Speech Interfaces for Point-of-Care Guideline Systems

Martin Beveridge, John Fox and David Milward

Cancer Research UK[1], London WC2A 3PX
{mb, jf, dm}@acl.icnet.uk

A major limiting factor in the acceptability of interactive guideline and decision support systems is the ease of use of the system in the clinic. A way to reduce demands upon users and increase flexibility of the interface is to use natural language dialogues and speech based interfaces. This paper describes a voice-based data capture and decision support system in which knowledge of underlying task structure (a medical guideline) and domain knowledge (disease ontologies and semantic dictionaries) are integrated with dialogue models based on conversational game theory resulting in a flexible and configurable interface.

Introduction

Natural language interfaces are likely to be important in future healthcare systems. However, their development is a greater challenge than applications that have been investigated to date (e.g. route planning, flight booking etc) which generally require only simple information look-up. Clinical systems may include complex reasoning and workflow management, and dialogue may need to be closely coupled to the underlying clinical context. This paper uses discourse analysis as a basis for combining dialogue techniques with models of clinical tasks (e.g. data-capture and decision-making) and ontological knowledge (e.g. about diseases and symptoms). The approach has three benefits: 1) dialogue is tied to natural clinical tasks, providing guidance and constraints for understanding input and interpreting intentions; 2) dialogues can be generated automatically from domain knowledge, and 3) the dialogue generator can be reconfigured for other domains without reprogramming.

The structure of dialogues

The *linguistic* structure of a discourse obviously includes the sequence of utterances that comprise the discourse, but in addition these utterances are considered to naturally aggregate into discourse segments (analogous to constituents in sentence syntax) [4] and the discourse must be understood in terms of functional aspects of these seg-

[1] This work was carried-out as part of the EU-funded HOMEY project (Home Monitoring through an Intelligent Dialogue System, IST-2001-32434). Many thanks to our partners for all their helpful comments and advice: Engineering (www.eng.it), Reitek (www.reitek.com), CBIM (www.cbim.it), ITC (www.itc.it), and L&C (www.landcglobal.com).

ments: so-called intentional, informational and attentional aspects. *Intentional* structure deals with two kind of relations between the intentions which underpin discourse segments [4], namely: *satisfaction-precedence (SP*) and *dominance (DOM)* relations. For example, the intention to obtain a clinical history "satisfaction-precedes" the intention to make a diagnosis, and is partly satisfied by or "dominates" the intention to find out a patient's age. Discourse can also be described in terms of *informational* (semantic) relationships between discourse segments [6]. For example, in order to generate the sentence "John will be treated urgently because his condition is life threatening" it is necessary to know about the semantic relation of causality between the notion of "life threatening condition" and "being treated urgently" in order to generate the appropriate linking word "because" (rather than "unless", "although" etc). Lastly, discourse can also be described in terms of the way it unfolds over time. In [4] this is represented through the notion of a dynamic *attentional state*, which describes all the objects, properties and relations that are salient at a particular point in a discourse. The attentional state coordinates the linguistic structure and non-linguistic representations such as intentions and information relations. Recently there has been a growing consensus that all three of the structures described above are required to represent discourse.

In attempting to describe *dialogue* (conversational discourse) one approach that has proven valuable is Conversational Game Theory [5]. This represents dialogue in terms of conversational *games*, a plan-based level associated with intentions, and a structural level consisting of sequences of conversational *moves* which specify the linguistic structure required to satisfy those intentions. Dialogues are thought-of as a series of games each aiming to achieve some sub-goal of the dialogue. The present approach to specifying natural language dialogue builds on a combination of these ideas.

Intentional Structure

Intentions are implicitly captured by the structure of games. For example, a game that is initiated by a *query-yn* move reflects an underlying intention on the part of the "initiating conversational partner" (ICP) that the "other conversational partner" (OCP) should intend that the ICP know if some state of affairs holds. If the OCP is cooperative then they will adopt this intention and make an appropriate *reply-yn* move. Dominance and satisfaction-precedence relations can be treated as relations between games. The initiating and response moves currently implemented in our system are shown in the following table. Each participant may also respond with the initiating move of a new sub-game whose intention sub-serves that of the parent game.

Initiating moves	**Response moves**
Explain: provide information not previously requested	*Acknowledge*: acknowledge and signal continuation
Instruct: provide instruction	*Reply-yn*: yes/no reply
Query-yn: yes/no query for unknown information	*Reply-w*: reply supplying a value
Query-w: complex (wh-)query for unknown information	

Information Structure

One of the problems associated with applying information relations to dialogue is determining the appropriate units that such relations should apply to. In text generation, they are applied to successive utterances, but in dialogue they may span more than one utterance and speaker [7]. For example, here is a dialogue fragment from our system when advising on whether a patient should be referred to a specialist:

1. *S*: Is there any nipple discharge? [*Query-yn*]
2. *U*: Yes [*Reply-yn*]
3. *S*: Ok… [*Acknowledge*]
4. *S*: And is it bloodstained? [*Query-yn*]
5. *U*: No [*Reply-yn*]
6. *S*: Ok. [*Acknowledge*]

In this example, the second *Query-yn* game (4, 5 and 6) elaborates the information provided in the first *Query-yn* game (1, 2 and 3).

Attentional State

There are likely to be several playable moves at any point in a dialogue, so it is necessary to determine which is the preferred move. This must take into account constraints imposed by the intentional and information structures. For example, if the intentional structure specifies that I_1 *SP* I_2 then I_2 should not be chosen until I_1 is satisfied; if I_1 *DOM* I_2 and I_1 *DOM* I_3 then I_2 and I_3 satisfy a common goal so they can be presented in succession, or by aggregating them (e.g. "what are the patient's age and sex?"). Information relations can also help to preserve dialogue coherence by ensuring that the next move made by the system is as semantically relevant as possible to previous moves. For instance: in the next example the user introduces the topic of nipple discharge and the system chooses its next move so as to continue that topic (rather than pursuing other intentions that it may have in its dialogue plan):

1. *S*: What is the patient's age?
2. *U*: They're thirty and they have severe nipple discharge
3. *S*: Ok…
4. *S*: And is the nipple discharge bloodstained?
5. *U*: No.
6. *S*: Ok…

Separating domain knowledge from dialogue knowledge

Among the aims of this work are the ability to exploit existing domain representation schemas in generating dialogue specifications automatically. The system includes domain knowledge of two types: *task knowledge* and *ontological knowledge*.

Task knowledge

Since the domain plan determines the tasks to be carried out (e.g. first take a patient history then make the decision), it provides a basis for deriving the intentional structure of a dialogue about that domain. In fact, the plan imposes certain obligations on the dialogue system in order that the process can be completed successfully, and the dialogue system must interact with the user to meet those obligations. This approach is consistent with the suggestion that dialogue structure is largely determined by task structure [4].

Intentional relations can be derived from relations between tasks in the domain plan. For example, preconditions of tasks within the plan can be considered to give rise to *satisfaction-precedence* relations in the intentional structure. Hence if task T_2 has preconditions such that it cannot be started until task T_1 has completed (e.g. you must take patient details before making a referral decision) then the relation I_1 *SP* I_2 can be inferred between the associated intentions I_1 and I_2. *Dominance* relations can similarly be inferred from decomposition relations between tasks, e.g. if T_1 is decomposed into T_2 and T_3 then the relations I_1 DOM I_2 and I_1 DOM I_3 can be inferred.

Ontological knowledge

Task-specific knowledge must be augmented with a conceptual model that describes general domain knowledge e.g. 'breast cancer is-a cancer', 'nipple discharge is-symptom-of breast cancer' and so forth. In our work the domain ontology forms the basis for deriving information relations between dialogue games. A fragment of the ontology (taken from [2]) is shown in Figure 1.

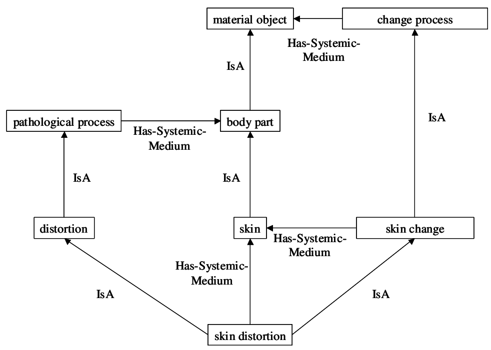

Fig. 1. A fragment of the domain ontology

Use of information in a cancer guideline

In other work our group has developed a system for advising doctors on whether patients require urgent referral for suspected cancer [1]. The system is accessed by a standard web browser that generates web pages for collecting patient data and reporting on results (see www.infermed.com/era).

For the present project we wish to have a voice-based mode for entering data into this system. The task knowledge component of this voice-based system is currently implemented in the PRO*forma* task representation language [1] using the Tallis toolset (www.openclinical.org/kpc). The domain ontology is implemented using an Ontology Browser developed by Language & Computing n.v. (L&C) [2]. A dialogue Engine uses the task descriptions provided by these components to create a high-level dialogue specification (HLDS) that describes the games to be played to complete current tasks. The HLDS is in turn used to create a sequence of moves that can be made by either participant at the current point in the dialogue. The result of this process is encoded as a VoiceXML document, which is then interpreted by a voice browser which controls automatic speech recognition (ASR) and text-to-speech (TTS) components. The voice browser and ASR are provided by Istituto Trentino di Cultura (ITC) [3] and are integrated, along with the Actor multilingual TTS produced by Loquendo (www.loquendo.com), into an interactive voice response (IVR) platform provided by Reitek Sp.A. The IVR platform typically handles telephony control, audio recording etc so that the dialogue system can be accessed over the phone.

Conclusions

An approach to building spoken dialogue systems that treats the dialogue model as having distinct high-level and low-level representations has been described. This uses current voice-based standards which are widely employed in commercial systems for the low-level elements (e.g. VoiceXML) whilst also expressing high-level notions of intention, information and attention which are required for flexible "conversational" dialogue. The high-level dialogue representation can be automatically derived from the domain knowledge (task and ontological knowledge), reducing the need to author dialogues by hand, and providing reconfigurability. A complete demonstrator has been implemented for the domain of breast cancer and is currently being evaluated.

References

1. J Bury M Humber, J Fox Integrating Decision Support with Electronic Referrals In R. Rogers , R. Haux and V. Patel (Eds). Medinfo. 2001. IOS Press, Amsterdam -
2. Ceusters, W., Beveridge, M. A., Milward, D., and Falavigna, D. (2002). Specification for Semantic Dictionary Integration, Deliverable D9, EU HOMEY Project, IST-2001-32434.
3. Falavigna, D. and Gretter, R. (1999). Flexible Mixed Initiative Dialogue over the Telephone Network. Proc. Of ASRU'00, 12th – 15th December, Colorado.
4. Grosz, B., and Sidner, C. (1986). Attention, Intention and the Structure of Discourse. *Computational Linguistics* 12(3):175-204.
5. Kowtko, J. C. and Isard, S. D. (1993). *Conversational Games Within Dialogue*, Research Paper 31, Human Communication Research Centre, Edinburgh.
6. Mann, W. D., and Thompson S. A. (1988). Rhetorical Structure Theory: Towards a functional theory of text organization. *Text,* 8(3):243-281.
7. Stent A. (2000). Rhetorical Structure in Dialog. *Proc. 2nd International Natural Language Generation Conference* (INLG'2000).