A human-machine dialogue system for CALL

Peter Vlugter, Alistair Knott
Dept of Computer Science
University of Otago
{pvlugter/alik}@cs.otago.ac.nz

Victoria Weatherall
School of Mäori, Pacific and Indigenous Studies
University of Otago
victoria.weatherall@stonebow.otago.ac.nz

Abstract

This paper describes a CALL tool which interacts with the student via a bilingual human-machine dialogue system. We focus on the modifications needed to the system's grammar and dialogue manager to cater for student errors and help to diagnose them.

1 Introduction: dialogue-based CALL

Dialogue is a common medium for ordinary language teaching and learning. A classroom teacher responding to questions, or asking questions, is engaging in dialogue; so is a student attempting a conversation with a native speaker and 'learning by doing' in the process. There are several reasons why dialogue is a useful environment for language learning. Firstly, in the case of an L2 dialogue, the medium of the interaction is itself the topic being learned. Secondly, dialogue is a means by which both the teacher and the student can shape the learning experience: for instance, the teacher can initiate various kinds of exercise, or the student can ask questions to clarify or extend what they currently know. Finally, the teacher can analyse the student's dialogue contributions to diagnose how well the student has assimilated the material to be learned.

In this paper, we describe a CALL system for Mäori, the indigenous language of New Zealand. The system engages in a mixed-initiative dialogue with the student in a mixture of English and Mäori. The system is an extension of a 'conventional' human-machine dialogue system. In Section 2, we outline the mechanisms for syntactic, semantic and dialogue processing in the conventional system. In Sections 3 to 5, we describe how these mechanisms need to be extended to support the kind of interactions needed in the CALL domain.

2 Overview of the Te Kaitito system

Our dialogue system, called Te Kaitito[1], is built using a collection of natural-language-processing modules for English and Mäori. When the student enters a sentence, it is first processed by a **sentence parser**, using a combined grammar of Mäori and English. The sentence parser is the LKB system (Copestake, 2002). The result of parsing is a set of possible syntactic structures, each associated with a semantic representation of the sentence. A **dialogue attachment module** computes how each of these semantic representations could be incorporated into the current dialogue context. A **disambiguation module** then selects one of these representations, based on a combination of syntactic preferences and considerations about ease of dialogue attachment. The **dialogue manager** then updates the dialogue context, and decides what to say in response, using information from the context and from its own private knowledge base of facts and goals. The response utterance is either generated by a simple template system, or by invoking the **sentence generator**, which uses the same grammar as the parser to turn a semantic representation back into a surface sentence. Sentence generation is also done by the LKB system.

2.1 Syntactic representations

The syntactic formalism used in LKB is HPSG (Pollard and Sag, 1994). We use a bilingual grammar, in which words and rules are associated with a LANGUAGE feature whose value is either MAORI or ENGLISH. To prevent sentences using rules from both languages, agreement is enforced for the LANGUAGE feature throughout the sentence.

2.2 Semantic representations

The semantic representations delivered by LKB are given in a language called Minimal Recursion Semantics (MRS; Copestake *et al.*, 1999). For dialogue processing, we convert MRS representations into a format related to presuppositional DRT (see Kamp *et al.*, in preparation). Very briefly, each utterance is a structure associated with a **dialogue act** (e.g. *query*), a **speaker**, an **addressee**, a **message** and a list of **bindings**. The message can be either a **question** or a **proposition**. Propositions contain an **asserted** part and a set of **presuppositions**.

[1] 'Te Kaitito' is Mäori for 'the improvisor' or 'the extempore speaker'.

Questions have a (possibly empty) set of **parameters** which represent the abstracted parts of the question and scope over a proposition. The bindings list contains a binding for the main variable of each presupposition in the message to a variable in the discourse context.

2.3 Example: a question and answer dialogue

Consider the simple question and answer dialogue given in Example 1.

(1) System: Nö hea koe?
(Where are you from?)
Student: Nö Ötepoti ahau.
(I'm from Ötepoti.)

The information state of the system before this dialogue occurs is given in Figure 1. The system is represented by a discourse entity a_1, the student by a_2, and within the common ground there is an entity a_3 which is a place named Ötepoti. The semantic representation for the question posed by the system is shown in Figure 2 and for the student's answer in Figure 3.[2] (Note the way in which 'you' in the question and 'I' in the answer both refer to the student.)

$$\begin{bmatrix} \textit{information state} & \\ \text{SYSTEM} & a_1 \\ \text{STUDENT} & a_2 \\ \text{COMMON GROUND} & \left\{ \begin{array}{c} \vdots \\ named(a_3, \bar{\text{O}}\text{tepoti}) \\ \vdots \end{array} \right\} \end{bmatrix}$$

Figure 1: The information state

$$\begin{bmatrix} \textit{query} & \\ \text{SPEAKER} & a_1 \\ \text{ADDRESSEE} & a_2 \\ \text{MSG} & \begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \{place(x_2)\} \\ \text{PROP} & \begin{bmatrix} \textit{proposition} & \\ \text{ASSERTED} & \{from(e_1, x_1, x_2)\} \\ \text{PRESUPS} & \{addressee(x_1)\} \end{bmatrix} \end{bmatrix} \\ \text{BINDINGS} & \{x_1 = a_2\} \end{bmatrix}$$

Figure 2: Nö hea koe? (Where are you from?)

$$\begin{bmatrix} \textit{answer} & \\ \text{SPEAKER} & a_2 \\ \text{ADDRESSEE} & a_1 \\ \text{MSG} & \begin{bmatrix} \textit{proposition} & \\ \text{ASSERTED} & \{from(e_2, x_3, x_4)\} \\ \text{PRESUPS} & \left\{ \begin{array}{l} speaker(x_3), \\ named(x_4, \bar{\text{O}}\text{tepoti}) \end{array} \right\} \end{bmatrix} \\ \text{BINDINGS} & \{x_3 = a_2, x_4 = a_3\} \end{bmatrix}$$

Figure 3: Nö Ötepoti ahau (I'm from Ötepoti)

When processing a question and answer pair like this, the system first matches the propositional part of the question to the proposition given by the answer. This is done by matching the asserted parts of each proposition and then checking that the corresponding bindings for any presupposed information also match. The answer can then be checked to see if it is also appropriate to the parameters of the question. If the matching is succesful then the answer is considered a proper answer to the question. When the question is a genuine question (where the system is asking for new information) the asserted part of the answer will be grounded, as shown in the updated information state given in Figure 4.

$$\begin{bmatrix} \textit{information state} & \\ \text{SYSTEM} & a_1 \\ \text{STUDENT} & a_2 \\ \text{COMMON GROUND} & \left\{ \begin{array}{c} \vdots \\ named(a_3, \bar{\text{O}}\text{tepoti}), \\ from(e_2, a_2, a_3) \\ \vdots \end{array} \right\} \end{bmatrix}$$

Figure 4: The updated information state

2.4 Utterance disambiguation and dialogue attachment

Utterances coming into the dialogue system can be ambiguous in several ways, due to multiple syntactic readings, multiple presupposition resolutions, or multiple dialogue interpretations. In our system, the main mechanism for resolving these ambiguities is **ease of dialogue attachment**.

If there are several alternative interpretations of a sentence, it is often appropriate to choose the one which is easiest to incorporate into the current dialogue context. For instance, if one interpretation contains presuppositions which cannot be resolved and another does not, it makes sense to choose the latter interpretation (all other things being equal). To take another example, if a sentence can be interpreted as an assertion or as an answer to a question, and a question has just been asked, it is

[2] These representations have been simplified; e.g. we omit details of scope underspecification, and tense and aspect.

interpreted as an answer, while if no question has been asked, it is interpreted as an assertion. (See Knott and Vlugter, 2003 for details.)

2.5 The CALL system

The CALL application we are developing is aligned with an introductory course in conversational Mäori running at Otago University. The coverage of our grammars and of the dialogue system is based on the examples of Mäori dialogues found in the course textbook. The student will then be able to play with a wide range of similar dialogues. Our treatment of errors is based on a study of student errors in coursework and exams (in preparation).

The system supports two kinds of dialogue. In **authoring mode**, the system interacts with a human author setting up the scenario for a single lesson. All initiatives in this mode come from the author. The author can make assertions, which go towards creating a knowledge base of facts, stored simply as a set of sentence representations. The author can also enter questions, which go towards creating an agenda of system goals; these are likewise stored as sentence representations.

In **student mode**, the system interacts with a student in a mixed-initiative dialogue, centered around one of the authored scenarios. The system can take initiatives by asking questions from its agenda of questions, or by creating teaching questions based on material from the common ground, or by asserting new facts about the scenario. The student can take initiatives by asserting new facts, or by asking questions.

3 Error diagnosis using syntax

The first thing the system needs to be able to do is to deal robustly with syntactic errors in the sentences entered by the student. For instance, accusative Mäori NPs require a case-marking particle "i", as illustrated in Example 2. This is often left out, as in Example 3.

(2) Kei te whai te kurï i te ngeru.
TAM chase the dog ACC the cat.
"The dog is chasing the cat"

(3) *Kei te whai te kurï te ngeru.

There has been a lot of work on how to implement error grammars using mal-rules which generate sentences containing grammatical mistakes (see Menzel and Schröder, 2003 for a review). In our system we take a simple approach in which the LANGUAGE feature of rules can be further defined for a number of independent errors, each of which takes a boolean value. For instance, the rule which allows an object NP without an accusative case-marker is defined for the sub-feature CASE-MARKER-ERROR=TRUE. The agreement requirements on the LANGUAGE feature mean that information about all the errors in the sentence will be visible on the parent node.

4 Error diagnosis in the dialogue manager

Sometimes a student's misconceptions about L2 can manifest themselves in L2 sentences which are actually syntactically correct. To illustrate, consider some additional possible answers to the question in Example 1:

(4) Q: Nö hea koe? (Where are you from?)
A1: Nö Ötepoti ahau. (I'm from Ötepoti.)
A2: #Nö Ötepoti koe. (You're from Ötepoti.)
A3: #Kei Ötepoti ahau. (I'm at Ötepoti.)

The semantic representation of the question was given in Figure 2. As before, the correct answer is A1. If the student answers using A2 or A3, she has made some kind of mistake, but note that the sentences themselves are both syntactically correct. In A2, the student has responded by repeating the same personal pronoun ("you") as was used in the question. In A3, she has the right pronoun, but she is answering a slightly different question.

In such cases, catching the error is a matter of recognising infelicitous dialogue acts. Specifically, these are cases where the student has an obligation to respond to a question, but makes an utterance which does not in fact provide the answer. We suggest that a general way to diagnose these mistakes is to modify the dialogue attachment module by relaxing the constraints on how answers are matched back to questions. For answer A2 (see Figure 5), this involves allowing a mismatch between the bindings made by the question and the answer; x_1 in the question, bound to a_2 (the student) is matched with x_3 in the answer, bound to a_1 (the system). For answer A3 (see Figure 6), the relaxation involves allowing a mismatch in the identity of the predicates *from* and *at*.

$$\begin{bmatrix} \textit{answer} & \\ \text{SPEAKER} & a_2 \\ \text{ADDRESSEE} & a_1 \\ \text{MSG} & \begin{bmatrix} \textit{proposition} & \\ \text{ASSERTED} & \{\textit{from}(e_2,\, x_3,\, x_4)\} \\ \text{PRESUPS} & \begin{Bmatrix} \textit{addressee}(x_3), \\ \textit{named}(x_4,\, \text{Ötepoti}) \end{Bmatrix} \end{bmatrix} \\ \text{BINDINGS} & \{x_3 = a_1,\, x_4 = a_3\} \end{bmatrix}$$

Figure 5: Nö Ötepoti koe (You're from Ötepoti)

$$\left[\begin{array}{ll} \textit{answer} & \\ \text{SPEAKER} & a_2 \\ \text{ADDRESSEE} & a_1 \\ \text{MSG} & \left[\begin{array}{ll} \textit{proposition} & \\ \text{ASSERTED} & \{at(e_2,\, x_3,\, x_4)\} \\ \text{PRESUPS} & \left\{\begin{array}{l} \textit{speaker}(x_3), \\ \textit{named}(x_4,\, \bar{\text{O}}\text{tepoti}) \end{array}\right\} \end{array}\right] \\ \text{BINDINGS} & \{x_3 = a_2,\, x_4 = a_3\} \end{array}\right]$$

Figure 6: Kei Ötepoti ahau (I'm at Ötepoti)

5 Interactions between syntax and dialogue-based error diagnosis

A final type of error needs to be picked up by a combination of information about grammatical errors and information about dialogue attachment. In these errors, the fact that the student has made a mistake of some kind is picked up by the error grammar, but in order to work out *what* error is made, information about dialogue attachment is needed. For instance, consider Example 5:

(5) Q: Kei hea te kurï? (Where is the dog?)
A1: Kei roto te kurï i te mära.
TAM in the dog ACC the garden.
A2: Kei roto i te mära te kurï.
TAM in ACC the garden the dog.
A3: *Kei roto te mära te kurï.

A1 and A2 are both syntactically correct answers to the question, which mean "the dog is in the garden". In Mäori locative sentences, the order of the subject and object DPs is free: in A1, the subject comes first, while in A2, the object comes first. This freedom is possible because the object is identifed by the accusative case-marker "i".

In A3, the case-marker as been left out, resulting in a syntactically incorrect sentence. As we saw in Section 3, leaving out the case-marker is a common syntactic error. However, because of free constituent order in locative sentences, there are two ways to fix this mistake: one is to add a case-marker to "the dog", and the other is to add a case-marker to "the garden". Both mistakes are frequently attested in our error analysis, so both these buggy rules need to be in our error grammar. The result is a sentence which has two interpretations containing mal-rules. Each parse is associated with a different semantic representation, just as regular parses are.

How can we choose between these alternatives? In fact, the principle of dialogue attachment described in Section 2.4 for ordinary ambiguity resolution extends straightforwardly to such cases. The analysis in which "i" is missing from "the garden" gets the interpretation that the dog is in the garden, while the analysis in which "i" is missing from "the dog" gets the interpretation that the garden is in the dog. The former interpretation can be attached to the dialogue context as an answer to the question just asked, while the latter cannot. The principle of ease of dialogue attachment therefore correctly dictates a preference for the former interpretation.

6 Summary and conclusions

We began by arguing that free, mixed-initiative human-computer dialogue would be a good medium for CALL. We then described a few ways in which an existing human-computer dialogue system has been extended to function as the back end of a dialogue-based CALL system. We found that many of the error-diagnosing and error-correcting functions which an intelligent CALL system should have are quite natural extensions of the dialogue system's ordinary functionality. In conclusion, we think there is considerable potential in mixed-initiative human-computer dialogue as a feasible and useful medium for CALL applications.

7 Acknowledgements

Thanks to the NZ Foundation for Research in Science and Technology for grant UOOX0209.

References

Copestake, A. 2002. *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford.

Copestake, A., Flickinger, D., Sag, I., and Pollard, C. 1999. *Minimal Recursion Semantics: An introduction*. Manuscript, CSLI, Stanford University.

Kamp, H., van Genabith, J., and Reyle, U. in preparation. *Discourse representation theory*. Handbook of Philosophical Logic.

Knott, A. and Vlugter, P. 2003. "Syntactic disambiguation using presupposition resolution" in *Proceedings of the 4th Australasian Language Technology Workshop (ALTW2003)*, Melbourne.

Menzel, W. and Schröder, I. 2003. *Error diagnosis for language learning systems*. Lecture notes for a course on Error diagnosis and feedback generation for language tutoring systems at the ELSNET Summer School, Lille, France.

Pollard, C. and Sag, I. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press.