

A Phrasebook Style Medical Speech Translator

Manny Rayner
Netdecisions Technology Centre
Westbrook Centre
Milton Road
Cambridge CB4 1YG
United Kingdom
manny.rayner@netdecisions.com

Pierrette Bouillon
University of Geneva
TIM/ISSCO
40, bvd du Pont-d'Arve
CH-1211 Geneva 4
Switzerland
pierrette.bouillon@issco.unige.ch

1 Introduction

The demo illustrates a lightweight speech translation architecture aimed at medical and other safety-critical application domains. The central requirement in all such domains is complete reliability. Assuming correct speech recognition, which can be checked by feeding back the recognition result, the user must be able to trust the system never to produce meaningless or incorrect translations. These demands appear to rule out most existing frameworks for speech translation. In particular,

- It is hard to see that any architecture which involves statistical or heuristic components can ever achieve the requisite degree of reliability.
- Rule-based system which produce ambiguous results (ambiguous parses, ambiguous generation strings) are also suspect, unless the ambiguity is resolved by deterministic methods.

With the above considerations in mind, we have implemented a rule-based transfer architecture, where both the analysis and generation grammars are designed so as to be completely free of ambiguity; this has been verified experimentally by running large numbers of randomly generated utterances through the system. We have aimed to produce not so much a general speech translator as a flexible voice-activated phrasebook. Our feeling has been that anything more ambitious will probably not achieve the desired level of robustness.

The architecture is implemented on top of the standard commercial Nuance Toolkit platform (Nuance, 2002), and runs comfortably on mid-range PCs. Recognition is guided by an annotated CFG grammar encoded in Nuance Grammar Specification Language (GSL) format; this acts as the language model, and also specifies how source-language semantic representations are to be produced. The grammar is not written by hand, but compiled from a compact unification grammar representation using the open source REGULUS tool (Rayner et al., 2001). Source language semantic representations are feature-value structures, constructed so that the representation of a clause is a flat feature-value list. Transfer and generation are carried out by a suite of simple programs implemented in SICStus Prolog. The transfer component maps sets of feature-value pairs to sets of feature-value pairs; most transfer rules map a single feature-value pair to a single feature-value pair. Generation is carried out by a DCG grammar. Output speech is produced using the Nuance Vocalizer TTS engine.

2 A prototype system

The prototype system translates spoken yes/no questions relating to the symptoms of hypoglycaemia from French into English, using a vocabulary of about 200 words. The basic scenario envisaged is that a French-speaking doctor suspects that an English-speaking patient may be suffering from some form of hypoglycaemia (low blood sugar). The symptoms of hypoglycaemia include anxiety, sweating, tachycardia, tremor,

faintness, headache, confusion, convulsions, and coma; one of the reasons we have chosen hypoglycaemia as a domain is that these symptoms can coincide with those relating to many other conditions, which can often necessitate a lengthy verbal examination.

Coverage is based on a French-language questionnaire constructed by the Association of Hypoglycaemics of Quebec (Thériault, 2002). In terms of content, all questions are of the basic form

“Do you
?(often/sometimes/ever/...)
(do something/experience symptom)
?(at some time/when you do something)”

The grammar provides enough phrasal patterns that it is possible to ask about most domain concepts in a natural way. The most common pattern is some version of

“*avez-vous ?<freq> <symptom> ?<time>*”
 (“do you suffer from ?<freq> <symptom>
?<time>”)

so for example

“*ressentez-vous des engourdissements ?*” →
“do you suffer from numbness?”

“*éprouvez-vous souvent des maux de tête le matin ?*” →
“do you often suffer from headache in the morning?”

“*avez-vous parfois des somnolences quand vous regardez la télé ?*” →
“do you sometimes suffer from sleepiness when you watch TV?”

However, it is often necessary to use other constructions, and the system consequently also supports use of the verb “*être*” (“be”), e.g.

“*êtes-vous émotive ?*” →
“are you emotional?”

intransitive and transitive verbs, e.g.

“*urinez-vous fréquemment la nuit ?*” →
“do you often urinate at night?”

“*manquez-vous toujours d’énergie l’après-midi ?*” →
(lit. “lack you always energy the afternoon?”)
“do you always suffer from lack of energy in the afternoon?”

past tense, e.g.

“*avez-vous déjà eu des convulsions ?*” →
“have you ever suffered from convulsions?”

and use of the “*est-ce que*” construction, e.g.

“*est-ce que vous êtes enceinte ?*” →
“*est-ce que* you are pregnant?”

This project is described in more detail in (Rayner and Bouillon, 2002).

References

- Nuance, 2002. <http://www.nuance.com>. As of 1 Feb 2002.
- M. Rayner and P. Bouillon. 2002. A flexible speech to speech phrasebook translator. In *Proceedings of ACL-02 Workshop on Speech-to-Speech Translation*, Philadelphia, PA.
- M. Rayner, J. Dowding, and B.A. Hockey. 2001. A baseline method for compiling typed unification grammars into context free language models. In *Proceedings of Eurospeech 2001*, pages 729–732, Aalborg, Denmark.
- M. Thériault, 2002. Questionnaire de dépistage pour adultes (in French). As of 15 Mar 2002.