Ontology-Driven Context Interpretation and Conflict Resolution in Dialogue-Based Home Care Assistance

Georgios Meditskos, Efstratios Kontopoulos, Stefanos Vrochidis, Ioannis Kompatsiaris

Information Technologies Institute, Centre for Research and Technology – Hellas
Thessaloniki, Greece
{gmeditsk, skontopo, stefanos, ikom}@iti.gr

Abstract. In this paper we present a framework for conversational awareness and conflict resolution in spoken dialogue systems for home care assistance. Conversational awareness is supported through OWL ontologies for capturing conversational modalities, while interpretation and incremental context enrichment is facilitated through Description Logics reasoning. Conflict resolution further assists the interaction with end users, facilitating exception handling and context prioritisation by coupling defeasible logics with medical and profile information.

Keywords: Ontologies, defeasible logics, dialogue-based systems, healthcare.

1 Introduction

Spoken dialogue systems aim to assist end users in satisfying their information needs, hiding the complexity of knowledge representation and query languages. Of the numerous domains of interest, conversational assistance in healthcare is a notable case where natural language interfaces provide unique solutions to patients and medical experts. In addition, multimodal dialogue-based systems overcome the limitations of dialogue systems that use speech as the only communication means, collecting and analysing information from multiple sources and modalities.

The presented framework focuses on enriching multimodal dialogue-based agents with (a) intelligent context aggregation for conversation understanding, and (b) conflict resolution of domain inconsistencies and conflicts. To this end, OWL ontologies are used for modelling multimodal information (e.g. verbal and non-verbal modalities) and the semantics that underpin the interpretation logic, while defeasible logics [1] provide the non-monotonic semantics needed to deliver advanced conflict resolution strategies.

2 Related Work

In the domain of natural language interfaces and dialogue-based systems, ontologies such as WordNet and BabelNet, provide the vocabulary and semantics for content disambiguation [2]. Ontologies and Description Logics (DL) [3] have also been used in

NLP for co-reference resolution [4]. In multimodal fusion, ontologies are used for fusing multi-level contextual information [5]. For example, [6] presents a framework for coupling audio-visual cues with multimedia ontologies. Relevant approaches are also described in [7] for multimedia analysis tasks. As far as defeasible reasoning is concerned, the non-monotonic semantics of the logic has been mainly used for building argumentative dialogue-based systems [8] or resolving conflictual arguments through counterarguments [9]. Through the use of DL reasoning for conversational awareness and defeasible rules for conflict resolution, this work focuses on conversation understanding and high-level conflict resolution.

3 Ontology-Driven Conversational Awareness

Contextual information, such as multimedia information (e.g. speech analysis, named entities and concepts) and video analysis (e.g. gestures, facial expressions) is mapped to ontological entities in a hierarchical manner. The topic hierarchy defines the way conversational observations can lead to the derivation of high-level interpretations. In terms of DL semantics, the `Topic` (root) class is defined as:

$$\texttt{Topic} \equiv \exists\texttt{contains.Observation} \tag{1}$$

For example, the recognition of a topic that indicates a pain problem is defined as:

$$\texttt{PainTopic} \equiv \texttt{Topic} \sqcap \exists\texttt{contains.HurtReference} \tag{2}$$

$$\texttt{HurtSpoken} \sqsubseteq \texttt{HurtReference} \tag{3}$$

Topics can be further specialized hierarchically, defining additional `contain` property restrictions. For example, for the recognition of certain symptoms of pain, e.g. headache based on language analysis and deictic gestures, (2) can be extended as:

$$\texttt{HeadacheTopic} \equiv \texttt{PainTopic} \sqcap \exists\texttt{contains.HeadReference} \tag{4}$$

$$\texttt{HeadReference} \equiv \texttt{HeadDeictic} \sqcup \texttt{HeadSpoken} \tag{5}$$

The hierarchical topic decomposition also facilitates the descriptive modelling of topic-related semantics, i.e. to model descriptive information that does not directly define the conversational topic but provides useful information to drive the interaction with the user (see Section 5). Descriptive context is modelled in terms of the `DescriptiveContext` hierarchy, whose root class is defined as:

$$\texttt{DescriptiveContext} \equiv \exists\texttt{requires.Concept} \tag{6}$$

The descriptive context of a topic is specified through one or more `requires` property assertions about domain concepts. For example, `PainTopic` can be further associated with structures denoting the intensity or the part of the body:

$$\begin{aligned}\texttt{PainTopic} \sqsubseteq \texttt{DescriptiveContext} \sqcap (&\exists\texttt{requires.Intensity} \\ &\sqcup \exists\texttt{requires.BodyPart})\end{aligned} \tag{7}$$

Similarly, the descriptive context of headache may contain structures relevant to sleep quality or coffee consumption:

$$\texttt{HeadacheTopic} \equiv \texttt{DescriptiveContext} \\ \sqcap (\exists \texttt{requires.SleepQuality} \sqcup \exists \texttt{requires.CoffeeConsumption}) \tag{8}$$

4 Context-based Reasoning and Conflict Resolution

Context-based reasoning aims at coupling the semantics of conversational awareness with background knowledge, such as medical and profile information, in order to acquire a better understanding of the situation, resolve conflicts and provide the most plausible responses. Each conversational topic t is associated with a defeasible rule base D_t that handles domain contextual semantics. Assuming that T is the set of all conversational topics supported ($\forall t \in T, T \sqsubseteq Topic$), we define

$$\forall t \in T, D_t = \{r_i : A(r_i) \looparrowright C(r_i)\}$$

where r_i is a unique label of the rule, $A(r_i)$ is the antecedent, $C(r_i)$ is the consequent and $\looparrowright$ indicates the rule type: strict ($\rightarrow$), defeasible ($\Rightarrow$) or defeater ($\in$). Intuitively, the detection of t triggers the inference mechanisms of the defeasible rule base D_t.

5 Use Case

We describe the simulated evaluation of our framework that is part of the KRISTINA agent [10] (Fig. 1) and involves interaction with users at a home in order to acquire information about their condition and suggest treatments for frequent problems. In one of the evaluation scenarios, the user informs the agent about feeling pain ("*I feel pain*"). The Dialogue Manager (DM) collects the incoming verbal observation, which involves a hurt reference captured by language analysis, and builds the current context:

$$\texttt{Topic(t1), HurtSpoken(h1), contains(t1, h1)} \tag{9}$$

The context is then passed to Conversational Awareness to interpret the topic and, according to axioms (2) and (3), it classifies `t1` in the `PainTopic` class. Next, the available descriptive context is collected. According to (7), `PainTopic` is associated with the `Intensity` and `BodyPart` concepts that are sent back to DM to decide upon next steps. In our scenario, it is assumed that DM decides to further enrich the current conversational context by asking the user where he hurts. The user points to his head and says: "*It hurts here*". Again, a hurt spoken reference is detected from speech analysis, as well as a deictic gesture to the head. Both observations are added to the topic instance `t1` (9) using `contains` assertions:

$$\texttt{HurtSpoken(h2), HeadDeictic(hd1), contains(t1, h2),} \\ \texttt{contains(t1, hd1)} \tag{10}$$

The new contextual information is passed to Conversational Awareness to reason again on the current context. The enriched context now satisfies (4), and `HeadacheTopic` becomes the current conversational topic that, along with its descriptive context `SleepQuality` and `CoffeeConsumption` (derived by (9)) are sent back to DM. DM

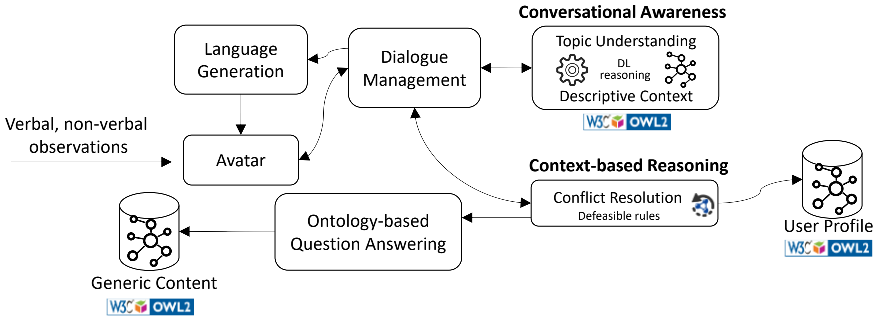

Fig. 1. Conceptual architecture of the simulated evaluation. Arrows visualize interactions.

decides not to further enrich the context (e.g. by asking questions about sleep problems or coffee consumption habits) and propagates the current context (`HeadacheTopic`) to Context-based Reasoning for generating appropriate responses.

The generic defeasible logics rule base for `HeadacheTopic` involves the following defeasible rules for relevant treatment recommendations:

$r_1: HeadacheTopic \Rightarrow recommendSleep$
$r_2: HeadacheTopic \Rightarrow recommendNoCoffee$
$r_3: HeadacheTopic \Rightarrow recommendMildPainkillers$

According to the elderly's profile, he suffers from frequent migraines and caffeine intolerance. Therefore, the following personalized rules are also considered:

$r_4: HeadacheTopic, Profile\_CaffeineIntolerance \in \neg recommendNoCoffee$
$r_5: HeadacheTopic, Profile\_Migraines \Rightarrow recommendStrongPainkillers$
$r_5 > r_3$ and $C = \{recommendMildPainkillers, recommendStrongPainkillers\}$

In addition, the scenario involves a sleep sensor that monitors night sleep quality and provides an assessment every morning. The following defeater enriches context-based reasoning by fusing sleep quality information that overrides r_1:

$r_6: HeadacheTopic, Logs\_GoodSleepQuality \in \neg\, recommendSleep$

The rule base of the example (via SPINdle [11]) finally recommends that the user should take strong painkillers for his headache, since he suffers from migraines, overriding other plausible recommendations based on profile and sleep-related information.

6 Conclusions

In this paper we presented a framework for conversational awareness and conflict resolution in spoken dialogue systems combining ontologies and defeasible reasoning. OWL is used to model multimodal input and the semantics that underpin the conversational logic, while defeasible rules provide the non-monotonic semantics needed to deliver intuitive knowledge representation and advanced conflict resolution.

We are currently conducting pilots for collecting additional data and evaluating the framework with more use cases. In parallel, we are working towards further enrichment of the fusion and interpretation capabilities of the framework, so as to support additional use cases, e.g. taking into account emotions and facial expressions.

Acknowledgements. This work has been partially supported by the H2020-645012 project "KRISTINA: A Knowledge-Based Information Agent with Social Competence and Human Interaction Capabilities".

References

1. Maier, F., Nute, D.: Well-founded semantics for defeasible logic. Synthese. 176, 243–274 (2010).
2. Damljanović, D., Agatonović, M., Cunningham, H., Bontcheva, K.: Improving habitability of natural language interfaces for querying ontologies with feedback and clarification dialogues. Web Semant. Sci. Serv. Agents World Wide Web. 19, 1–21 (2013).
3. Baader, F.: The description logic handbook: theory, implementation, and applications. Cambridge university press (2003).
4. Prokofyev, R., Tonon, A., Luggen, M., Vouilloz, L., Difallah, D.E., Cudré-Mauroux, P.: SANAPHOR: Ontology-Based Coreference Resolution. In: International Semantic Web Conference. pp. 458–473. Springer (2015).
5. Dourlens, S., Ramdane-Cherif, A., Monacelli, E.: Multi levels semantic architecture for multimodal interaction. Appl. Intell. 38, 586–599 (2013).
6. Perperis, T., Giannakopoulos, T., Makris, A., Kosmopoulos, D.I., Tsekeridou, S., Perantonis, S.J., Theodoridis, S.: Multimodal and ontology-based fusion approaches of audio and visual processing for violence detection in movies. Expert Syst. Appl. 38, 14102–14116 (2011).
7. Atrey, P.K., Hossain, M.A., El Saddik, A., Kankanhalli, M.S.: Multimodal Fusion for Multimedia Analysis: A Survey. Multimed. Syst. 16, 345–379 (2010).
8. Modgil, S., Prakken, H.: The ASPIC+ framework for structured argumentation: a tutorial. Argum. Comput. 5, 31–62 (2014).
9. Prakken, H.: On dialogue systems with speech acts, arguments, and counterarguments. In: European Workshop on Logics in Artificial Intelligence. pp. 224–238. Springer (2000).
10. Wanner, L., Blat, J., Dasiopoulou, S., al, et: Towards a Multimedia Knowledge-Based Agent with Social Competence and Human Interaction Capabilities. In: Proceedings of the 1st International Workshop on Multimedia Analysis and Retrieval for Multimodal Interaction. pp. 21–26. ACM (2016).
11. Lam, H.-P., Governatori, G.: The making of SPINdle. In: International Workshop on Rules and Rule Markup Languages for the Semantic Web. pp. 315–322. Springer (2009).