Hypothesis-Driven Agent Dialogues for Dementia Assessment

Chunli Yan, Helena Lindgren
{chunli, and helena}@cs.umu.se

Department of Computing Science, Umeå University, SE-901 87 Umeå, Sweden

Abstract. The purpose of our work is to support learning in human agents by providing agent-based dialogues tailored to the health professional’s focus and shifts of focus. A hypothesis-driven inquiry dialogue system were developed and implemented in a decision-support system for dementia diagnosis, visualizing the dialogue line and resulting arguments. The software agents utilize knowledge repositories and structures of ACKTUS, which is a Semantic Web-based platform for developing knowledge-based support applications. The results from an initial evaluation study involving domain expert physicians indicate that the approach is highly interesting to and useful for physicians for supporting decision making.

1 Introduction

In order to explore the potentials in supporting early diagnosis of dementia, the Dementia Management and Support System (DMSS-R) was developed [1]. DMSS-R is developed for physicians who may not often meet dementia patients in daily work. The system provides the user an overview of potential diagnoses and their strength based on a set of clinical guidelines and diagnostic criteria in the domain. The system follows a diagnostic reasoning procedure. However, sometimes it is not easy for less skilled physicians and physicians in stressful situations to make a decision [2]. They may fall into a causal reasoning procedure for the purpose to verify hypotheses. We explore in this paper argumentation-based agent dialogues [3] for providing the physician support also in this situation.

The contribution of this paper is the supplementary functionality in the form of a multi-agent dialogue system based on formal argumentation for the purpose to allow the user to select a hypothetical diagnosis and make the system elaborate the support and contradictory arguments for the hypothesis. The arguments are based on information about a patient and a set of clinical diagnostic guidelines.

In the medical domain, it is common that data is stored in different databases. However, the different information are often needed in order to make a well-founded decision. Merging the data from the different databases into one may not be desired or practical, partly because of the following reasons: 1) the size of the database will be otherwise too large and would affect the speed for querying or modifying data; 2) the knowledge used for each reasoning procedure only occupies a very small part of the whole database; and 3) the data in the different
databases may be inconsistent. In our system, we simulate two agents that have different knowledge bases, based on different sources of data. Because of the reasons above they need to jointly collaborate with each other to make a decision regarding a topic.

A prototype implementation of agent-based diagnostic dialogues integrated as a part of DMSS-W, the web-version of DMSS-R, was developed and evaluated in a pilot study with potential end users. Our approach is to extend and implement the inquiry dialogue systems presented in [4] and adjust the rules for dialogue execution by allowing the agents to open warrant inquiry dialogues nested within argument inquiry dialogues.

2 Methods, Material and Subjects

DMSS-W is developed using the ACKTUS platform, which domain expert physicians can use for modifying the knowledge content and interaction design [5, 6]. The multi-agent system (MAS) was developed under JADE (Java Agent Development Framework) platform. JADE was selected because it is the best-known and most widely used [7] among platforms that integrates the FIPA standard. The structure with the MAS, DMSS-W and ACKTUS is shown in Figure 1.

We conducted a pilot evaluation study for the purpose to receive primarily an initial response to our dialogue approach and secondarily on the rest of DMSS-W and the content modifiable by using ACKTUS. Key domain experts were asked to participate, based on their roles as developers of dementia care at national level and dementia research at the forefront internationally, partly as authors of one of the consensus guidelines exemplified in our user study. Three physicians participated who are leading work on developing national guidelines for dementia care, a quality registry, research consortiums and more. Moreover, an occupational therapist participated who was responsible for developing the care for dementia patients in a county council. Since it is a small sample, we only give a summary of their comments, and refer details to future evaluation studies when a larger and more diverse group of participants are involved.

3 Reasoning Support by Hypothesis-Driven Agent Dialogues

The main purpose of DMSS-W is to provide decision support about an individual patient to the physician. When a physician meets and investigates a patient, the patient’s symptoms are entered into DMSS-W. These symptoms will be used as defeasible facts in the following reasoning process. DMSS-W is built upon a knowledge base named ACKTUS-dementia. In this RDF-repository, there are knowledge about dementia diseases captured in an ontology. From this knowledge repository defeasible rules can be extracted into different formats, e.g., If the patient has (not) the symptom1 and symptom2, then the patient has disease A.

For implementing the agent-based dialogues, we adapt the approach developed by Black and Hunter [4], with some essential extensions, partly introduced
Fig. 1. Information flow between the systems.

Fig. 2. DMSS-W showing a part of a dialogue between PA and DA.
in [8]. We construct two agents: a Physician Agent (PA) and a Domain Agent (DA). Both agents have the aspiration of gaining new knowledge, which is accomplished cooperatively through inquiry dialogues [9]. PA is used to simulate a physician who has little experience from dementia assessments, and DA to simulate a domain expert. DA builds its belief base on knowledge retrieved from the ACKTUS-dementia repository. PA builds its belief base on the knowledge stored in the physician’s user profile and on information about a patient stored in a patient case repository. In addition, patient data entered into the system is used as PA’s beliefs (Figure 1).

The human actor selects a hypothetical diagnosis as the topic for the dialogue from a drop down list, PA receives this topic and initiates a dialogue with DA about this topic (Figure 2). So PA has an extra function to interact between the MAS and DMSS-W (Figure 1). After the dialogue ended, the PA sends the dialogue moves back to DMSS-W, and DMSS-W translates the dialogue moves to sentences that human can understand and displays them to the user (Figures 2 and 3).

![Fig. 3. The resulting arguments generated by a dialogue in our example, based on two conflicting diagnostic criteria (Source A and B).](image)

### 3.1 Building Inquiry Dialogues and Arguments

Black and Hunter proposed a method to generate inquiry dialogues [4]. There are two subtypes of inquiry dialogues in their system: warrant inquiry (wi) dialogues and argument inquiry (ai) dialogues. wi dialogues allow two agents to share knowledge to jointly construct dialectical argumentation trees; ai dialogues allow two agents to share knowledge to jointly construct arguments to be fed into the dialectical tree. ai dialogues can be nested within a wi dialogue, however a wi dialogue cannot be nested within an ai dialogue in their framework.

We apply the two types of inquiry dialogues in our approach [4]. The topic of an ai dialogue is a defeasible rule. During an ai dialogue, two agents jointly
Table 1. Algorithm for dialogue implementation.

1 **Initiating Agent:**
Sends the topic of the dialogue.
The patient case and the topic of the dialogue is identified, initiating each agents belief bases (PBQ).

2 **Agent:**
Receive and Check MessageType
If RecMessageType==⟨open wi⟩
   Update its PBQ
   CheckPBQ();
If RecMessageType==⟨close wi⟩
   If this agents previous move was also close wi, it means this wi dialogue is now terminated.
   Update RS
   Return to top ai dialogue
   If there is no top ai dialogue, the whole dialogue terminates
   Else CheckQS();
If RecMessageType==⟨open ai⟩
   CheckQS();
If RecMessageType==⟨close ai⟩
   If this agents previous move was also close ai, it means this ai dialogue is now terminated.
   Update RS
   CheckPBQ();
   Else CheckQS();
If RecMessageType==⟨assert wi⟩
   CheckPBQ();
If RecMessageType==⟨assert ai⟩
   CheckQS();

*Note 1.* RS means result store, QS means query store and RS and QS are shared between the two agents. PBQ means possible beliefs queue, and each agent has its own PBQ. The functions CheckPBQ() and CheckQS() are found in Table 2.
generate a valid argument, however, if at least one premise of the rule can not be proved to be true, the dialogue gives no result. The topic of a wi dialogue is a defeasible fact. Within a wi dialogue with the topic $\alpha$, several arguments may be generated with the derivation of each argument is $\alpha$ or $\neg \alpha$. A result ($\alpha$ or $\neg \alpha$ is true or unknown) will be given when closing the wi dialogue. If more than one argument is generated and they are in conflict, we use a strategy to solve this problem and decide which one ($\alpha$ or $\neg \alpha$) will win, which is presented in the next section.

Table 2. Algorithms for selecting content and type of the next locution.

<table>
<thead>
<tr>
<th>Function CheckQS($QS$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>If $QS$ is null,</td>
</tr>
<tr>
<td>send close ai dialogue</td>
</tr>
<tr>
<td>Else</td>
</tr>
<tr>
<td>Poll first premise.</td>
</tr>
<tr>
<td>Check RS based on this premiseID</td>
</tr>
<tr>
<td>If it is null,</td>
</tr>
<tr>
<td>Send open wi dialogue.</td>
</tr>
<tr>
<td>Update its PBQ.</td>
</tr>
<tr>
<td>Else</td>
</tr>
<tr>
<td>Send assert ai premiseID.</td>
</tr>
<tr>
<td>If this premiseIDs result is false or unknown,</td>
</tr>
<tr>
<td>Empty QS.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Function CheckPBQ($PBQ_a$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>If $PBQ_a$ is null,</td>
</tr>
<tr>
<td>send close wi dialogue</td>
</tr>
<tr>
<td>Else</td>
</tr>
<tr>
<td>Poll first belief</td>
</tr>
<tr>
<td>If it is a defeasible fact,</td>
</tr>
<tr>
<td>Send assert wi fact</td>
</tr>
<tr>
<td>Else</td>
</tr>
<tr>
<td>Send open ai dialogue</td>
</tr>
<tr>
<td>Update QS.</td>
</tr>
</tbody>
</table>

Note 2. RS means result store, QS means query store and RS and QS are shared between the two agents. PBQ means possible beliefs queue, and each agent has its own PBQ.

In practice, we let the agents investigate each premise of a domain belief when an argument inquiry dialogue has been opened, by opening a new warrant inquiry dialogue with the premise as the topic. This is a major difference in our approach compared to [4]. The two kinds of dialogues are nested within each other alternately, which means that an ai dialogue is always nested within a
wi dialogue and a wi dialogue except the top one is always nested within an ai
dialogue.

The dialogues are implemented using the algorithm presented in Table 1
and the algorithms for selecting locution and content of a locution described in
Table 2. The algorithms are generic in order to be applicable in a situation when
more than two agents participate in the dialogue. Participating agents have a
predefined order to send their messages, according to their turn and the dialogue
is concluded and terminated after all of them have made a close move.

3.2 Argument Evaluation

In our system, there are several types of knowledge sources and each knowl-
dge source has a preference level, which corresponds to a confidence level. A
preference level is a natural number that the smaller the number is, the higher
confidence level the knowledge source has. Each rule is related to a knowledge
source, which gives the rule the same preference level as its knowledge source.
Consequently, arguments are evaluated based on their preference levels, which
are based on the confidence in the underlying clinical knowledge. This approach
follows how evidence-based knowledge is evaluated in health care practice. This
way we can also include knowledge with less confidence such when a single expert
physician contributes with rules-of-thumb, not (yet) validated through evidence-
based studies. If the knowledge source is a guideline that every physician should
follow according to authorities, it will have higher confidence level (smaller pref-
rence level). The same condition holds with facts. If a fact is obtained from
applying validated assessment instruments (e.g., for memory functions), it is
more objective and it has smaller preference level. In contrary, if it comes from
what the patient tells about e.g., his memory difficulties, this fact has a bigger
preference level since the patient may not remember details about his problems.
Arguments are generated from a set of facts and rules, so each argument has a
preference level according to the facts and rules it uses. Our strategy to solve
the conflicted argument is the following: argument with smaller preference level
will defeat the one with bigger preference level. If they have same preference
level, we count the number of how many arguments support the topic and how
many arguments attack the topic. The one with more supporters will remain
undefeated. If the number of supporters and attackers are the same, no one will
win and the wi dialogue result will be unknown.

We show a very simple example about how two agents jointly build arguments
and get the result in Figure 4. Suppose belief base of Agent1 is: (¬a, 4), (a → c, 2)
and belief base of Agent2 is: (b, 3), (b → a, 3). Agent1 wants to know c is true
or false, so it opens a wi dialogue with c as its topic. The dialogue proceeds
as in Figure 4. From this example, we can see there is a conflict about a with
two arguments ((¬a, 4), (¬a, 4)) and ((b, 3), (b → a, 3), (a, 3)). Since the second
argument has smaller preference level than the first one, the result is a is true.
4 Evaluation

Performance Results  A comparison was made by implementing Example 2 in [4] using our approach (Figure 5). While their approach generated 23 moves, our approach generated 37 moves. However, using our approach, agents reached a consensus about the topic when terminating the dialogue (the 37th move in this example). Using their approach, only some arguments are generated. Furthermore, nesting warrant inquiry dialogues was found necessary also for translating the dialogues into a format executable by a computer and achieving an efficient implementation of the dialogues.

Improving Human Readibility  In order to improve readability for a human agent, we applied some methods for reducing the number of lines in the visualization of a dialogue, and mapped the locutions into statements in natural language as follows.

For evaluation purposes the dialogue line generated in our inquiry dialogue implementation was interpreted into natural language sentences, which were expected to be perceived more natural and readable by a human agent. A simple mapping of locutions was done and presented to the human agent, as shown in the Figures 2 and 3. Since the agents are collaborative, and for saving space in our illustration, the two close moves for dialogues are visualized as one common agreeing expression of the outcome of the (sub-)dialogue.

Now we give the detailed description for Figure 2 and 3. All figures related to this example are screenshots from DMSS-W or ACKTUS. Here, PA retrieves the patient information (facts), and the topic selected by the physician from
Fig. 5. Example 2 in Black and Hunter in table on the left, and same example implemented using our approach in the table to the right.

Fig. 6. The facts (patient’s symptoms).

Fig. 7. An example of the s-nodes used for creating rules.
the DMSS-W user interface. The selected topic is *MCI is present*. The facts are shown in Figure 6 and the two agents have their own rules, retrieved from ACKTUS-dementia. Totally there are 31 rules used in this dialogue for building arguments. An example of the s-nodes used for creating rules is shown in Figure 7. The rule based on the information in the figure is retrieved by the DA. Only when all premises of a rule are fulfilled, the conclusion can be reached. When PA makes an open wi move with *MCI is present* as its topic, the dialogue between the two agents starts. Then the two agents check if they have relevant rules or facts. Since DA has four rules related to this, it opens four ai dialogues in turn. Within each ai dialogue, there are some sub-wi dialogues with the premises being the topics of the dialogues. Within each wi dialogue, there may be some sub-ai dialogues and some assert moves. The *assert move* is used when an agent can contribute with a fact, or can present a valid argument, e.g., whether to believe that the topic of the wi dialogue is true or false. When all the rules and facts related to *MCI is present/absent* are considered, the top dialogue (with *MCI is present* as its topic) terminates and the final result (the topic is true/ false/ unknown) is obtained.

This example is provided in Figure 2 and 3 targeting the presence of Mild Cognitive Impairment (MCI), a dialogue taking place in our scenario between the physician and the domain agent in the physician’s use of DMSS-W. In this dialogue also the presence of dementia needs to be assessed as a sub-topic and related cognitive deficiencies. Consequently, the list showing the dialogue line becomes very long. Therefore, an improved visualization of the dialogue where the sub-topics are more distinguishable was implemented before evaluating the dialogues with physicians.

**Pilot Evaluation Study with End Users** The dialogue functionality was seen as very interesting by the physicians. Especially that one can get an overview of conflicting views based on the different guidelines (Figure 3). One of the physicians envisioned a larger number of conflicting arguments than what was showing in our example in some cases and wondered how this could be handled. Moreover, she also wondered how to integrate the more subtle reasons a physician can have based on observations. She was satisfied with the functionality provided by ACKTUS, where the sources of the different conflicting views can be managed. She also thought that the idea of having the possibility to enter her reasons for deviating from clinical guidelines, which can be accomplished by using the knowledge structures of ACKTUS, was very interesting.

The physicians also expressed that DMSS-W, and especially the agent part, can fill an education function for the less experienced physicians.

5 Related Work

Formal argumentation frameworks and argumentation schemes have the potential to improve the reasoning and communication capabilities within and between software agents (e.g., [10]).
There are several proposals of formal frameworks for capturing the factors influencing an argumentation-based dialogue, and for resolving which arguments are valid and which are defeated. The most influential is the Dung semantics of argumentation frameworks [11], which have been extended and applied in many example cases, implementations of reasoners and argumentation software for different purposes (e.g., [12, 13]). Another example is the formal framework presented by Black and Atkinson [14] combining inquiry dialogue argumentation and practical reasoning (i.e., reasoning about actions). The combination of these approaches makes the reasoning easier among the agents with different perspectives on a subject. They have taken a formal specification of the two mentioned types of dialogues, extended it with additional critical questions, which make it possible to make more arguments while trying to find some consensus. The critical questions are generic, which is a difference from our work where the knowledge and associated questions are mainly domain-specific. However, it will be investigated in future work in what way our agent protocols conform to their generic framework.

6 Conclusions

In order to provide support to a physician in their reasoning about diagnoses, a hypothesis-driven agent dialogue system was developed. The prototype system DMSS-W, which is being developed for supporting physicians in dementia diagnostics, was extended with a dialogue space where inquiry dialogues between agents are visualized. The human actor can select the topic for a dialogue, and initiate a simulation of a dialogue between a human actor and a software agent. Patient data entered into the system is used as the defeasible facts in the Physician Agent’s belief base. The agents utilize domain knowledge extracted from an ACKTUS knowledge repository, which is developed by domain professionals using ACKTUS-dementia.

The implemented solution builds on earlier work on inquiry dialogue systems, modified and extended in order to obtain a more efficient implementation and a dialogue which can follow the natural focus shifts of a human actor.

The approach was found highly interesting by the domain experts who participated in a pilot evaluation study, and our results will be further evaluated with also subjects with lesser experience from and knowledge about dementia diseases.

References