Adaptive Human-Agent Dialogues for Reasoning about Health

by

Jayalakshmi Baskar

Licentiate Thesis, September 2014
Abstract

The aim of this research is to develop new theories, methods and technology, which enables adaptive and personalised dialogues between a human and a software agent, to handle everyday queries about health that are perceived as meaningful and useful to the human. Some of the challenges to build such human-agent dialogue system are the following. The agent needs to have knowledge about the human, the topic of the dialogue, the knowledge domain of the topic, and also about the physical and social environment. Moreover, the agent must know about itself, its role, purpose and limitations. It must know how to be cooperative and be able to behave and express with empathy while conducting a dialogue activity. In some situations, it needs to reason and make decisions about a topic together with the human and about its own behavior. To be able to do this, it needs the capability to evaluate its behavior in the context in which the dialogue takes place. These challenges are addressed by developing formal semantic models to provide the agent with tools to build their knowledge and to be able to reason and make decisions. These models were developed based on literature studies, theories of human activity, argumentation theory, personas and scenarios.

The models were formalised and implemented using Semantic Web technology, and integrated into a human-agent dialogue system. The system was evaluated with a group of therapists and a group of elderly people, who showed curiosity and interest in having dialogues with a software agent on various topics.

The formal models that the agent constructs are adapted to the specific situation and to the human actor participating in a dialogue. They are based on four models: a model with knowledge about the human actor, a model of itself, a domain model, and a dialogue activity model. The dialogue activity is based on argumentation schemes, which function as patterns of reasoning and for the dialogue execution. These models allow the agent and the human actor to conduct flexible and nested sub-dialogues with different purposes within a main dialogue about a topic. The agent can adapt its moves to the human actor’s trail of reasoning, to the human’s priorities and goals, and to some extent behave in an empathic way during the dialogue, and in this way adapt to the human’s emotional state. A method for the agent to be able to evaluate its behavior was also developed and evaluated. The proportion of appropriate moves in relation to the local context of earlier moves in the dialogue was 90% in the pilot study, which indicates that the agent’s strategies for selecting moves can be improved.

Future research will focus on further development of reasoning methods, learning and assessment methods, and interface design. The results will be applied to additional knowledge domains to test its domain independence and will be evaluated with different groups of potential users.
Syftet med forskningen som presenteras i denna avhandling är att utveckla nya metoder, teorier och tekniker som möjliggör dialoger mellan människan och en mjukvaruagent, där dialogerna uppfattas av människan som meningsfulla och till hjälp i vardagsfrågor kring hälsa.

För att agenten ska uppfattas intelligent, social och empatisk, och kunna ha något att bidra med, behöver agenten formaliserad och exekverbar kunskap om människan och hennes omgivning, ämnesområdet för dialogen, om sig själv, sin roll och begränsningar. Dessutom behöver agenten ha kunskap om hur en dialog ska föras, hur resonera tillsammans med människan för att komma fram till beslut, lösningar, ny kunskap och möjliga nya aktiviteter för att människan ska nå sina mål. Agenten behöver också kunskaper om hur bete sig och kunna utvärdera sitt eget beteende.

Formella semantiska modeller utvecklades för att ge agenten verktyg för att bygga sin kunskap och för att kunna resonera och fatta beslut. Dessa modeller baseras på teoretiska modeller för mänsklig aktivitet, personas, scenarios och formell argumentationsteori. Modellerna implementerades i en semantic webstruktur baserad på RDF/OWL, som i sin tur integrerades i ett mänskliga-dialog system. Systemet utvärderades i en pilotstudie med en grupp av terapeuter och en grupp äldre personer. Resultatet visade på ett intresse och nöje med att använda systemet, och det föreslogs kunna vara till nytta även i andra områden.


Framtida forskning kommer att fokuseras på vidareutveckling av resonemangsmetoder, lärande- och utvärderingsmetoder, samt gränssnittsdesign. Resultaten kommer att appliceras på ytterligare kunskapsdomäner och utvärderas med olika grupper av potentiella användare.
Preface

This thesis is based upon the following three papers, which will be referred to in the text by their Roman numerals (Paper I, II & III). The order of the papers reflects the academic progress.


Other publications


A software agent is defined as a computer system that is situated in some environment and that is capable of autonomous action in this environment to meet its design objectives [1]. Software agents have a wide range of applications such as supporting an expert’s decision making, accessing and making use of distributed data sources or the coordination of the execution of assistive technology for healthcare activities [2].

Assistive technology is any product or service designed to enable independence for people. Assistive technology can be supportive, preventive or responsive [3]. Kennedy and co-workers define two types of assistive technology [4]: 1) active assistive technology which continues to process the health-related information during interactive use and may adapt its responses; 2) passive assistive technology which do not process information related to health or behavior change. For example, an interactive system can process user choices or preferences regarding presentation format (eg., video or text) and adapt to these choices during a session. However, this is not active assistance because the responses are not related to the semantic content of the health messages, but only to their formatting. In this work the definition of active assistive technology is applied. The purposes of active assistive technology include the following, targeting the human’s needs: to increase knowledge and ability to perform activities, assist in deciding about actions to make, and promote changes of unhealthy behavior (i.e., in the form of behavior change systems [5, 4]). This thesis focuses on dialogues between a human actor and an active assistive technology in the form of an intelligent software agent.

A dialogue means that two participants exchange verbal messages or so-called speech acts, that take the form of moves in a sequence of exchanges [6]. The goal of a human-agent dialogue system is to communicate with a human, with a coherent structure. To successfully achieve this, the agents need to adapt themselves to the humans interacting with them. However, a number of challenges need to be tackled to build a system that is flexible and adaptive to its environment.
1.1 Research Challenges

Building a dialogue system, that is able to conduct and complete a dialogue with a human, is challenging. The system must be able to participate in a purposeful dialogue that includes ability to reference to items mentioned earlier in that dialogue and also to keep track of different sub-dialogues. It must also know the topic and the main goal of a dialogue to be able to make appropriate moves. The major challenge is to endow these capabilities that allow them to take a role of a pal (friend) that is usually only taken by a human, thus, making an agent less as a tool that people use and more as a peer or pal that people can relate to. As a consequence, hereafter, the concept of human actor is used instead of user. Thus, the whole interaction process is a dynamic one that relies on the capabilities of the agent to learn and adapt to the human actor. We identified the following capabilities as needed for an agent to be able to conduct meaningful dialogues with a human:

1. Knowledge about the human actor,
2. Knowledge about the topic of a dialogue and its domain,
3. Knowledge about the physical and social environment,
4. Knowledge about self, its role, purpose, limitations etc.,
5. Knowing how to conduct the dialogue activity,
6. Knowing how to reason and make decisions about its own behavior and about a topic together with human actor,
7. Knowing how to be social and empathic, i.e., behave,
8. Knowing how to evaluate its own behavior in the context in which it is performed.

To different extent, these challenges have also been described in literature (e.g., [7, 8, 2, 5, 9]). A key asset to address these challenges is knowledge, in a broad sense, which needs to be formalised and executable for an agent to have a dialogue with a human actor. However, some tacit knowledge will always remain in the human actor, which requires that the agent and the human actor cooperate in reasoning and decision-making as well as in improving the agent’s behavior (Figure 1).

1.2 Objectives

This research aims at developing formal models for building the agent’s dialogue and reasoning capabilities, which are based on an activity-centered and holistic view on human-agent interaction. The developed models are aimed to be integrated into a suitable cognitive architecture and implemented in a prototype dialogue system. The objective of the research is to achieve dialogues about health-related topics, which are perceived as meaningful to the human agent.
Figure 1: Level of complexity increases with increasing dependency on the human actor.

1.3 Thesis Outline

The remaining chapters of this thesis are organized as follows. Chapter 2 describes the methodology used. Chapter 3 focuses on representing and sharing knowledge, addressing challenges relating to an agent’s knowledge about self, the human actor, the knowledge domain of the topic of a dialogue, and about the environment (Challenges 1-4). Chapter 4 focuses on the knowledge relating to how to conduct the dialogue activity, to reason, make decisions, how to behave and evaluate its behavior (Challenges 5-8). The contributions of this research are presented in Chapter 5. Finally, some directions of future work are discussed in Chapter 6.
Chapter 2

Methodology

A range of different research methods and theories have been applied in the process of developing theories, models, methods and technology for enabling adaptive human-agent dialogues. These methods are briefly described and motivated in the following sections.

2.1 Literature Review

A literature study was conducted to investigate how earlier research work presents the agent’s role, the purpose of dialogue systems, the knowledge model of the agent, the user, the domain and the dialogue activity, who develops the knowledge model, sources for the knowledge, representation formats and how generic the representation is, and issues regarding the dialogue execution. The literature review provides a base for building the generic conceptual models as a part of this research.

2.2 Persona and Scenario

The semantic models for adaptive human-agent dialogues are designed partly based on personas and scenarios to meet our objectives [10, 11]. A persona functions as a knowledge artefact. In other words, the persona represents a group of users who share common goals, attitudes and behaviors when interacting with a particular product or service [12]. The scenario is a detailed description of how the persona will interact with the intelligent software agent to have a dialogue about a topic [13]. Personas help us remember the target audience by creating a vivid picture of the requirements and help as a guide in the design phase.
2.3 Theories of Human Activity

For understanding and modeling the human actor’s aims, resources and behavior, Activity Theory [14] was used. Activity Theory originated in the work of the Russian psychologist Lev Vygotsky [15] in the early 20th century and has been developed by Kaptelinin among others [16]. Activity Theory is a theoretical framework developed to analyse human’s activity and development of skills within the environment where the activity takes place. Every activity is directed towards a person’s motive which corresponds to a need. The motive is crystallized in an objective, which defines the activity and is the focus of an activity. Activity is always mediated through the use of instruments, or tools, which are not in focus when conducting the activity. In this research the software agent is treated as an actor rather than a tool, since the purpose is to keep the focus on the topic of a dialogue. The instruments for conducting a dialogue are different sources of knowledge and mediating devices. Activity Theory provides a systemic framework of human activity and has been integrated in the resulting formal models of human-agent dialogues.

2.4 Argumentation Theory

As the Activity Theory emphasizes, over a period of time, both the human’s and agent’s knowledge changes as they develop. This changing knowledge results in a conflict with existing information/beliefs and a need to handle the conflicts and potential breakdown situations, in terms of Activity Theory. Argumentation theory was used for handling new information and justification of beliefs [17, 18]. The main idea of argumentation is to structure the reasoning in the way that rules, which support a conclusion, can be defeated when new information arises. This is called non-monotonic reasoning, and reflects how people reason in everyday life. Consequently, the Argumentation Theory was a natural choice for enabling reasoning as a part of dialogues, that may be perceived as “natural” to the human actor.

2.5 Prototyping and Observations in User Studies

A pilot evaluation study was conducted involving a group of five female professionals in occupational therapy and physiotherapy, specialized in the needs of older adults, and a group of eleven older adults, six women and five men. The study was formative, with the results aimed to inform further development. The participants were observed using the prototype system and interviewed. A likert scale was integrated for measuring appropriateness of the agent’s behavior, and evaluated as a potentially useful method for evaluating the agent’s behavior.
2.6 Summary

The described methods and theories enable us to know the state of the art, specify the requirements from the human actor’s perspective, relate a dialogue to the human’s activity, enable reasoning when the dialogue participant’s have conflicting beliefs and to evaluate our results. These methods and theories have been used to build the formal models, partly in collaboration with domain experts such as physicians and therapists, implement the results and evaluate the results with human actors.
Chapter 3

Representing and Sharing Knowledge

Knowledge representation is concerned with designing and implementing ways of representing information so that agents can use it for the purpose of communicating among systems and people with a common representation format, the agents can generate decisions and new knowledge, and bridging distributed and scattered knowledge sources. The types of knowledge we are interested in is typically incomplete, uncertain and ambiguous - thus requires representation formats, that can treat the knowledge in a sound and complete way. Knowledge is typically divided into factual and procedural knowledge [19]. Factual knowledge refers to representations of objects. The procedural knowledge is knowing how to do things to reach a particular objective or goal. An intelligent agent needs reasoning strategies, as well as representations of human activities, to put the human’s actions into a meaningful context, predict and plan activity.

In an intelligent environment which aims at providing adaptive support tailored to the human actor, there are multiple sources of information to consider for an agent to have a dialogue with that human actor [20]. Some of the sources can be information about the human actor’s daily activities observed by an activity recognition system at home, the information about the human actor’s medical health condition obtained from domain professionals and relatives, and the human’s preferences etc. An intelligent interactive behavior also requires social and emotional intelligence, which builds social interoperability.

In the following sections, the representation formats applied in this research for representing and sharing knowledge are introduced. Also approaches for representing the knowledge about the human actor, the agent and the dialogue context are briefly described.
3.1 Representation Formats

A range of knowledge representation formats have evolved for different purposes, e.g., for reasoning over the World Wide Web (WWW). The Semantic Web was initially introduced by Tim Berners-Lee [21] and has evolved over the past 15 years with a growing research community, which explores how knowledge can be embedded, re-used and developed on the WWW in a more intelligent way. The Semantic Web can be viewed as “clouds” of interlinked pieces of information, where sometimes there are links between clouds, but most often they have content that is difficult to process and extract meaning by other systems [22]. The Semantic Web is built using primarily two representation formats: the Resource Description Format (RDF)\(^1\) and the Web Ontology Language (OWL)\(^2\).

RDF is a W3C (World Wide Web Consortium) standard for describing web resources. A web resource is any identifiable information on the web. It helps to ensure that the meaning of a web resource is interpreted as the author/publisher intended. RDF is a graph-based data model with labeled nodes and directed, labeled edges, which makes it a flexible model for representing data. The nodes and edges can be augmented with additional information, and the edges represent the relationship between two resources. Consequently, the structures can be used for reasoning. An example is provided in Figure 2, which shows an instance of an event in the Actor repository developed as a part of this research (Paper I [23]).

The Web Ontology Language (OWL) extends the RDF. OWL is an international standard for encoding and exchanging ontologies (i.e., terminology models) and is designed to support the Semantic Web. OWL is based on Description Logics [24]. Description Logic is a family of logics that are decidable fragments

\(^{1}\)http://www.w3.org/RDF/

\(^{2}\)http://www.w3.org/OWL/
of first-order predicate logic. These logics focus on describing classes and roles, and have a set-theoretic semantics. RDF and OWL are the underlying representation formats for the models developed as a part of this thesis.

FIPA (Foundation for Intelligent Physical Agents) protocol is the most common format for agent’s communication and knowledge sharing. It facilitates rapid development of multiagent systems. Several platforms have been developed that support FIPA messaging but these platforms are restricted to multi-agent systems without a human actor’s involvement. In FIPA-ACL (Asynchronous Communication Language), the agent moves are called communication acts. The FIPA’s communication acts have corresponding moves in our approach, for example, \((\text{query}, \text{if} \equiv \text{ask}), (\text{inform} \equiv \text{tell})\). However, for enabling human-agent interaction using FIPA-ACL, a translation service is required to present the messages through a user-friendly interface, and it does not support argument-based interaction.

Argument Interchange Format (AIF) was developed for visualizing and sharing different arguments and argumentation schemes [25]. The purpose of AIF is to function as a common format for exchanging arguments on the WWW. The AIF model has been modeled as a Semantic Web ontology by Rahwan and colleagues [26]. With this AIF model, arguments may be posted/shared by individuals on the WWW, in a structured format. The ambition with the AIF initiative and the research presented in this thesis is to provide instruments for querying the web and conduct argumentative reasoning with resources, i.e., knowledge, collected from different sources. Therefore, an extended AIF ontology, which is integrated in the ACKTUS\(^4\) platform is used in this work for providing the agent tools for reasoning [27].

### 3.2 Representing the Actor

An actor needs a perception of “self” to be able to relate to other agents. Consequently, a software agent needs explicit models of self and of the human actor to reason about health issues with the human. The agent also needs behavioral knowledge which is generic regardless of the dialogue topic at focus [7].

#### 3.2.1 The Human Actor

Human actors have different knowledge, learning styles, interests, background and preferences regarding information presentation over the Internet. This has paved way to research on intelligent user interfaces that can be designed to recognize the goals and characteristics of the human actor and adapt accordingly. To achieve adaptability, it is important to observe the human actor’s behavior, and make predictions based on those observations. The method of obtaining information pertaining to an individual human from such observations is known

\(^3\)http://www.fipa.org/

\(^4\)ACKTUS- Activity Centered modelling of Knowledge and interaction Tailored to USers, http://acktus.cs.umu.se/
as user modeling (e.g., [28]). A user model, or a simple user profile, may consist of information collected by filling questionnaires, by observing user-actions, or by making inferences. Personalisation aims at providing human actors a system with the content and behavior that they need and desire without necessarily requiring the human actors to specify it explicitly [29].

User models have been used in the healthcare domain for instance, to provide management and retrieval of a person’s or patient’s data [30]. Two techniques available to model an adaptive web system are feature-based modelling and stereotype-based modelling. An example of the stereotype-based model is the User Modelling and Profiling Service (UMPS), which provides the methodology for context dependent personalisation and adaptivity of applications and services in the Amigo environment [31]. Normally, a stereotype model ignores the features and uses the stereotype as a whole. In UPMS the user profiles are built first based on stereotypes and explicit human input, and in the second step these profiles are refined using the interaction/context history.

In this thesis, the feature-based modelling is used to create user models (i.e., human actor model). Feature-based modelling attempts to model specific features such as the individual human actor’s knowledge, interests, goals etc. The majority of modern web systems use the feature-based approach to represent and model information about the human actors. As a part of this thesis, the ACKTUS core ontology was further extended to incorporate information about a user’s environment as presented in Paper I [23], for the purpose of forming a user model, which an agent can use for adapting a dialogue to the human actor. In this model, we extract information related to the chosen topic of a dialogue, combine this with information about the human actor’s motives and preferences as they have assessed this in dialogues about goals and priorities. Based on the local context of a dialogue, a user model is created by the agent, which contains a priority of goals and which gives the agent an instrument for deciding about adaptation. The extended model contains concepts such as space, properties such as action-is-situated-in, to connect the situatedness of human activity with higher level reasoning about purpose and what to do next.

3.2.2 The Software Agent

The characteristics of an intelligent software agent are borrowed from both the domain of cognitive psychology and the domain of distributed intelligent systems [19]. In general, an agent for human-agent dialogues is represented by a cognitive architecture with components for perception, planning, desire, intention, execution and communication [32, 7]. A perception module analyses the data obtained from the environment and available information. It aggregates this information into beliefs. A planning module constructs and outputs abstract plans according to a situation and a desire generator uses these as input. The desire generator interprets an abstract plan as desires; it embodies a representation of current plans. An intention generator refines a desire into intentions. An execution module performs the necessary actions of an intention. A communication module allows the agent to interact; it is used by other
modules such as the perception module to interpret received messages and by
the execution module to send messages.

The most widely used cognitive architecture is the BDI architecture [19]. In
this architecture, the state of the agent is represented by three structures: its
Beliefs, its Desires and its Intentions. The agent’s set of beliefs corresponds
to its knowledge and represents its model of the world. Its desires correspond
to its goals and provide some sort of ordering between states. The intentions
 correspond to the concept of plans, and are the things it has decided to do.
The intentions of a BDI agent may be defined at various levels of abstraction;
for example, an agent may intend to buy a book but may not have decided
which shop to buy it from. A BDI agent gradually refines its intentions to
elementary actions that can be executed. Several cognitive agent architectures
have extended this BDI model to integrate features like planning, learning,
interaction and reasoning.

The canonical model of an agent described by Fox and coworkers [8] is an
extension of the BDI model (e.g., [7]) and provides formal semantics for the
concepts argumentation, decision-making, action, belief, goal (desire) planning
(intention) and learning for agents. Literature describes implementation in sys-
tems, however, mainly for multi-agent reasoning and decision making without
active participation of a human agent in the process [8].

A cognitive architecture of an intelligent agent that can have a dialogue with
a human on health-related topics is presented in [20] (Paper II). The Agent
model represents the behavioral knowledge of an agent. Coach Agent was mod-
eled as a separate project using the ACKTUS platform [33], which builds the
Agent Knowledge repository in Figure 3. As a consequence, the Coach Agent
shares the ACKTUS core ontology with the human actor and the domain mod-
els. The core ontology functions as the common vocabulary in the dialogues.
Moreover, several agents with different characteristics, with some shared models
of behavior, can be created. Each instantiated agent stores its specific learnt
knowledge in a repository, (structured the same way as the Human Actor repos-
itory, the repository for human actor’s information), based on events, e.g., a
dialogue occasion with the human actor.

3.3 Representing Contextual Knowledge

In addition to the human actor’s changing needs, abilities and wishes, the agent
needs to adapt to contextual factors such as the knowledge related to the dia-
logue topic, which typically concerns a particular knowledge domain (e.g., pain,
worries, etc), and the social and physical context, in which the dialogue takes
place.

3.3.1 Epistemic Context

The epistemic contextual knowledge of human-agent dialogues applied in this
thesis work is related to health, e.g., pain, worries, wellbeing, gait, cognition
etc. A dialogue topic is framed by a representation of the knowledge domain, which defines the reasoning context and associated questions, rules, knowledge sources etc. The knowledge is modeled by professionals who are experts in relevant domains, by using the ACKTUS platform [27]. The factual knowledge is modeled using an extended ICF(International Classification of Functioning, Disability and Health)\(^5\) concept model and procedural knowledge is represented using the extended AIF model.

### 3.3.2 Physical and Social Context

The knowledge about the human actor's activities deals with what is being done, its purpose and how a particular activity is performed. The user may describe explicitly what activity is being performed and their environment, while the latter involves observation of both environment and activity, e.g., by the use of sensors as a part of a ubiquitous computing environment.

In dialogues between a software agent and a human actor about health-related topics, it is essential to reason about daily activities to situate a health-related topic in a context. The identification of activity needs to be related to level of importance to the user, user’s preferences and goals, to reason about how the system should act upon the new knowledge, e.g., guidance in activity execution etc. Activity theory is useful for categorizing and interpreting observations to represent knowledge about purposeful activities [16]. The egocentric

\(^5\)http://www.who.int/classifications/icf/en/
conceptual model covers primarily concepts related to sensors, effectuators, interaction devices and their properties and their relation to the situative spaces [34]. As a part of our work, the ACKTUS core ontology was extended with models of activities of daily living, situated in a physical or virtual environment for the purpose to build more holistic models of the human actor and her activities. This was done by fusing the activity-theoretical model, the conceptual model of egocentric interaction and the results of hierarchical task analyses of selected human activities (presented in Paper I [23]).
Chapter 4

Conducting the Dialogue Activity

The purpose of a dialogue can be to allow the human actor to propose a claim about a topic and to participate in an elaborated discussion about the topic, with well-founded arguments in favor and against the claim. However, this is not always the case, since the coherence of a dialogue depends on its goal it is important to identify the various types of dialogues. Walton and Krabbe [6] have proposed six main categories of human dialogue that are based on the overall goal of the respective dialogue. The categories integrated in the dialogue activity model developed as a part of this work are the following:

- **Information-seeking dialogues**, where one participant seeks the answer to some question from another participant. The participant’s aim is to gain knowledge or pass on knowledge to the other participant without having to reason to come to a conclusion about the topic of the question. The clinical interview as conducted in the healthcare domain is an example, where information is collected, however, the reasoning about e.g., diagnosis, or interventions may be done by other persons.

- **Inquiry dialogues** occur when the participants collaborate to search for an answer to some question and to validate a claim about a topic [35, 36]. Reasoning takes place, where the agents utilize their respective knowledge, which can be different and incomplete, and new knowledge is derived in cooperation.

- In **Deliberation dialogues** agents collaborate to decide what course of action should be adopted in some situation [37, 38]. In the healthcare domain, the decisions about interventions aimed at improving a person’s daily life and medical condition are typical examples.

- **Persuasion dialogues** involve one agent seeking to persuade another to endorse a statement that this agent currently do not [39, 40]. Initially,
agents involved in persuasion dialogues will have conflicting opinions about e.g., the importance of consulting the nurse about a pain condition, or when a changed behavior is judged to be beneficial for a person’s health, e.g., stop smoking, eat healthier, etc. [27]

This thesis will primarily focus on exploring the application of the three categories suitable for cooperative behavior: information-seeking, inquiry and deliberation dialogues. In addition, persuasive dialogues are included since one of the main obstacles to improving health is changing behavior.

We interpret and formalize the dialogue activity into a hierarchy of tasks with different goals and semantics, which builds up a model of the dialogue activity (Figure 4). In the dialogue activity model presented in Paper III [41], we identified the following generic set of actions to serve the overall motive of a dialogue, which is addressing the topic of the dialogue: *information-seeking, inquiry, deliberative, and persuade*, related to the common goals of different types of dialogues as described previously: seek information, construct/build new knowledge, decide upon actions to be made and lead the other agent to change opinion. To these actions, we added the actions *organize dialogue, decide* and *support dialogues*. Using this dialogue activity model, the nested dialogues about topics related to the main topic can be accomplished in our framework.

Some human-agent dialogues involve reasoning and decision-making aimed at reaching conclusions about new knowledge or action to take. Some other dialogues aim to collect information or pass the information, or provide emotional support to the human actor without reasoning. In the following sections, the two purposes are further described.
4.1 Reasoning, Generating New Information and Making Decisions

The software agents use their formal representations, or models of the world (i.e., knowledge) for reasoning, decision making and action selection. Rule-based approaches are the most commonly used approaches for reasoning about health-related conditions [42]. Typically, rules are built using classical monotonic logics, where premises and conclusions cannot be changed and adding new information is not allowed. However, humans possess the ability to identify the information that is appropriate to the context of the situation and change reasoning as new information appears. The classical logics require a fixed set of facts that are either true or false (called closed-world assumption). These fixed set of facts are used to make sound inferences to come to valid conclusions. Consequently, classical rule-based systems are not well suited for cooperative (plausible) reasoning with humans.

Given a situation with conflicting information, an agent is faced with the problem of deciding what it could reasonably believe. As advocated in various non-monotonic inference formalisms such as default logic [43], it is often possible to identify multiple reasonable positions. This idea has been adopted in abstract argumentation theory [44], which provides a formal framework for handling conflicting positions, potentially due to uncertain or inconsistent knowledge. This theory views logical derivations as abstract arguments (nodes in a graph), and conflicts as defeat relations (directed arcs) over these arguments without considering the content of a node or an argument.

Argumentation-based dialogues have been widely studied as a framework to support agreement among agent dialogues (e.g., [45, 35, 39, 46]). The arguments can be instances of argumentation schemes. The argumentation schemes are patterns of reasoning and regulate how a particular type of dialogue can be conducted. The argumentation schemes are also useful in finding missing premises, analysing arguments and evaluating them. These schemes represent defeasible inferences that are useful in reasoning about a plausible hypothesis under conditions of uncertainty and lack of knowledge. Several argumentation schemes have been defined and applied in various research studies [47, 48, 49]. As a part of this work existing schemes have been adapted and integrated, as well as new developed schemes in a dialogue activity model (Paper III [41], Figure 3).

McBurney and Parsons give an overview of dialogue games for argumentation in [45]. A dialogue between two participants is seen as a game in which each individual has objectives and a set of legal moves which can be used to obtain those objectives. In this game-theoretic approach to dialogues the speech acts are viewed as moves in a game and semantics indicating whether moves are appropriate at a specified time and are formulated as rules of the game [35]. An inquiry dialogue system has been presented by Black and Hunter [35]. Their system uses the game-theoretic approach and restricts each dialogue to one single topic, and a small set of moves (open, assert and close).
Since we need a more complex adoption of dialogues with different topics for accomplishing a dialogue that is perceived as natural by the human, we extended the set of moves to include the following moves as valid actions for the agent to take: open, close, pause, resume, ask, assert, affirm, inform, remind, and alert (Figure 4).

The content of the dialogue type support dialogue introduced in Paper III [41], is typically the outcome of an earlier conducted deliberation dialogue, where the human agent and the Coach Agent have agreed upon a plan of actions to be conducted. The actions performed as support dialogues are one of the following: inform, i.e., provide the human agent with information or advice, remind the person of actions to make, and alert the person when important things need to be done. A remind and alert move are formal arguments, which contain the information about what is to be done (a claim), together with the motivations (the grounds), which support the claim (Paper III [41]).

4.2 Adaptive and Social Behavior

In a human-agent dialogue, an agent must adapt to the human actor. We identified three types of adaptation needed for an agent to exhibit adaptive and social behavior. Firstly, an agent needs to adapt to the human actor’s line of thinking. This is accomplished through the dialogue activity model in combination with a domain model, which allows flexibility in conducting a dialogue (Figure 3). Secondly, the agent needs to adapt to the human actor’s needs, preferences and ability. This accomplished through building a model of the human actor (actor model), which is used in the selection of topics and contents of dialogues e.g., tailored advices. Finally, the agent must also adapt to the human actor’s emotional states and norms, which puts additional demand on situatedness and evaluation of the agent behavior.

Norms are an important part of human social systems, governing many aspects of cooperative decision-making [50]. The study of norm emergence, compliance, and adoption has resulted in new architectures and standards for social agents. Norms play an important role in determining the behavior of people in society. The norms have been used as a computational mechanism for creating coordinated action within multi-agent systems [51, 52]. Previous work on modeling norm life-cycles can be organized into two categories: internal and external. In the first category, norms are characterized as arising from internal mental processes that can be specified using cognitive modeling techniques, and social behavior is viewed as the outcome of internalizing external preferences. The agents are able to acquire new norms, rather than relying on preexisting constructs, and can reason and value about norm compliance. In the second category, the focus is on social interactions, and game-theoretic models are used to quantify the bottom-up process of recognizing and complying with norms in the external social system. Convergence occurs when agents arrive at a mutually agreed upon utility maximization strategy. A limitation of this type of system is that the agents lack a sense of normative expectation and do not distinguish
between a strategy and a social norm [53].

Bickmore and co-workers [54] state that the establishment of norms is crucial in domains in which a human actor is attempting to undergo a change in behavior. They designed relational agents that build and maintain long-term, social-emotional relationships with their human actors [5].

In this thesis, for the Coach Agent to be able to conduct dialogues similar to natural dialogues, it was necessary to include generic behavioral knowledge. The behavioral knowledge needs to be generic regardless of which type of dialogue or topic is at focus. The initial version of the agent model developed as a part of this thesis work has been enriched with structures, which function as tools for the Coach Agent to use in the adaptation of its behavior. The Coach Agent associates each action with a concept that is common for both the human and the agents such as starting, sustaining and ending a conversation. These concepts were used to define certain behaviors that are typically appropriate in a phase of a dialogue, for example, "hi" is associated to concept "starting a conversation". Similarly, for the dialogue moves, such as affirm, the agent emphasizes and affirms the emotional state, which the human actor expresses, by using empathic statements such as "does not sound good" and "I see". This enables the Coach Agent to simulate the behavior of a participant who listens with empathy.
Chapter 5

Contributions

Paper I

The contributions of Paper I [23] include the extension of the ACKTUS core ontology with the user and domain models for incorporating situated models for environments and activities. This was done based on theoretical analysis, and comparisons with existing formal and informal models of human activity. As a consequence, the *event ontology* building the Actor repository was created as a supplement to the actor and activity models. Based on these models, the agent constructs its knowledge about the human actor, the topic of a dialogue and the environment, which corresponds to the Challenges 1-3 described in Section 1.1. These ontologies formed the basis of an initial implementation of a human-agent dialogue system, with a focus on the information-seeking type of dialogues using the JADE\(^1\) platform. The agent architecture enforced by JADE was found to be a limitation for more flexible human-agent interaction, which is required for other types of dialogues.

Paper II

The contribution of Paper II is the design and partial implementation of the architecture of a cognitive agent potentially able to engage in a dialogue with a human actor. The architecture consists of different components including the dialogue adaptation strategies and reasoning strategies, in addition to the models presented in Paper I. This paper focuses in particular on the dialogue manager, building on the idea of schemes. A basic strategy for adaptation to the human actor’s line of thinking, goals and priorities was formed.

\(^1\)Java Agent Development Environment, http://jade.tilab.com/
Paper III

Paper III supplements Paper I and Paper II, addressing the agent’s knowledge about self, its role, purposes and limitations, by taking a holistic and activity-centered perspective on the human-agent dialogue activity. The contributions of this paper include an agent model, a dialogue activity model and the integration of the human actor model, the agent model, the dialogue activity model and the domain model for building the belief base of the agent (see challenges 4, 5, 7 and 8, described in Section 1.1). The human-agent dialogue system introduced in Paper II was further developed based on these semantic models, and integrated in the web-based support application Vardagsvis (I-Help), designed for older adults. A pilot evaluation study of the prototype system was done with five therapists and eleven older adults to investigate the overall idea of a dialogue system for supporting everyday issues, how a sense-making dialogue would unfold and what topics would be interesting to elaborate upon.

In general, the available dialogue topics were considered interesting and relevant. The participants suggested additional health-related topics for human-agent dialogues relating to eating habits and eating disorders. Enthusiasm and curiosity was expressed about the idea and the wish to use the dialogue system merely for the fun of it: "...this is fun, let me try another one!"

A purpose was also to evaluate the appropriateness of each dialogue move made by the agent. A method for the agent to be able to evaluate its behavior was developed and evaluated. Two types of appropriateness were identified and evaluated: context-related appropriateness and topic-related appropriateness. The context-related appropriateness relates the agent’s move to the context in the dialogue line in which it occurs (i.e., measures the agent’s adaptivity). The proportion of appropriate moves in relation to the local context of earlier moves in the dialogue line was 90% in the pilot study, which indicates that the agent’s strategies for selecting moves can be improved. The topic-related appropriateness relates the agent’s move to the knowledge domain (i.e., evaluates the knowledge domain). The approach was found to add significant value to the evaluation study, without disturbing the participant from participating in the dialogue.
Chapter 6

Future Work

Several directions of research are considered as part of future work. The main topics relate to continued development of the modules in the cognitive agent architecture. One obvious line of future work is to develop reasoning and adaptation strategies for the agent to improve its ability to adapt, to the individual and to the situatedness of context during a dialogue. Another interesting line of future work is to build a learning algorithm for the agent to plan next dialogue move based on the human actor’s feedback data collected by the evaluation method developed in Paper III.

The additional topics for future work relate to how the dialogue can be mediated to the human actor through purposefully designed user interfaces, and user environments. The potential mediation of dialogues with human actors using avatars and humanoid service robots will be investigated, as well as using modalities for communication other than the text-based prototypes. A full-scale prototype will be built and embedded with other applications to test its domain independence. The adaptivity, agent behaviors and knowledge context will be evaluated with different user groups of the human-agent dialogue system.
Bibliography


