GE NLTOOLSET: DESCRIPTION OF THE SYSTEM AS USED FOR MUC-4

George Krupka, Paul Jacobs and Lisa Rau
Artificial Intelligence Laboratory
GE Research and Development
Schenectady, NY 12301 USA
E-mail: rau@crd.ge.com
Phone: (518) 387 - 5059
and
Lois Childs and Ira Sider
Management and Data Systems Operation
GE Aerospace

Abstract

The GE NLTOOLSET *is a set of text interpretation tools designed to be easily adapted to new domains. This report summarizes the system and its performance on the MUC-4 task.*

INTRODUCTION

The GE NLTOOLSET aims at extracting and deriving useful information from text using a knowledge-based, domain-independent core of text processing tools, and customizing the existing programs to each new task. The program achieves this transportability by using a core knowledge base and lexicon that adapts easily to new applications, along with a flexible text processing strategy that is tolerant of gaps in the program's knowledge base.

The language analysis strategy in the NLTOOLSET uses fairly detailed, chart-style syntactic parsing guided by conceptual expectations. Domain-driven conceptual structures provide feedback in parsing, contribute to scoring alternative interpretations, help recovery from failed parses, and tie together information across sentence boundaries. The interaction between linguistic and conceptual knowledge sources at the level of linguistic relations, called "relation-driven control" was added to the system in a first implementation before MUC-4.

In addition to flexible control, the design of the NLTOOLSET allows each knowledge source to influence different stages of processing. For example, discourse processing starts before parsing, although many decisions about template merging and splitting are made after parsing. This allows context to guide language analysis, while language analysis still determines context.

The NLTOOLSET, now in Version 3.0, has been developed and extended during the three years since the MUCK-II evaluation. During this time, several person-years of development have gone into the system. The fundamental knowledge-based strategy has remained basically unchanged, but various modules have been extended and replaced, and new components have been added while the system has served as a testbed for a variety of experiments. The only new module added for MUC-4 was a mechanism for dealing with spatial and temporal information; most of the other improvements to the system were knowledge base extensions, enhancements to existing components, and bug fixes.

The next section briefly describes the major portions of the NLTOOLSET and its control flow; the remainder of the paper will discuss the application of the Toolset to the MUC-4 task.

SYSTEM OVERVIEW

Processing in the NLTOOLSET divides roughly into three stages: (1) *pre-processing*, consisting mainly of a pattern matcher and discourse processing module, (2) *linguistic analysis*, including parsing and semantic interpretation, and (3) *post-processing*, or template filling. Each stage of analysis applies a combination of linguistic, conceptual, and domain knowledge, as shown in Figure 1.

	Pre-processing	**Analysis**	**Post-processing**
Syntax	Tagging, bracketing	Parsing, attachment	Scoring
Semantics	Collocations, cluster analysis	Sense disambiguation, role mapping	Role extension
Domain Templates	Template activation	Ambiguity pruning, recovery	Default filling

Figure 1: Stages of data extraction

The pre-processor uses lexico-semantic patterns to perform some initial segmentation of the text, identifying phrases that are template activators, filtering out irrelevant text, combining and collapsing some linguistic constructs, and marking portions of text that could describe discrete events. This component is described in [1]. Linguistic analysis combines parsing and word sense-based semantic interpretation with domain-driven conceptual processing. The programs for linguistic analysis are largely those explained in [2, 3]—the changes made for MUC-4 involved mainly some additional mechanisms for recovering from failed processing and heavy pruning of spurious parses. Post-processing includes the final selection of templates and mapping semantic categories and roles onto those templates. This component used the basic elements from MUCK-II, adding a number of specialized rules for handling guerrilla warfare, types, and refines the discourse structures to perform the template splitting and merging required for MUC-3 and MUC-4.

The control flow of the system is primarily from linguistic analysis to conceptual interpretation to domain interpretation, but there is substantial feedback from conceptual and domain interpretation to linguistic analysis. The MUC-4 version of the Toolset includes a version of a strategy called *relation-driven control*, which helps to mediate between the various knowledge sources involved in interpretation. Basically, relation-driven control gives each linguistic relation in the text (such as subject-verb, verb-complement, or verb-adjunct) a preference score based on its interpretation in context. Because these relations can apply to a great many different surface structures, relation-driven control provides a means of combining preferences without the tremendous combinatorics of scoring many complete parses. Effectively, relation-driven control permits a "beam" strategy for considering multiple interpretations without producing hundreds or thousands of new paths through the linguistic chart.

The knowledge base of the system, consisting of a feature and function (unification-style) grammar with associated linguistic relations, and a core sense-based lexicon, still proves transportable and largely generic. The core lexicon contains over 10,000 entries, of which 37 are restricted because of specialized usage in the MUC-4 domain (such as *device*, which always means a bomb, and *plant*, which as a verb usually means to place a bomb and as a noun usually means the target of an attack). The core grammar contains about 170 rules, with 50 relations and 80 additional subcategories. There were 23 MUC-specific additions to this grammatical knowledge base, including 8 grammar rules, most of them dealing with unusual noun phrases that describe organizations in the corpus.

The control, pre-processing, and transportable knowledge base were all extremely successful for MUC-4; remarkably, lexical and grammatical coverage, along with the associated problems in controlling search and selecting among interpretations, proved not to be the major stumbling blocks for our system. While the program rarely produce an incorrect answer as a result of a sentence interpretation error, it frequently fails to distinguish multiple events, resolve vague or subtle references, and pick up subtle clues from non-key sentences. These are the major areas for future improvements in MUC-like tasks.

ANALYSIS OF TST2-0048

Overview of Example

TST2-0048 is faily representative of how the NLTOOLSET performed on MUC-4. The program successfully interpreted most of the key sentences but missed some references and failed to tie some additional information in to the main event. As a result, it filled two templates for what should have been one event and missed some additional fills. The program thus derived 53 slots out of a possible 52, with 34 correct, 19 missing, and 19 spurious for .65 recall, .64 precision, and .35 overgeneration. We made no special effort to adapt the system or fix problems for this particular example; in fact, we used TST2 as a "blind" test and did not do any development on that set at all.

Detail of Message Run

This example is actually quite simple at the sentence level [1]: The sentences are fairly short and grammatical, especially when compared to some of the convoluted propaganda stories, and TRUMP had no real problems with them. The story is difficult from a discourse perspective, because it returns to the main event (the attack on Alvarado) essentially without any cue after describing a background event (the attack on Merino's home). In addition, the story is difficult and a bit unusual in the implicit information that is captured in the answer key—that the seven children, because they were home when Merino's house was attacked, are targets. Most of the difference between our system's response and the correct templates was due to these two story-level problems.

The program made one or two other minor mistakes; for example, it was penalized for filling in "INDIVIDUAL" as a perpetrator (from the phrase *AN INDIVIDUAL PLACED A BOMB ON THE ROOF OF THE ARMORED VEHICLE*), an apparently correct fill that could have been resolved to "URBAN GUERRILLAS". It missed the SOME DAMAGE effect for the vehicle, which should have been inferred from the fact that the story later says the roof of the vehicle collapsed.

The system correctly parsed most of the main sentences, correctly linked the accusation in the first sentence to the murder of the Attorney General in the same sentence, and correctly separated the second event, which was distinguished by the temporal expression *5 days ago*.

As explained earlier, the Toolset uses pattern matching for pre-processing, followed by discourse processing, parsing and semantic interpretation, and finally template-filling. The pre-processor in this example filters out most of the irrelevant sentences (and, in this case, two of the relevant ones), recognizes most of the compound names (e.g. SALVADORAN PRESIDENT-ELECT ALFREDO CRISTIANI and ATTORNEY GENERAL ROBERTO GARCIA ALVARADO). The pre-processor marks phrases that activate templates (such as A BOMB PLACED and CLAIMED CREDIT), brackets out phrases like source and location (ACCORDING TO CRISTIANI and IN DOWNTOWN SAN SALVADOR), and tags a few words with part-of-speech to help the parser (e.g. auxiliaries (HAS), complementizers (THAT), and certain verbs following "to" (COLLAPSE)).

The last stage of pre-processing is a discourse processing module, which attempts a preliminary segmentation of the input story using temporal, spatial, and other cues, event types, and looking for certain definite and indefinite descriptions of events. In this case, the module identifies five potential segments. The first three turn out to be different descriptions of the same event (the killing of Alvarado), but they are later correctly merged into one template. The fourth segment is correctly identified as a new event (the attack on Merino's home). The fifth segment (describing the injury to Alvarado's bodyguards) is correctly treated as a new description, but is never identified as being part of the same event as the attack on Alvarado.

Linguistic analysis parses each sentence and produces (possibly alternative) semantic interpretations at the sentence level. These interpretations select word senses and roles, heavily favoring domain-specific senses. The parser did fail in one important sentence in TST2-0048: In the sentence "A 15-YEAR-OLD NIECE OF

[1] See Appendix F for the text and answer templates for the example.

MERINO'S WAS INJURED", it could not parse the apostrophe-s construct. This was a harmless failure because it occurs between a noun phrase and a verb phrase, and one of the parser's recovery strategies attaches any remaining compatible fragments that will contribute to a template fill.

The interpretation of each sentence is interleaved with domain-driven analysis. The conceptual analyzer, TRUMPET, takes the results of interpreting each phrase and tries to map them onto domain-based expectations, determining, for example, the appropriate role for the FMLN in "ACCUSED THE FMLN" as well as associating "support" events (such as accusations and effects) with main events (such as attacks or bombings). Because the discourse pre-processing module is prone to error, TRUMPET has begun to play a major role in resolving references as well as in guiding semantic interpretation.

Post-processing maps the semantic interpretations onto templates, eliminating invalid fills (in this case none), combining certain multiple references (in the attack on Alvarado), and "cleaning up" the final output.

Interpretation of Key Sentences

The TRUMP parser of the NLTOOLSET successfully parsed and interpreted the first sentence (S1) and correctly applied conjunction reduction to get Cristiani as the accuser and get the "SUSPECTED OR ACCUSED BY AUTHORITIES" fill. Embedded clauses are typically handled in much the same way as main clauses, except that the main clauses often add information about the CONFIDENCE slot. The system correctly treats the main event and the accusing as a single event, in spite of ignoring the definite reference "THE CRIME". In our system, linking an accusation (C-BLAME-TEMPLATE in the output below) to an event is the default.

The following is the pre-processed input and final sentence-level interpretation of S1:

Pre-processed input:

```
[byline: SAN SALVADOR, 19 APR 89 (ACAN-EFE) --] [bracket: [TEXT]] [fullname: SALVADORAN PRESIDENT-ELECT
ALFREDO CRISTIANI] CONDEMNED THE TERRORIST KILLING OF [fullname: ATTORNEY GENERAL ROBERTO GARCIA
ALVARADO] AND (comp: ACCUSED THE FARABUNDO MARTI NATIONAL LIBERATION FRONT [bracket: (FMLN)] OF) THE CRIME.
```

Interpretation:

```
Calling Trumpet with FINAL Interpretation:
(COORDCONJ_AND1
  (R-PART
    (VERB_ACCUSE1 (R-REL-TIME *PAST*)
                  (R-PATIENT
                    (TERRORIST-NAME_FMLN1 (R-NAME FMLN)
                                          (R-PART (C-ENTITY))))
                  (R-COMMUNICATOR
                    (FULLNAME_ALFREDO-CRISTIANI*1 (R-NUMBER *SINGULAR*)
                                                  (R-NAME ALFREDO-CRISTIANI*)
                                                  (R-NATIONALITY
                                                    (C-NATION-NAME_EL-SALVADOR1-QUAL))))
                  (R-ACCUSATION
                    (NOUN_CRIME1 (R-NUMBER *SINGULAR*)
                                 (R-DEFINITE (DET_THE1))))))
  (R-PART
    (VERB_CONDEMN1 (R-REL-TIME *PAST*)
                   (R-PATIENT
                     (C-ACT-OF-VERB_KILL1 (R-NUMBER *SINGULAR*)
                                          (R-EFFECT (C-DEAD-QUAL))
                                          (R-DEFINITE (DET_THE1))
                                          (R-CAUSE
                                            (C-VERB_TERRORIZE1-ER
                                              (R-NUMBER *SINGULAR*)
                                              (R-INHERENT-ACTIVITY
                                                (VERB_TERRORIZE1))))
                                          (R-EFFECTED
                                            (FULLNAME_ROBERTO-GARCIA-ALVARADO1
                                              (R-NUMBER *SINGULAR*)
                                              (R-NAME ROBERTO-GARCIA-ALVARADO)))))
                   (R-COMMUNICATOR
                     (FULLNAME_ALFREDO-CRISTIANI*1 (R-NUMBER *SINGULAR*)
```

```
                                                   (R-NAME ALFREDO-CRISTIANI*)
                                                   (R-NATIONALITY
                                                     (C-NATION-NAME_EL-SALVADOR1-QUAL))))))
TRUMPET WARN: Splitting connective COORDCONJ_AND1 into parts
Activating new sense
(C-BLAME-TEMPLATE (R-REL-TIME *PAST*)
                  (R-MODALITY (C-QUALIFIER))
                  (R-POLARITY (C-QUALIFIER))
                  (R-PERPETRATOR
                    (TERRORIST-NAME_FMLN1 (R-NAME FMLN)
                                          (R-PART (C-ENTITY))))
                  (R-ACCUSATION
                    (NOUN_CRIME1 (R-NUMBER *SINGULAR*)
                                 (R-DEFINITE (DET_THE1))))
                  (R-ACCUSER
                    (FULLNAME_ALFREDO-CRISTIANI*1 (R-NUMBER *SINGULAR*)
                                                  (R-NAME ALFREDO-CRISTIANI*)
                                                  (R-NATIONALITY
                                                    (C-NATION-NAME_EL-SALVADOR1-QUAL)))))
Activating new sense
(C-REPORT-TEMPLATE (R-REL-TIME *PAST*)
                   (R-MODALITY (C-QUALIFIER))
                   (R-POLARITY (C-QUALIFIER))
                   (R-OBJECT
                     (C-DEATH-TEMPLATE (R-NUMBER *SINGULAR*)
                                       (R-MODALITY (C-QUALIFIER))
                                       (R-POLARITY (C-QUALIFIER))
                                       (R-DEFINITE (DET_THE1))
                                       (R-TARGET
                                         (FULLNAME_ROBERTO-GARCIA-ALVARADO1
                                           (R-NUMBER *SINGULAR*)
                                           (R-NAME ROBERTO-GARCIA-ALVARADO)))
                                       (R-PERPETRATOR
                                         (C-VERB_TERRORIZE1-ER
                                           (R-NUMBER *SINGULAR*)
                                           (R-INHERENT-ACTIVITY
                                             (VERB_TERRORIZE1))))))
                   (R-REPORTER
                     (FULLNAME_ALFREDO-CRISTIANI*1 (R-NUMBER *SINGULAR*)
                                                   (R-NAME ALFREDO-CRISTIANI*)
                                                   (R-NATIONALITY
                                                     (C-NATION-NAME_EL-SALVADOR1-QUAL)))))
TRUMPET WARN: Breaking out core templates (C-DEATH-TEMPLATE)
....
TRUMPET WARN: Linking (special) C-REPORT-TEMPLATE as filler for R-SUPPORT of C-DEATH-TEMPLATE
TRUMPET WARN: Linking (special) C-BLAME-TEMPLATE as filler for R-SUPPORT of C-DEATH-TEMPLATE
Adding TERRORIST-NAME_FMLN1 from C-BLAME-TEMPLATE to R-PERPETRATOR of C-DEATH-TEMPLATE
```

The next set of examples sentences (S11-13) are more difficult. There was one parser failure, with a successful recovery. As we have mentioned, we correctly identify this as a new event based on temporal information, but filter out S12 because it has no explicit event reference. This is not a bug—this sort of implicit target description is fairly infrequent, so we chose not to address it at this stage.

Pre-processed input:

```
GUERRILLAS ATTACKED MERINO'S HOME (location: IN SAN SALVADOR) [ago: 5 DAYS AGO] WITH EXPLOSIVES.
[filtered: THERE WERE SEVEN CHILDREN, INCLUDING FOUR OF THE VICE PRESIDENT'S CHILDREN, IN THE HOME AT THE
TIME.] A (age: 15-YEAR-OLD) NIECE OF MERINO'S WAS INJURED.
```

Interpretation:

```
Calling Trumpet with FINAL Interpretation:
(VERB_ATTACK1 (R-REL-TIME *PAST*)
              (R-INSTRUMENT (NOUN_EXPLODE-IVE-X (R-NUMBER *PLURAL*)))
              (R-PATIENT
                (NOUN_HOME1 (R-NUMBER *SINGULAR*)
                            (R-OBJECTHOLDER
```

```
                          (FULLNAME_FRANCISCO-MERINO*1
                            (R-NUMBER *SINGULAR*)
                            (R-NAME FRANCISCO-MERINO*)))))
          (R-DATE
            (C-DATE-OF-OCCURRENCE (R-RELATIVE NO)
                                  (R-YEAR |89|)
                                  (R-DAY |14|)
                                  (R-MONTH |4|)))
          (R-AGENT (NOUN_GUERRILLA1 (R-NUMBER *PLURAL*)))
          (R-LOCATION (CITY-NAME_SAN-SALVADOR1 (R-NAME SAN-SALVADOR))))
Activating new sense
(C-ATTACK-TEMPLATE (R-REL-TIME *PAST*)
                   (R-MODALITY (C-QUALIFIER))
                   (R-POLARITY (C-QUALIFIER))
                   (R-LOCATION (CITY-NAME_SAN-SALVADOR1 (R-NAME SAN-SALVADOR)))
                   (R-DATE
                     (C-DATE-OF-OCCURRENCE (R-RELATIVE NO)
                                           (R-YEAR |89|)
                                           (R-DAY |14|)
                                           (R-MONTH |4|)))
                   (R-INSTRUMENT (NOUN_EXPLODE-IVE-X (R-NUMBER *PLURAL*)))
                   (R-TARGET
                     (NOUN_HOME1 (R-NUMBER *SINGULAR*)
                                 (R-LOCATION
                                   (CITY-NAME_SAN-SALVADOR1
                                     (R-NAME SAN-SALVADOR)))
                                 (R-OBJECTHOLDER
                                   (FULLNAME_FRANCISCO-MERINO*1
                                     (R-NUMBER *SINGULAR*)
                                     (R-NAME FRANCISCO-MERINO*)))))
                   (R-PERPETRATOR (NOUN_GUERRILLA1 (R-NUMBER *PLURAL*))))


Calling Trumpet with FRAGMENT Interpretation:
(VERB_INJURE1 (R-REL-TIME *PAST*)
              (R-EFFECT (C-INJURY))
              (R-EFFECTED
                (NOUN_NIECE1 (R-NUMBER *SINGULAR*)
                             (R-DEFINITE (DET_A1))
                             (R-POSSESSES
                               (FULLNAME_FRANCISCO-MERINO*1
                                 (R-NUMBER *SINGULAR*)
                                 (R-NAME FRANCISCO-MERINO*))))))
Activating new sense
(C-INJURY-TEMPLATE (R-REL-TIME *PAST*)
                   (R-MODALITY (C-QUALIFIER))
                   (R-POLARITY (C-QUALIFIER))
                   (R-TARGET
                     (NOUN_NIECE1 (R-NUMBER *SINGULAR*)
                                  (R-DEFINITE (DET_A1))
                                  (R-POSSESSES
                                    (FULLNAME_FRANCISCO-MERINO*1
                                      (R-NUMBER *SINGULAR*)
                                      (R-NAME FRANCISCO-MERINO*))))))

...
TRUMPET WARN: Linking (special) C-INJURY-TEMPLATE as filler for R-TARGET-EFFECT of C-BOMBING-TEMPLATE
```

The system filters S21 (this is an omission, because "ESCAPED UNSCATHED" should be recognized as an effect), but successfully interprets S22 and resolves "ONE OF THEM" to "BODYGUARDS". Note that it is the pronoun "THEM", not "ONE", that gets resolved, using a simple reference resolution heuristic that looks for the most recent syntactically and semantically compatible noun phrase. However, this action results in a penalty rather than a reward because the system does not tie the injury to the attack on Alvarado at the beginning of the story.

Pre-processed input:

```
[filtered: ACCORDING TO THE POLICE AND GARCIA ALVARADO'S DRIVER, WHO
ESCAPED UNSCATHED, THE ATTORNEY GENERAL WAS TRAVELING WITH TWO
BODYGUARDS.]  [numword: ONE] OF THEM WAS INJURED.
```

Interpretation:

```
Calling Trumpet with FINAL Interpretation:
(VERB_INJURE1 (R-REL-TIME *PAST*)
              (R-EFFECT (C-INJURY))
              (R-EFFECTED
                (C-NUMBER (R-VALUE |1|)
                          (R-WHOLE (PNOUN_THEM1 (R-NUMBER *PLURAL*))))))
Activating new sense
(C-INJURY-TEMPLATE (R-REL-TIME *PAST*)
                   (R-MODALITY (C-QUALIFIER))
                   (R-POLARITY (C-QUALIFIER))
                   (R-TARGET
                     (C-NUMBER (R-VALUE |1|)
                               (R-WHOLE (PNOUN_THEM1 (R-NUMBER *PLURAL*))))))
Applying TR_WHOLE-OF-PARTS transform on NUMWORD_ONE1 for (AND (OR C-HUMAN C-HUMAN-GROUP NOUN_BODY3) EXP-NEG-TARGET)

Adding PNOUN_THEM1 from C-INJURY-TEMPLATE to R-TARGET of C-BOMBING-TEMPLATE
TRUMPET DANGER: Resolving exp PNOUN_THEM1 to :EXISTING TOK BODYGUARDS
TRUMPET DANGER: Resolving exp PNOUN_THEM1 to :EXISTING TOK BODYGUARDS
```

Comparison of Program Answers with Answer Key

The NLTOOLSET results for TST2-0048 were the following templates (Annotations have been added in lower case preceded by %, and blank slot (-) fills have been deleted to save space).

```
0.  MESSAGE: ID                     TST2-MUC4-0048
1.  MESSAGE: TEMPLATE               1
2.  INCIDENT: DATE                  - 19 APR 89
3.  INCIDENT: LOCATION              EL SALVADOR: SAN SALVADOR (CITY)
4.  INCIDENT: TYPE                  BOMBING
5.  INCIDENT: STAGE OF EXECUTION    ACCOMPLISHED
6.  INCIDENT: INSTRUMENT ID         "BOMB"
7.  INCIDENT: INSTRUMENT TYPE       BOMB: "BOMB"
8.  PERP: INCIDENT CATEGORY         TERRORIST ACT
9.  PERP: INDIVIDUAL ID             "URBAN GUERRILLAS"
                                    "INDIVIDUAL" % spurious fill
10. PERP: ORGANIZATION ID           "FARABUNDO MARTI NATIONAL LIBERATION FRONT"
11. PERP: ORGANIZATION_CONFIDENCE   SUSPECTED OR ACCUSED BY AUTHORITIES: "FARABUNDO MARTI NATIONAL LIBERATION FRONT"
12. PHYS TGT: ID                    "HIS VEHICLE"
13. PHYS TGT: TYPE                  TRANSPORT VEHICLE: "HIS VEHICLE"
14. PHYS TGT: NUMBER                1: "HIS VEHICLE"
16. PHYS TGT: EFFECT OF INCIDENT    - % missed SOME DAMAGE: "HIS VEHICLE"
18. HUM TGT: NAME                   "ROBERTO GARCIA ALVARADO"
19. HUM TGT: DESCRIPTION            "ATTORNEY GENERAL": "ROBERTO GARCIA ALVARADO"
                                    % missed fills DRIVER
                                    % missed fills BODYGUARDS
                                    "ATTORNEY GENERAL" % spurious fill
20. HUM TGT: TYPE                   GOVERNMENT OFFICIAL: "ROBERTO GARCIA ALVARADO"
                                    LEGAL OR JUDICIAL: "ATTORNEY GENERAL" % spurious
                                    % missed fills CIVILIAN: "DRIVER"
                                    % missed fills SECURITY GUARD: "BODYGUARDS"
21. HUM TGT: NUMBER                 1: "ROBERTO GARCIA ALVARADO"
                                    1: "ATTORNEY GENERAL" % spurious
                                    % missed fills: 1: "DRIVER"
                                    % missed fills: 2: "BODYGUARDS"
                                    % missed fills: 1: "BODYGUARDS"
23. HUM TGT: EFFECT OF INCIDENT     DEATH: "ATTORNEY GENERAL" % spurious
                                    DEATH: "ROBERTO GARCIA ALVARADO"
                                    % missed fills: NO INJURY: "DRIVER"
                                    % missed fills: INJURY: "BODYGUARDS"
24. HUM TGT: TOTAL NUMBER           -
```

```
0.  MESSAGE: ID                       TST2-MUC4-0048
1.  MESSAGE: TEMPLATE                 2
2.  INCIDENT: DATE                    14 APR 89
3.  INCIDENT: LOCATION                EL SALVADOR: SAN SALVADOR (CITY)
4.  INCIDENT: TYPE                    BOMBING
5.  INCIDENT: STAGE OF EXECUTION      ACCOMPLISHED
6.  INCIDENT: INSTRUMENT ID           "EXPLOSIVES"
7.  INCIDENT: INSTRUMENT TYPE         EXPLOSIVE: "EXPLOSIVES"
8.  PERP: INCIDENT CATEGORY           TERRORIST ACT
9.  PERP: INDIVIDUAL ID               "GUERRILLAS"
10. PERP: ORGANIZATION ID             "FARABUNDO MARTI NATIONAL LIBERATION FRONT"
11. PERP: ORGANIZATION CONFIDENCE     SUSPECTED OR ACCUSED BY AUTHORITIES: "FARABUNDO MARTI NATIONAL LIBERATION FRONT"
12. PHYS TGT: ID                      "MERINO'S HOME"
13. PHYS TGT: TYPE                    CIVILIAN RESIDENCE: "MERINO'S HOME" % missed type should be GOVERNMENT OFFICE OR RESIDEN
14. PHYS TGT: NUMBER                  1: "MERINO'S HOME"
19. HUM TGT: DESCRIPTION              "NIECE OF MERINO"
                                     % missed fill -  "CHILDREN"
                                      % missed fill - "VICE PRESIDENT'S CHILDREN"
20. HUM TGT: TYPE                     CIVILIAN: "NIECE OF MERINO"
                                      % missed fill CIVILIAN: "CHILDREN"
                                      % missed fill CIVILIAN: "VICE PRESIDENT'S CHILDREN"
21. HUM TGT: NUMBER                   1: "NIECE OF MERINO"
                                      % missed fill  7: "CHILDREN"
                                      % missed fill 4: "VICE PRESIDENT'S CHILDREN"
23. HUM TGT: EFFECT OF INCIDENT       INJURY: "NIECE OF MERINO"
                                      DEATH: "NIECE OF MERINO"
24. HUM TGT: TOTAL NUMBER             - % missed fill 7

% completely spurious template; should have been merged with template 1
0.  MESSAGE: ID                       TST2-MUC4-0048
1.  MESSAGE: TEMPLATE                 3
2.  INCIDENT: DATE                    - 19 APR 89
3.  INCIDENT: LOCATION                EL SALVADOR
4.  INCIDENT: TYPE                    BOMBING
5.  INCIDENT: STAGE OF EXECUTION      ACCOMPLISHED
6.  INCIDENT: INSTRUMENT ID           "BOMB"
7.  INCIDENT: INSTRUMENT TYPE         BOMB: "BOMB"
8.  PERP: INCIDENT CATEGORY           TERRORIST ACT
10. PERP: ORGANIZATION ID             "FARABUNDO MARTI NATIONAL LIBERATION FRONT"
11. PERP: ORGANIZATION CONFIDENCE     POSSIBLE: "FARABUNDO MARTI NATIONAL LIBERATION FRONT"
16. PHYS TGT: EFFECT OF INCIDENT      DESTROYED: "-"
19. HUM TGT: DESCRIPTION              "BODYGUARDS"
20. HUM TGT: TYPE                     SECURITY GUARD: "BODYGUARDS"
21. HUM TGT: NUMBER                   PLURAL: "BODYGUARDS"
23. HUM TGT: EFFECT OF INCIDENT       INJURY: "BODYGUARDS"
```

Some of the missing information in the response template comes from failing to tie information in to the main event or failing to recover implicit information. This is the case with the damage to the vehicle, which is described in passing, the children who were in Merino's home, and the driver who escaped unscathed. Almost all the rest of the departures owe to some aspect of reference resolution—from failing to recognize the injury to the bodyguards as part of Alvarado's murder, to the extra fills "INDIVIDUAL" and "ATTORNEY GENERAL" that were co-referential with others. One of these turned out to be a simple bug, in that the title "ATTORNEY GENERAL" in our system was interpreted as a different type (GOVERNMENT OFFICIAL) from the noun phrase "ATTORNEY GENERAL" (LEGAL OR JUDICIAL); thus the system failed to unify the references. However, the general problem of reference resolution is certainly one of the main areas where future progress can come.

The other illustrative problem with this example is the degree to which relatively inconsequential facts can be pieced together into an interpretation. There is no theoretical reason why our system didn't know about different forms of damage to vehicles, but we certainly wouldn't want to spend a lot of time encoding this sort of knowledge. This turned out to be a rather tedious part of the MUC task. We did go so far as to have template filling heuristics, for example, that tell the system: (1) When vehicles explode near buildings, it is the buildings and not the vehicles that are the targets, (2) When parts of buildings are destroyed or damaged (e.g. "the bomb shattered windows") this means that the buildings sustained some damage, and

(3) When body parts are damaged (e.g. "the bomb destroyed his head"), it is the owner of the body parts that is affected. However, such rules only scratch the surface of the reasoning that contributes to template filling.

While the reference resolution problem is quite general and very interesting from a research perspective, the reasoning problem seems more MUC-specific, and it's hard to separate general reasoning issues from the peculiar details of the fill rules.

Aside from these problems, our system performed pretty well on this example, as for MUC on the whole. The recall and precision for this message were both over .60, with the program recovering most of the information from the text. As is typical from our MUC experience, the local processing of sentences was very accurate and complete, while the general handling of story level details and template filling had some loose ends.

SUMMARY AND CONCLUSION

MUC-4 is a very difficult task, combining language interpretation at many levels with a variety of rules and strategies for template filling. The examples here illustrate some of the important characteristics of our system as well as where future progress can be made. Not surprisingly, the major problems that remain after MUC-4 are very similar to the ones that we identified at the end of MUC-3. This by itself might seem discouraging, but the fact that the system did much better on MUC-4 suggests that we can expect more improvements in the future. While there is a class of phenomena that we haven't really begun to address (the body of world knowledge that contributes to interpreting events), there is also the ripe problem of interpreting text in context, in which MUC has given the field a leg up.

References

[1] Paul S. Jacobs, George R. Krupka, and Lisa F. Rau. Lexico-semantic pattern matching as a companion to parsing in text understanding. In *Fourth DARPA Speech and Natural Language Workshop*, pages 337–342, San Mateo, CA, February 1991. Morgan-Kaufmann.

[2] Paul Jacobs and Lisa Rau. SCISOR: Extracting information from on-line news. *Communications of the Association for Computing Machinery*, 33(11):88–97, November 1990.

[3] Paul S. Jacobs. TRUMP: A transportable language understanding program. *International Journal of Intelligent Systems*, 7(3):245–276, 1992.