematic errors in second-order false belief tasks are due to first-order theory of mind strategy selection: A computational modeling study. *Frontiers in Psychology*, 8.

Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., and Fike, A. (2002). MetaTutor: A metacognitive tool for enhancing self-regulated learning. In Pirrone, R., Azevedo, R., and Biswas, G., editors, *Cognitive and Metacognitive Educational Systems: Papers from the AAAI Fall Symposium (FS-09-02)*.

Azevedo, R., Witherspoon, A., Graesser, A., McNamara, D., Chauncey, A., Siler, E., Cai, Z., Rus, V., and Lintean, M. (2009). Metatutor: Analyzing self-regulated learning in a tutoring system for biology. In *AIED*, pages 635–637.

Baker, R., Corbett, A., Koedinger, K., and Roll, I. (2006). Generalizing detection of gaming the system across a tutoring curriculum. In *Intelligent Tutoring Systems: 8th International Conference, ITS 2006, Jhongli, Taiwan, June 26-30, 2006*, volume 4053 of *Lecture Notes in Computer Science*, pages 402–411. Springer.

Bandura, A. (2011). Social cognitive theory. *Handbook of social psychological theories*, 2012:349–373.

Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81.

Beach, L. and Connolly, T. (2005). *The psychology of decision making: People in organizations.* Sage.

Bilange, E. (1991). A task independent oral dialogue model. In *Proceedings of the Fifth Conference of the European Chapter of the Association for Computational Linguistics*, pages 83–88, Berlin, Germany. Association for Computational Linguistics.

Boersma, P. and Weenink, D. (2009). Praat: doing phonetics by computer. computer program. Available at `http://www.praat.org/`.

Bohus, D. and Rudnicky, A. (2003). Ravenclaw: Dialog management using hierarchical task decomposition and an expectation agenda. In *Proceedings of the 8th European Conference on Speech Communication and Technology (EUROSPEECH 2003)*, Geneva, Switzerland.

Borst, J. and Anderson, J. (2015). Using the ACT-R cognitive architecture in combination with fMRI data. In *An introduction to model-based cognitive neuroscience*, pages 339–352. Springer.

Bos, J., Klein, E., Lemon, O., and Oka, T. (2003). DIPPER: description and formalisation of an information-state update dialogue system architecture. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*, pages 115–124.

Bos, J. and Oka, T. (2002). An inference-based approach to dialogue system design. In *Proceedings of the 19th international conference on Computational linguistics-Volume 1*, pages 1–7. Association for Computational Linguistics.

Boser, B., Guyon, I., and Vapnik, V. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM.

Bothell, D. (2004). ACT-R 6.0 Reference Manual, Working Draft.

Bunt, A. and Conati, C. (2003). Probabilistic student modelling to improve exploratory behaviour. *User Modeling and User-Adapted Interaction*, 13(3):269–309.

Bunt, H. (1994). Context and dialogue control. *THINK Quarterly 3(1)*, pages 19–31.

Bunt, H. (1999). Dynamic interpretation and dialogue theory. In Taylor, M., Neel, F., and D., B., editors, *The structure of multimodal dialogue II*, pages 139–166. John Benjamins, Amsterdam.

Bunt, H. (2000). Dialogue pragmatics and context specification. In Bunt, H. and Black, W., editors, *Abduction, Belief and Context in Dialogue: studies in computational pragmatics*, pages 81–105. John Benjamins, Amsterdam.

Bunt, H. (2007). Multifunctionality and multidimensional dialogue act annotation. In *Communication - Action - Meaning, A Festschrift to Jens Allwood. E. Ahlsèn et al. (ed.)*, pages 237–259. Göteburg University Press.

Bunt, H. (2009). The DIT++ taxonomy for functional dialogue markup. In Heylen, H., Pelachaud, C., Catizone, R., and Traum, D., editors, *Proceedings of the AAMAS 2009 Workshop 'Towards a Standard Markup Language for Embodied Dialogue Acts' (EDAML 2009)*, pages 13–25, Budapest.

Bunt, H. (2012). The semantics of feedback. In *Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL 2012)*, pages 118–127.

Bunt, H. (2014a). Annotations that effectively contribute to semantic interpretation. In *Computing Meaning*, volume 4. Springer, Dordrecht.

Bunt, H. (2014b). A context-change semantics for dialogue acts. In Harry Bunt, J. B. and Pulman, S., editors, *Computing Meaning*, volume 4. Springer, Dordrecht.

Cadilhac, A., Asher, N., Benamara, F., and Lascarides, A. (2013). Grounding strategic conversation: Using negotiation dialogues to predict trades in a win-lose game. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 357–368.

Carberry, S. (1990). *Plan recognition in natural language dialogue. ACL-MIT Press Series in Natural Language Processing*. Bradford Books, MIT Press, Cambridge, Massachusetts.

Carletta, J. (2006). Announcing the AMI meeting corpus. *The ELRA Newsletter*, 11(1):3–5.

Chi, M., Siler, S., Jeong, H., Yamauchi, T., and Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science*, 25(4):471–533.

Chin, J., Diehl, V., and Norman, K. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 213–218. ACM.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clark, H. H. and Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of memory and language*, 50(1):62–81.

Cohen, P. and Perrault, R. (1979). Elements of a plan-based theory of speech acts. *Cognitive science*, 3(3):177–212.

Core, M., Lane, C., and Traum, D. (2014). Intelligent tutoring support for learners interacting with virtual humans. In Sottilare, R., Graesser, A., Hu, X., and Goldberg, B., editors, *Design Recommendations for Intelligent Tutoring Systems*, volume 2, pages 249–257. U.S. Army Research Laboratory, Orlando, FL, USA.

Cox, A. (1958). The duty to bargain in good faith. *Harvard Law Review*, pages 1401–1442.

Dahlbaeck, N. and Jonsson, A. (1998). A coding manual for the Linköping dialogue model. Unpublished manuscript.

Danieli, M. and Gerbino, E. (1995). Metrics for evaluating dialogue strategies in a spoken language system. In *Proceedings of the 1995 AAAI spring symposium on Empirical Methods in Discourse Interpretation and Generation*, volume 16, pages 34–39.

De Dreu, C., Weingart, L., and Kwon, S. (2000). Influence of social motives on integrative negotiation: a meta-analytic review and test of two theories.

Dzikovska, M., Moore, J., Steinhauser, N., and Campbell, G. (2011). Exploring user satisfaction in a tutorial dialogue system. In *Proceedings of the 12th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGdial 2011)*, pages 162–172. Association for Computational Linguistics.

Ebhotemhen, E., Petukhova, V., and Klakow, D. (2017). Incremental dialogue act recognition: token- vs chunk-based classification. In *Proceedings of the 18th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, Stockholm, Sweden.

Efstathiou, I. and Lemon, O. (2015). Learning non-cooperative dialogue policies to beat opponent models:the good, the bad and the ugly. *Proceedings of the 19th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL 2015 - goDial)*, page 33.

Fisher, R. and Ury, W. (1981). *Getting to yes: Negotiating agreement without giving in*. Harmondsworth, Middlesex: Penguin.

Frampton, M. and Lemon, O. (2009). Recent research advances in reinforcement learning in spoken dialogue systems. *The Knowledge Engineering Review*, 24(4):375–408.

Fraser, N. (1998). Assessment of interactive systems. In *Handbook of standards and resources for spoken language systems*, pages 564–615. Mouton de Gruyter.

Gama, C. (2004). Metacognition in interactive learning environments: The reflection assistant model. In Lester, J. C., Vicario, R. M., and Paraguacu, F., editors, *Intelligent Tutoring Systems: 7th International Conference, ITS 2004, Maceió, Alagoas, Brazil, August 30 - September 3, 2004*, volume 3220 of *Lecture Notes in Computer Science*, pages 668–677. Springer.

Georgila, K., Nelson, C., and Traum, D. (2014). Single-agent vs. multi-agent techniques for concurrent reinforcement learning of negotiation dialogue policies. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 500–510.

Georgila, K. and Traum, D. (2011). Reinforcement learning of argumentation dialogue policies in negotiation. In *Twelfth Annual Conference of the International Speech Communication Association*.

Georgila, K., Wolters, M., and Moore, J. (2010). Learning dialogue strategies from older and younger simulated users. In *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 103–106. Association for Computational Linguistics.

Gnjatovic, M. and Rösner, D. (2008). *Emotion adaptive dialogue management in human-machine interaction*. Citeseer.

Gonzalez, C. and Lebiere, C. (2005). Instance-based cognitive models of decision-making. In Zizzo, D. and Courakis, A., editors, *Transfer of knowledge in economic decision making*. Macmillan.

Gross, R. (2016). *Psychology: The science of mind and behaviour*. Hodder Education.

Guhe, M. and Lascarides, A. (2014). Persuasion in complex games. *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL 2014 - DialWatt)*, page 62.

Harley, J., Bouchet, F., and Azevedo, R. (2013). Aligning and comparing data on emotions experienced during learning with metatutor. In *International Conference on Artificial Intelligence in Education*, pages 61–70. Springer.

Henderson, J., Lemon, O., and Georgila, K. (2008). Hybrid reinforcement/supervised learning of dialogue policies from fixed data sets. *Computational Linguistics*, 34(4):487–511.

Henderson, M., Thomson, B., and Young, S. (2013). Deep neural network approach for the dialog state tracking challenge. In *Proceedings of the 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGdial 2013)*, pages 467–471.

Henderson, M., Thomson, B., and Young, S. (2014). Robust dialog state tracking using delexicalised recurrent neural networks and unsupervised adaptation. In *Spoken Language Technology Workshop (SLT), 2014 IEEE*, pages 360–365. IEEE.

Hindriks, K., Jonker, C., and Tykhonov, D. (2007). Analysis of negotiation dynamics. In *International Workshop on Cooperative Information Agents*, pages 27–35. Springer.

Hone, K. S. and Graham, R. (2001). Subjective assessment of speech-system interface usability. In *Seventh European Conference on Speech Communication and Technology*.

Ickes, W., Holloway, R., Stinson, L., and Hoodenpyle, T. (2006). Self-monitoring in social interaction: The centrality of self-affect. *Journal of personality*, 74(3):659–684.

Janarthanam, S. and Lemon, O. (2009). A two-tier user simulation model for reinforcement learning of adaptive referring expression generation policies. In *Proceedings of the 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGdial 2009)*, pages 120–123. Association for Computational Linguistics.

Jordan, J. and Roloff, M. (1997). Planning skills and negotiator goal accomplishment the relationship between self-monitoring and plan generation, plan enactment, and plan consequences. *Communication research*, 24(1):31–63.

Keizer, S., Bunt, H., and Petukhova, V. (2011). Multidimensional dialogue management. In van den Bosch, A. and Bouma, G., editors, *IMIX book*. Springer.

Kelley, H. and Stahelski, A. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of personality and social psychology*, 16(1):66.

Kim, S., D'Haro, L. F., Banchs, R., Williams, J., and Henderson, M. (2015). Dialog state tracking challenge 4.

Lafferty, J., McCallum, A., Pereira, F., et al. (2001). Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the 18th International Conference on Machine Learning (ICML 2001)*, pages 282–289, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Lapina, V. and Petukhova, V. (2017). Classification of modal meaning in negotiation dialogues. In *Proceedings of the 13th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-13)*, pages 59–70, Montpellier, France.

Larsson, S. and Traum, D. (2000). Information state and dialogue management in the Trindi dialogue move engine toolkit. *Natural Language Engineering*, 6(3-4):323–340.

Lauria, S., Bugmann, G., Kyriacou, T., Bos, J., and Klein, A. (2001). Training personal robots using natural language instruction. *IEEE Intelligent systems*, 16(5):38–45.

Lax, D. and Sebenius, J. (1992). The manager as negotiator: The negotiators dilemma: Creating and claiming value. *Dispute resolution*, 2:49–62.

Lebiere, C., Wallach, D., and West, R. (2000). A memory-based account of the prisoners dilemma and other 2x2 games. In *Proceedings of International Conference on Cognitive Modeling*, pages 185–193.

Lee, H. S., Betts, S., and Anderson, J. (2015). Learning problem-solving rules as search through a hypothesis space. *Cognitive science*.

Lemon, O., Bracy, A., Gruenstein, A., and Peters, S. (2001). The witas multi-modal dialogue system i. In *Proceedings of the 7th European Conference on Speech Communication and Technology (EUROSPEECH 2001)*.

Lemon, O., Cavedon, L., and Kelly, B. (2003). Managing dialogue interaction: A multi-layered approach. In *Proceedings of the 4th SIGdial workshop on Discourse and Dialogue*.

Lewis, M., Yarats, D., Dauphin, Y. N., Parikh, D., and Batra, D. (2017). Deal or no deal? end-to-end learning for negotiation dialogues. *arXiv preprint arXiv:1706.05125*.

Lim, M. Y., Dias, J., Aylett, R., and Paiva, A. (2012). Creating adaptive affective autonomous npcs. *Autonomous Agents and Multi-Agent Systems*, 24(2):287–311.

Litman, D. and Silliman, S. (2004). Itspoke: An intelligent tutoring spoken dialogue system. In *Demonstration papers at HLT-NAACL 2004*, pages 5–8. Association for Computational Linguistics.

Logan, G. (1988). Toward an instance theory of automatization. *Psychological review*, 95(4):492.

López-Cózar, R., Callejas, Z., and Mctear, M. (2006). Testing the performance of spoken dialogue systems by means of an artificially simulated user. *Artificial Intelligence Review*, 26(4):291–323.

López-Cózar, R., Espejo, G., Callejas, Z., Gutiérrez, A., and Griol, D. (2009). Assessment of spoken dialogue systems by simulating different levels of user cooperativeness. *Methods*, 1(8):9.

Mairesse, F. and Walker, M. (2005). Learning to personalize spoken generation for dialogue systems. In *INTERSPEECH*, pages 1881–1884.

Malchanau, A., Volha, P., Harry, B., and Dietrich, K. (2015). Multidimensional dialogue management for tutoring systems. In *Proceedings of the 7th Language and Technology Conference (LTC 2015)*, Poznan, Poland.

Marewski, J. and Link, D. (2014). Strategy selection: An introduction to the modeling challenge. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1):39–59.

Martin, D., Cheyer, A., and Moran, D. (1999). The open agent architecture: A framework for building distributed software systems. *Applied Artificial Intelligence*, 13(1-2):91–128.

Meijering, B., Van Rijn, H., Taatgen, N., and Verbrugge, R. (2012). What eye movements can tell about theory of mind in a strategic game. *PloS one*, 7(9):e45961.

Möller, S. (2004). *Quality of telephone-based spoken dialogue systems*. Springer Science & Business Media.

Moore, D., Cheng, Y., McGrath, P., and Powell, N. (2005). Collaborative virtual environment technology for people with autism. *Journal of the Hammill Institute on Disabilities*, 20(4):231243.

Morbini, F., DeVault, D., Sagae, K., Gerten, J., Nazarian, A., and Traum, D. (2014). Flores: a forward looking, reward seeking, dialogue manager. In *Natural interaction with robots, knowbots and smartphones*, pages 313–325. Springer.

Mory, E. (2004). Feedback research revisited. *Handbook of research on educational communications and technology*, 2:745–783.

Nasoz, F. and Lisetti, C. (2007). Affective user modeling for adaptive intelligent user interfaces. In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, pages 421–430. Springer.

Nass, C., Jonsson, I.-M., Harris, H., Reaves, B., Endo, J., Brave, S., and Takayama, L. (2005). Improving automotive safety by pairing driver emotion and car voice emotion. In *CHI'05 Extended Abstracts on Human Factors in Computing Systems*, pages 1973–1976. ACM.

Nass, C. and Lee, K. M. (2000). Does computer-generated speech manifest personality? an experimental test of similarity-attraction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 329–336. ACM.

Nielsen, J. (2012). User satisfaction vs. performance metrics. *Nielsen Norman Group*.

Nijboer, M., Borst, J., van Rijn, H., and Taatgen, N. (2016). Contrasting single and multi-component working-memory systems in dual tasking. *Cognitive psychology*, 86:1–26.

Nouri, E., Georgila, K., and Traum, D. (2017). Culture-specific models of negotiation for virtual characters: multi-attribute decision-making based on culture-specific values. *AI & society*, 32(1):51–63.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobreperez, P., Woods, S., Zoll, C., and Hall, L. (2004). Caring for agents and agents that care: Building empathic relations with synthetic agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 194–201. IEEE Computer Society.

Peltason, J. and Wrede, B. (2011). The curious robot as a case-study for comparing dialog systems. *AI magazine*, 32(4):85–99.

Petukhova, V. (2011). *Multidimensional Dialogue Modelling. PhD dissertation*. Tilburg University, The Netherlands.

Petukhova, V., Bunt, H., and Malchanau, A. (2017). Computing negotiation update semantics in multi-issue bargaining dialogues. In *Proceedings of the SemDial 2017 (SaarDial) Workshop on the Semantics and Pragmatics of Dialogue*, Saarbrücken, Germany.

Petukhova, V., Bunt, H., Malchanau, A., and Aruchamy, R. (2015). Experimenting with grounding strategies in dialogue. In *Proceedings of the GoDial 2015 Workshop on the Semantics and Pragmatics of Dialogue*, Goteborg, Sweden.

Petukhova, V., Gropp, M., Klakow, D., Schmidt, A., Eigner, G., Topf, M., Srb, S., Motlicek, P., Potard, B., Dines, J., et al. (2014). The DBOX corpus collection of spoken human-human and human-machine dialogues. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. European Language Resources Association (ELRA).

Petukhova, V., Stevens, C., de Weerd, H., Taatgen, N., Cnossen, F., and Malchanau, A. (2016). Modelling multi-issue bargaining dialogues: Data collection, annotation design and corpus. In *Proceedings 9th International Conference on Language Resources and Evaluation (LREC 2016)*. ELRA, Paris.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain sciences*, 1(04):515–526.

Raiffa, H., Richardson, J., and Metcalfe, D. (2002a). *Negotiation analysis: the science and art of collaborative decision making*. Harvard University Press.

Raiffa, H., Richardson, J., and Metcalfe, D. (2002b). *Negotiation analysis: The science and art of collaborative decision making*. Harvard University Press.

Reinecke, A. (2003). Designing commercial applications with life-like characters. *Lecture notes in computer science*, pages 181–181.

Rich, C. and Sidner, C. (1998). Collagen: A collaboration manager for software interface agents. *User Modeling and User-Adapted Interaction*, 8:3:149–184.

Riedl, M. and Stern, A. (2006). Believable agents and intelligent story adaptation for interactive storytelling. In *International Conference on Technologies for Interactive Digital Storytelling and Entertainment*, pages 1–12. Springer.

Ritter, S., Anderson, J., Koedinger, K., and Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic bulletin & review*, 14(2):249–255.

Roll, I., Aleven, V., McLaren, B., and Koedinger, K. (2007). Can help seeking be tutored? searching for the secret sauce of metacognitive tutoring. In Luckin, R., Koedinger, K., and Greer, J., editors, *Artificial Intelligence in Education: Building Technology Rich Learning Contexts that Work*, volume 158 of *Frontiers in Artificial Intelligence and Applications*, pages 203–210. IOS Press.

Rus, V., Lintean, M., and Azevedo, R. (2009). Automatic detection of student mental models during prior knowledge activation in metatutor. *International Working Group on Educational Data Mining*.

Sadek, D. (1991). Dialogue acts are rational plans. In *Proceedings of the ESCA/ETRW Workshop on the Structure of Multimodal Dialogue*, pages 19–48, Maratea, Italy.

Salvucci, D. and Taatgen, N. (2008). Threaded cognition: an integrated theory of concurrent multitasking. *Psychological review*, 115(1):101.

Salvucci, D. and Taatgen, N. (2010). *The multitasking mind*. Oxford University Press.

Schunk, D. H. (2012). *Learning theories an educational perspective sixth edition*. Pearson.

Sebenius, J. (2007). Negotiation analysis: Between decisions and games. *Advances in Decision Analysis: From Foundations to Applications*, page 469.

Seneff, S., Hurley, E., Lau, R., Pao, C., Schmid, P., and Zue, V. (1998). Galaxy-ii: A reference architecture for conversational system development. In *Proceedings of the 5th International Conference on Spoken Language Processing*.

Serban, I., Sordoni, A., Bengio, Y., Courville, A., and Pineau, J. (2016). Building end-to-end dialogue systems using generative hierarchical neural network models. In *AAAI*, volume 16, pages 3776–3784.

Siegler, R. and Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General*, 127(4):377.

Singh, M., Oualil, Y., and Klakow, D. (2017). Approximated and domain-adapted lstm language models for first-pass decoding in speech recognition. In *Proceedings of the 18th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, Stockholm, Sweden.

Singh, S., Litman, D., Kearns, M., and Walker, M. (2002). Optimizing dialogue management with reinforcement learning: Experiments with the njfun system. *Journal of Artificial Intelligence Research*, 16:105–133.

Smith, L., Pruitt, D., and Carnevale, P. (1982). Matching and mismatching: The effect of own limit, other's toughness, and time pressure on concession rate in negotiation. *Journal of Personality and Social Psychology*, 42(5):876.

Stevens, C., de Weerd, H., Cnossen, F., and Taatgen, N. (2016a). A metacognitive agent for training negotiation skills. In *Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016)*.

Stevens, C., Taatgen, N., and Cnossen, F. (2016b). Instance-based models of metacognition in the Prisoner's Dilemma. *Topics in cognitive science*, 8(1):322–334.

Sutton, R. and Barto, A. (1998). *Reinforcement learning: An introduction*, volume 1(1). MIT press Cambridge.

Tinsley, C., O'Connor, K., and Sullivan, B. (2002). Tough guys finish last: The perils of a distributive reputation. *Organizational Behavior and Human Decision Processes*, 88(2):621–642.

Traum, D., Bos, J., Cooper, R., Larsson, S., Lewin, I., Matheson, C., and Poesio, M. (1999). A model of dialogue moves and information state revision. TRINDI project deliverable D2.1.

Traum, D., Marsella, S., Gratch, J., Lee, J., and Hartholt, A. (2008). Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In *International Workshop on Intelligent Virtual Agents*, pages 117–130. Springer.

Turunen, M., Hakulinen, J., Raiha, K.-J., Salonen, E.-P., Kainulainen, A., and Prusi, P. (2005). An architecture and applications for speech-based accessibility systems. *IBM Systems Journal*, 44(3):485–504.

Van Helvert, J., Petukhova, V., Stevens, C., de Weerd, H., Börner, D., Van Rosmalen, P., Alexandersson, J., and Taatgen, N. (2016). Observing, coaching and reflecting: Metalogue - a multi-modal tutoring system with metacognitive abilities. *EAI Endorsed Transactions on Future Intelligent Educational Environments*, 16(6).

Van Rij, J., Van Rijn, H., and Hendriks, P. (2010). Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language*, 37(3):731–766.

Vanlehn, K. (2006). The behavior of tutoring systems. *International journal of artificial intelligence in education*, 16(3):227–265.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media.

Veksler, V., Myers, C., and Gluck, K. (2012). An integrated model of associative and reinforcement learning. Technical report, AIR FORCE RESEARCH LAB WRIGHT-PATTERSON AFB OH.

Walker, M., Fromer, J., and Narayanan, S. (1998). Learning optimal dialogue strategies: A case study of a spoken dialogue agent for email. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 2*, pages 1345–1351. Association for Computational Linguistics.

Walker, M., Kamm, C., and Litman, D. (2000). Towards developing general models of usability with paradise. *Natural Language Engineering*, 6(3-4):363–377.

Walker, M., Litman, D., Kamm, C., and Abella, A. (1997). PARADISE: A framework for evaluating spoken dialogue agents. In *Proceedings of the 8th conference on European Chapter of the Association for Computational Linguistics*, pages 271–280. Association for Computational Linguistics.

Walton, R. and McKersie, R. (1965). *A behavioral theory of labor negotiations: An analysis of a social interaction system*. Cornell University Press.

Watkins, M. (2003). Analysing complex negotiations. *Harvard Business Review, December*.

Wen, T.-H., Vandyke, D., Mrksic, N., Gasic, M., Rojas-Barahona, L., Su, P.-H., Ultes, S., and Young, S. (2017). A network-based end-to-end trainable task-oriented dialogue system. In *Proceedings of the 15th European Chapter of the Association for Computational Linguistics (EACL 2017)*, Valencia, Spain.

Williams, J., Raux, A., Ramachandran, D., and Black, A. (2013). The dialog state tracking challenge. In *Proceedings of the 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGdial 2013)*, pages 404–413.

Williams, J. and Young, S. (2007). Partially observable markov decision processes for spoken dialog systems. *Computer Speech & Language*, 21(2):393–422.

Xiao, G. and Georgila, K. (2018). A comparison of reinforcement learning methodologies in two-party and three-party negotiation dialogue. In *Proceedings of the 31st International Florida Articial Intelligence Research Society Conference*, pages 217–220.

Xu, W. and Rudnicky, A. (2000). Task-based dialog management using an agenda. In *Proceedings of the ANLP-NAACL 2000 Workshop on Conversational Systems*, pages 42–47.

Young, S. (2000). Probabilistic methods in spoken–dialogue systems. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 358(1769):1389–1402.

Young, S., Gašić, M., Thomson, B., and Williams, J. D. (2013). Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179.

Yu, H.-F., Huang, F.-L., and Lin, C.-J. (2011). Dual coordinate descent methods for logistic regression and maximum entropy models. *Machine Learning*, 85(1-2):41–75.

Zhu, J., Zou, H., Rosset, S., Hastie, T., et al. (2009). Multi-class AdaBoost. *Statistics and its Interface*, 2(3):349–360.