

Analyzing Free Text Segments of Emergency Service Reports

Doug Arnold & Bill Williams
Essex & BAE Systems

Language and Computation Day,
8 Oct, 2010

Basics

- ▶ Emergency Reports can contain Free Text segments.

Basics

- ▶ Emergency Reports can contain Free Text segments.
- ▶ These are multi-participant dialogues,

Basics

- ▶ Emergency Reports can contain Free Text segments.
- ▶ These are multi-participant dialogues,
- ▶ apparently drawn from a lexically and grammatical restricted subset of English (a ‘sublanguage’).

Basics

- ▶ Emergency Reports can contain Free Text segments.
- ▶ These are multi-participant dialogues,
- ▶ apparently drawn from a lexically and grammatical restricted subset of English (a ‘sublanguage’).
- ▶ Can these fragments be processed to produce useful, accurate, information (inferences)?

Fragments

- ▶ The fragments we are dealing with are derived from a Wizard of Oz scenario, but seem authentic;
- ▶ In the form we process them, they consist of four components:
 - ▶ an indication of the *source* of the message ('the speaker');
 - ▶ the *location* of the incident under discussion;
 - ▶ the *time* of the message;
 - ▶ the message itself;

Example

```
src   loc  time  message
pol00 102  2104  rtc reported in coval lane
pol01 27   2105  one car crashed into bus stop.
           car on fire. driver and passenger trapped.
pol01 27   2107  car is focus registration pq04abc.
           please check details.
pol01 27   2110  require assistance to close coval lane.
amb01 100  2110  we have an incident on coval lane:
           a car on fire.
fir02 27   2111  on scene of accident.
           there is a car on fire with two occupants.
```

Project Goals

- ▶ A working demonstration system ('feasability of concept')
- ▶ Using non-proprietary software
- ▶ Analysis of state of the art
- ▶ Short time-frame (December 2009)

Approach (1)

- ▶ Adapt a general purpose parser and grammar, e.g.
 - ▶ C&C Tools
 - ▶ ERG (HPSG)
 - ▶ Pargram Grammar (LFG)
- ▶ This needs investigation, but initially it looks unattractive, because of the linguistic peculiarities of the fragments.

Approach (2)

- ▶ Build a special purpose grammar/parser
 - ▶ What syntactic Framework?
 - ▶ What semantic Framework?

Syntactic framework:

- ▶ Candidates
 - ▶ CG
 - ▶ LFG
 - ▶ HPSG
 - ▶ Something simpler (DCG)

Semantic framework:

- ▶ Candidate(s): DRT is appealing:
 - ▶ conceptually simple
 - ▶ discourse potential
 - ▶ Blackburn and Bos (1999) provides an implementation
 - ▶ SDRT (Asher and Lascarides, 2003) extension is promising

Discourse

- ▶ Multi-participant
- ▶ Variety of illocutionary acts (assertion, question, request, . . .)
- ▶ No clear structure of 'turn taking'

Lexical

- ▶ relatively small lexicon, reflecting narrow domain
- ▶ some genuine openness (e.g. proper names)
- ▶ some specialized vocabulary and abbreviations (*rtc*, *thru*)

Grammatical

▶ article drop

▶ copula drop

(1) a. car crashed on coval lane. driver
trapped.

b. A car has crashed on Coval Lane. The driver is
trapped.

(2) a. rtc on coval lane. car on fire.

b. A road traffic accident has occurred on Coval Lane.
The car is on fire.

Plan/Milestones

date	task
20/10	State of Art Survey
1/10	Scenario
26/10	Design of Demo
23/11	Implementation
??/12	Demo/Presentation

How hard can it be?

```
src    loc  time message
pol01 27  14:43:16 Burglar chased across town park.
                Fell down old well.
pol01 28  14:56:31 I can seeeeeeeeeeeeee hiiiiiiiiiiiiim!
```

References I

Nicholas Asher and Alex Lascarides. *Logics of Conversation*. Studies in Natural Language Processing. Cambridge University Press, 2003.

Patrick Blackburn and Johan Bos. *Working with Discourse Representation Theory: An Advanced Course in Computational Semantics*. 1999. URL <http://homepages.inf.ed.ac.uk/jbos/comsem/book2.html>. Vol II of Representation and Inference for Natural Language: A first course in computational semantics.