WAYNE STATE UNIVERSITY: DESCRIPTION of the UNO NATURAL LANGUAGE PROCESSING SYSTEM as USED for MUC-6

Lucja Iwańska, Mary Croll, Taewan Yoon, Maria Adams
Wayne State University
Department of Computer Science
Detroit, Michigan 48202, USA
lucja@cs.wayne.edu
ph. (313) 577-1667, fx: (313) 577-6868

INTRODUCTION

Our Research Hypothesis

The UNO natural language processing (NLP) system implements a Boolean algebra computational model of natural language [Iwańska, 1992] [Iwańska, 1993] [Iwańska, 1994] [Iwańska, 1996b] and reflects our research hypothesis that natural language is a very expressive, yet computationally tractable knowledge representation and reasoning system with its own representational and inferential machinery. One of our goals is to experimentally demonstrate that an NLP system that closely parallels the representational and inferential characteristics of natural language allows one to achieve an in-depth processing (eg., querring automatically from texts created knowledge bases or entity classification), with close-to-real-time, high-recall-and-precision performance.

Why and What for MUC-6

MUC6 Tasks Not Inconsistent with Our Goals

We view the MUC6 tasks as not inconsistent with our goal of providing further experimental evidence in support of our research hypothesis. Also, we have been conducting experiments in non-committal processing with controlled resource allocation such as exploiting local interpretations, taking advantage of local contexts, and performing expensive computations such as disambiguation only if needed. We have also been experimenting with flexible processing such as undoing various decisions and independent processing of related tasks. The NE MUC-6 task is a good test for such experiments.

Coreference– Most Interesting Task

We were particularly interested in the coreference task, with the NE task largely considered as a preparation stage for it because:

1. Reference resolution is critical for processing pretty much any text in any domain;
2. We believe that mechanism for resolving references to named entities is basically the same regardless of the entities' types. Identifying certain named entities and therefore resolving references to them may be facilitated by the availability of extra constraints on the beginning or the end of the name. For example, identifying personal and company names is facilitated by the fact that personal names often start with a title such as "*Ms.*" and company names often end with a corporate extension such as "*Ltd*".

3. We are in the process of demonstrating that it is natural language for acquiring knowledge, and not acquiring knowledge in order to process natural language. Our very encouraging preliminary results show that definite anaphora facilitates acquisition of taxonomic knowledge [Iwańska, 1996a], including the "*type-subtype*" and "*part-subpart*" relations. These results demonstrate that approaches to definite anaphora resolution that rely primarily on the availability of such knowledge, a standard approach in all NLP systems we are aware of, are misguided.

Successful Techniques for Developing and Testing In-Depth Processing Large Number of Texts

Finally, by participating in MUC6 we hoped to share our experience and learn from the other participants about successful techniques for developing and testing in-depth processing for a large, more than two thousands, number of texts and for developing large knowledge bases.

UNO MODEL OF NATURAL LANGUAGE

Closely Mimics Representation and Reasoning Inherent in Natural Language

The UNO model closely mimics the representation and reasoning inherent in natural language because its translation procedure, data structures of the representation, and inference engine are motivated by the semantics and pragmatics of natural language. One payoff of this close correspondence to natural language is the capability of automatically creating and querring knowledge bases from textual documents[1].

Sentences asserting properties of (sets of) individuals, sentences describing subtyping relations, including extentional type definitions, as well as intensional definitions of concepts such as

1. "*John is a neither good nor hard-working nurse*", "*Not many students did well*"
2. "*Dobermans, poodles and terriers are dogs*"
3. "*Elephant- a huge, thick-skinned, mammal with very few hairs, with a long, flexible snout, and two ivory tusks*"

are uniformly represented by the following type equations:

$type == \{ < P_1, TP_1 >, < P_2, TP_2 >, \ldots, < P_n, TP_n > \}$

Their left-hand side, "type", is the representation of a noun phrase, the largest type, or the name of a concept; the right-hand side is a two-element set: a property value P, and TP- a set of $< t, p >$ elements representing the fact that the property value P holds at a temporal interval t with the probability p.

Individual property values, temporal intervals and probabilities are represented by the sets $[a_1, a_2, \ldots, a_n]$ whose elements a_i are terms, record-like structures consisting of a *head*, a type symbol, and a *body*, a list of attribute-value pairs: *attribute => value*. For example, a complex noun "*sick, very unhappy woman*" is represented by

[*woman*(*health* => *sick*,
happy => (*not happy*)(*degree* => *very*))]

whose only term has the type "*woman*" as its head and two attributes: "*health*" with the value *sick*; "*happy*" with the value (*not happy*)(*degree* => *very*). Semantically, this data structure represents this subset of individuals of the type "*woman*" for which the attribute "*health*" has the value "*sick*" and the function "*happy*" has the value "*very unhappy*".

Various relations such as entailment or its dual subsumption (set-inclusion) and negation (set-complement) are automatically computed, and the intuitively and formally correct results are guaranteed to hold. The UNO NLP system uses such type equations bi-directionally: for answering questions about the properties of a particular individual, and for matching particular properties against the properties of individuals in its knowledge base.

[1] Needless to say, limited by the subset of English covered by the model.

Solid Computational and Mathematical Framework in Tact with Linguistic Theories

The UNO representation offers a solid computational and mathematical framework in tact with linguistic theories. Updating knowledge base and automated inferencing is done by the same semantically clean computational mechanism of performing Boolean operations on the representation of natural language input and the representation of previously obtained information stored in the knowledge base.

The underlying knowledge representation formalisms with the computable Boolean algebras with set-theoretic and interval-theoretic semantics allows one to capture semantics of different syntactic categories because sets and intervals underlie semantics of many syntactic categories: common nouns, intransitive verbs, and adjectives can be thought of as denoting sets of persons or objects that possess properties denoted by the words; adjectives and adverbs are functions mapping sets of objects into sets of objects; determiners are functions mapping sets of objects into sets of sets of objects, and the denotations of proper nouns are sets of sets of objects [Dowty et al., 1981] [Barwise and Cooper, 1981] [Keenan and Faltz, 1985] [Hamm, 1989].

The same machinery is used as a metalanguage for describing and propagating arbitrary Boolean constraints, including dictionary entries describing morphological and grammatical constraints. The data structures are partially specified, negative constraints are propagated via unification, and the nonmonotonicity of negation [Pereira, 1987] is not problematic.

The UNO model shares many computational characteristics with the programming language LIFE [Aït-Kaci and Richard Meyer and Peter Van Roy, 1993] because the efficiently computable calculus that underlies LIFE [Aït-Kaci, 1986] is extended in the UNO model to handle negation and generalized quantifies.

Some of the linguistic theories that the UNO model encompasses and (or) extends include insights of the Montague semantics of natural language [Montague, 1973] [Dowty et al., 1981], the Boolean algebra mathematical models of [Keenan and Faltz, 1985], the theory of generalized quantifiers [Barwise and Cooper, 1981] [Hamm, 1989], the theory of the pragmatic inference of quantity-based implicature of [Horn, 1972] [Horn, 1989], and the theory of negation in natural language of [Horn, 1989].

Recent Extension– Temporal Reasoning

We have recently extended the UNO model to incorporate temporal reasoning. This extension demonstrates that important inferences about time can be captured by a general representation and reasoning mechanism inherent in natural language, many aspects of which are closely mimicked by the UNO model. We have shown that computing logical, context-independent and non-monotonic, context-dependent inferences for temporal and non-temporal objects is almost exactly analogous.

Theory and Practise

We are committed to addressing research problems with a strong promise for facilitating processing natural language input. For example, we had decided to extena the UNO model of natural language to handle temporal information because virtually all real-life tasks involve handling some aspects of time. There is a large body of existing work on morphologically marked time and aspect, but we had decided against handling this type of temporal information because it necessarily requires high recall and precision of performing sentential-level parsing, a task that no NLP system, including our system, can perform well. Instead, we had decided to address temporal information from explicit temporal expressions because we can extremely reliably recover such expressions via local parsing.

Our natural-language-based temporal reasoner was developed and tested on more than three hundreds 1989 "*Wall Street Journal*" articles. We somewhat randomly chose a batch of 300 WSJ articles and using SGML-like marks, an opening mark $< TE >$ and a closing mark $< /TE >$, we had marked all expressions of different syntactic categories that contained any information pertaining to time. Incidentaly, the number of articles coincides with the number of the MUC-6 development data.

Our temporal reasoner automatically extracts explicit temporal expressions from on-line textual documents and creates their representation. This representation allows the system to compute entailed logical, context-independent, deductive inferences and facilitates computing context-dependent, non-monotonic inferences, including implicature, specialization, and generalization.

For any set of English temporal expressions, their information content can be computed and compared, which allows the system to compute answers to the "*Yes-No*" questions about various aspects of time, answers to the "*When ?*", "*How long ?*" and "*How often ?*" queries of the resulting knowledge base and, to a limited extent, temporal ordering of the events described in the documents.

UNO NLP SYSTEM

In this section, we briefly comment on the most distinct characteristics of our system, as requested by the MUC-6 program committee; technical details can be found in the references provided earlier. Whenever possible, we illustrate our system's capabilities with the examples from the walkthrough article.

Demonstrated by the pre-MUC6 Research and Implementation

Reasoning with Explicit Negative, Disjunctive, and Conjunctive Information at All Syntactic Levels

One consequence of this rather unique capability is flat taxonomies; complex Boolean types need not be stored explicitly, which prevents the unnecessary, but common, exponential growth of a knowledge base. For example, the UNO system does not need to explicitly store the entailment (subsumption) relation between the complex disjunctive type "*power-boat or sail-boat*" and the lexically-simple type "*boat*" because this relation is directly (and cheaply) computed from the UNO representation of these two expressions, common nouns in this case.

Flat taxonomies are highly desirable because, among others, they facilitate the ease and quality increase of knowledge base maintenance.

Handling general negation in natural language allows the UNO system to correctly compute possible interpretations of the sentence "*Mr. Dooner doesn't see a creative malaise permeating the agency*", one of which is that Mr. Dooner sees something else. This correct interpretation is automatically preferred by the system because of the context created by the immediately following sentence "*He points to several campaigns with pride, including the Taster's Choice commercials that are like a running soap opera*".

Adverbial and Adjectival Modification at All Syntactic Levels

Adverbial and adjectival modification also contributes to the flatness of our taxonomies. For example, the UNO system automatically computes the relations between the following pairs of expressions denoting types and qualitative frequency values:

"*smooth process*"	"*not very smooth process*"
"*no immediate plans*"	"*plans*"
"*huge models*"	"*small models*"
"*unusual*"	"*not extremely unusual*"

Nonstandard Quantifiers

Some of the nonstandard quantifiers that our system can handle include vague quantifiers involving the determiner "*many*". The UNO system automatically computes relations between the following expressions:

"*many differences*" "*very many differences*" "*not very many differences*"

Reasoning with Uncertainty and Qualitative Probabilistic Reasoning without Underlying Numeric Values

Just like the other systems participating in MUC-6, our system often misses a high-level relation between entities described in a sentence because our parser does not attempt or fails to compute full sentential-level parsing. However, even without such a quality parse, the system is capable of automatically computing relations such as the relation between the following two expressions

"*possible acquisitions*" "*possible, but not very likely, acquisitions*"

While the lack of a quality parse may prevent the system from understanding a high-level relation, in this case, what exactly was said about possible acquisitions, it understands the difference between "*possible, but not very likely*" and "*possible*" acquisitions.

Temporal Reasoning

The UNO system does not only identify explicit temporal expressions, but also automatically reasons with them. Some of the expressions uniquely handled by our system include the following:

1. **Temporal frequency quantifiers**

   "*Not very often*" "*Very often*"

2. **Infinite number of temporal relations**

   "*Not in the immediate future*" "*immediately precedes*"

   "*Shortly before or after*" "*Not long before*" "*Five or six days before*"

3. **Uncertain and underspecified (explicitly negative) temporal information**

   "*It did not happen on April 22, 1992*"

   "*X happened in April, 1992, or X happened in May, but not early May, 1992*"

Reasoning with Underspecified Information

The UNO system computes the fact that the following expressions differ in their information content and that the first has strictly more information than the second:
"*X happened in May, 1992*" "*X happened in May, but not early May, 1992*"

Automatic Combination of Information

Our system offers a highly efficient Boolean "*meet*" operation which is mathematically guaranteed to combine information in the most general way. This UNO operation provides an alternative to ad-hoc merging employed by many other systems. We must note, however, that currently this advantage of our system is rarely realized in practise because it strongly depends on a correct, quality parse of input sentences.

Uniform Natural-Language-Based Representation of Taxonomic and Temporal Reasoning

As we explain later, this uniformity of our representation greatly simplifies our architecture and control.

Work Done during June-October 1995

Below we elaborate on the work done between the end of June and beginning of October 1995. We briefly comment on the scope, importance and quantity of the tasks we had decided to do.

1. **Implemented changes necessary to participate in MUC6**

   For example, before MUC-6, the UNO system was preserving the exact image of the input text, but it did not keep track of the correspondence between the resulting knowledge base and the actual pieces of input text from which the knowledge base resulted. For MUC-6, we had to redesign our bookkeeping structures in order to be able to do this. We devoted a considerable effort to this task because such a change directly affects every processing stage.

   Another piece of code we had to develop were functions for chosing markables and outputting SGML-tagged text.

2. **Improved the accuracy of identifying unmarked sentential boundaries**

   Our pre-MUC-6 system was quite good in correctly identifying sentential boundaries in newspaper articles. A simple list of 200-or-so standard abbreviations and the sensitivity to the most common occurences of periods in numbers was largely responsible for this good performance. We wanted the system to perform this task near-perfectly because it would improve this generally needed capability and because based on the existing literature, we expected the distance expressed in a number of sentences to be a very important factor in computing pronominal referents.

3. **Improved handling punctuation**

   It is fair to say that before MUC-6, we largely ignored, as opposed to handled, punctuaction. Now we can say that we can handle periods, exclamations, most unpaired singlequotes, most commas, and some dashes. While handling punctuation is good in general, specifically for MUC-6, it is needed for processing numbers and facilitating identification of apposites.

4. **Completed half-built semantic hierarchy of structured geographical knowledge.**

   The UNO NLP hierarchy of geographical knowledge contains major geographical information about all countries, including capital cities, major and important cities, towns, ports, suburbs, local settlements, geographical and political regions that divide land such as provinces, islands, major ports and airports, landmarks, monetary, length, area, and volume systems, official languages, major political organizations, waters such as seas, lakes, and rivers, and geographical landmarks and points of interest such as mountains, hills, woods, and national parks.

   This geographical knowledge is encoded in our uniform, general-purpose UNO knowledge representation; UNO NLP system supports geographical reasoning with its general inferencing mechanism.

   We certainly did not need to do this for MUC-6, a simple gazetteer list would do (an approach adopted by most MUC-6 participants). However, we put a lot of effort into encoding this hierarchy because:

   (a) With the existence of this hierarchy, we further substantiate our claim that natural language is a powerful and efficient knowledge representation system and add geographical knowledge to the list of uniformly represented and reasoned about types of knowledge. Right now, our system can reason about geographical region containment in the exact analogous fashion as about type subset relation and temporal interval containment.

   (b) We wanted to demonstrate that the same in-depth mechanism of some aspects of geographical reasoning can be efficiently used to perform a much lesser, MUC-6-like task of marking locations.

5. **Identified common named types, including organization types and existing named entities**

   We identified more than 100 types other than the type "*organization*" and developed sizable knowledge bases and dictionaries with the actual existing, classified named entities of these types. We decided to do it in order to experimentally substantiate our belief that reference mechanism for named entities of different types is basically the same for all entity types, and that references can be computed by the same piece of code.

6. **Implemented numbers and personal names**

   Our effort of making the UNO system process numbers and improving handling personal names was strictly related to MUC-6.

7. **Developed and tested a general approach to handling abbreviations, acronyms and aliases**

   We spent much more effort on abbreviations, acronyms and aliases than originally planned. First, such short forms are very common in written language. Second, handling such short forms resembles handling semantic ambiguity. And third, short forms fit with our ongoing research on context.

8. **Implemented quick-and-dirty pronoun resolution**

   While this code appears to perform well, it breaks occasionally and needs to be further debugged.

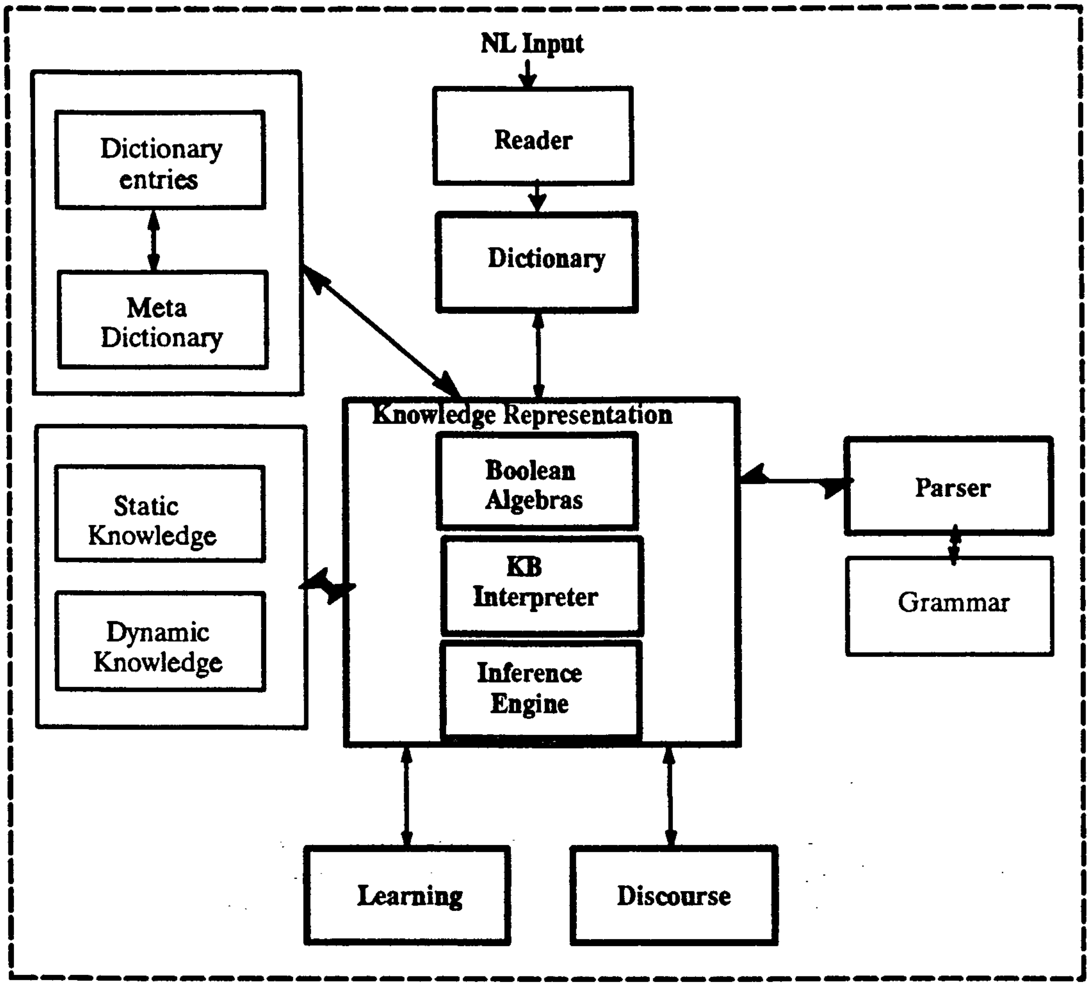

Figure 1: Modules of the UNO natural language processing system

UNO Architecture

The practical significance of the uniformity of the UNO natural-language-based representation and inference, is a simple and flexible architecture of our NLP system:

1. All UNO modules access the knowledge representation module and share its uniform representation.
2. There is no need for external specialists such as knowledge representation systems or temporal reasoners. Our system uniformly represents and reasons with taxonomic, temporal and geographical knowledge.
3. With no external specialists, no interfaces to access them are needed, and therefore there is no need to translate between incompatible representations.

   An NLP system that needs to perform tasks beyond information extraction and to exhibit some in-depth processing such as question answering virtually always calls some external specialists, typically knowledge representation systems. As reported in the literature, the necessity to translate between the representation of the NLP system and such an external specialist is very hard to do and it tremendously complicates control [Palmer et al., 1993] [exp, 1996].

UNO NLP Modules

The UNO NLP system consists of the following modules: **Reader**, **Dictionary**, **Parser**, **Knowledge Representation**, **Discourse**, and **Learning**.

The first three, **Reader**, **Dictionary**, and **Parser**, are modules of the BILING system, a NLP system processing a large number of narratives written by bilingual English/Spanish students [Iwańska, 1989]. The changes to these old modules include augmenting the parser to produce the UNO representation of sentences, enhancing the morphological analyzer to handle prefixes, and supplying the reader with structures for storing the information gained at various stages of processing. The **Knowledge Representation** module implements the theory behind the UNO model of natural language.

The **Reader module** contains functions for breaking input text into documents, paragraphs, sentences, and words. It recognizes abbreviated phrases, contractions, punctuation etc. This module also contains functions for creating various structures from strings and LISP s-expressions, and routines for initializing global variables used by other modules.

The **Dictionary module** contains functions for creating, updating, loading and checking consistency of the UNO dictionary, and functions for performing morphological analysis of the input. Metaknowledge about the dictionary describes its content: it lists known features, specifies feature applicability to different syntactic categories, describes possible and default values of different features (the default values are not shown explicitly in the entries). This metaknowledge facilitates maintaining consistency of the dictionary.

The dictionary is used by the morphological analyzer for supplying each input word with syntactic, semantic, and pragmatic information. The morphological analyzer can recognize and generate various forms of nouns and verbs, for example, *cry*, *cries*, *crying*, *cried*, adjectives, for example, *angrier*, derive adverbs from adjectives, for example, *slowly*, etc. The morphological analyzer handles both prefixes, eg. the prefix *im* in the word *impossible*, and suffixes, eg. *less* in the word *brainless*.

The **Parser module** contains functions implementing a chart parser [Winograd, 1983] [Earley, 1985]. The parser produces both syntactic parse trees and the UNO semantic representation of the natural language input. The grammar allows a limited context-sensitivity via features on lexical categories and non-terminals. Each grammar rule is supplied with the name of a function translating the recognized expressions of natural language into the UNO representation.

The **Knowledge Representation module** consists of the **Boolean algebras module**, **Knowledge Base Interpreter**, and the **Inference Engine module**. The **Boolean algebras module** implements the UNO knowledge representation formalisms and some standard Boolean algebras such as predicate calculus and the powerset of a finite set. The module contains functions for deriving the disjunctive normal form of a complex Boolean expression independently of its algebra, as well as functions for creating the representation of the element that this complex Boolean expression stands for.

The **Knowledge Base Interpreter** implements the interpreter of the sets of type equations encoding taxonomic, temporal and geographical knowledge.

The **Inference Engine module** implements the UNO algorithm for representing and utilizing knowledge derived from natural language sentences. This algorithm updates the dynamic knowledge bases of the UNO system.

The **Discourse** module implements anaphora resolution, functions for identifying referring expressions and computing referents if needed, and discourse processing, functions for computing certain discourse structures that facilitate maintaining dynamic knowledge bases.

The **Learning module** consists of functions mixing statistics and inductive learning techniques and is used for corpus analysis and definite-anaphora-based knowledge acquisition.

Control

Flexible, non-sequential control with all modules accessing the **Knowledge Representation** module.

Speed

Our system is reasonably fast. For MUC-6, it took slightly under half a minute to process a typical WSJ article in the development set.

WALKTHROUGH ARTICLE AND MORE

Somewhat Expected and Unexpected Problems

Given our extremely ambitious goals for such a short period of time, and particularly knowledge engineering large knowledge bases prevented us from doing coreference. We fully expected that we will not be able to successfully complete improving our existing anaphora resolution code and make it work reliably with the newly created large knowledge bases.

However, we did not expect a failure of the function printing out the markings. This buggy piece of code is about the easiest to fix, but at the same time, it is the most damaging in terms of the score. We had to not print the recognized organization names and turn off processing of the "HL" and "DATELINE" parts of the articles.

Our Official, Unofficial, and Very Unofficial Score

Our scores are not very meaningful because despite the fact that the problematic markings constituted only a small percentage of all the markings, 8% for the first six test articles, the scorer failed to score anything, but four shortest articles; this can be seen in the official results table below- the "DD" slot containing a date in a canned "Month/Day/Year" format shows only four matches.

In order to obtain an unofficial score, we had to edit our results, which we did strictly according to the trace. Even this unofficial score does not reflect well our real performance, only shows that we did something. For those things that we were able to mark, our own **very unofficial** estimate is that we performed in the high nineties in both recall and precision (as can be seen in the enclosed sample article). For example, we correctly marked both expressions problematic for most other systems: "*the 21st century*" and "*Hollywood*".

OFFICIAL

* * * TOTAL SLOT SCORES * * *

SLOT	POS	ACT—	COR	PAR	INC—	SPU	MIS	NON—	REC	PRE	UND	OVG	ERR	SUB
<enamex>	909	9—	7	0	0—	2	902	0—	1	78	99	22	99	0
type	909	9—	5	0	2—	2	902	0—	0	56	99	22	99	28
text	909	9—	5	0	2—	2	902	0—	0	56	99	22	99	28
subtotals	1818	18—	10	0	4—	4	1804	0—	0	56	99	22	99	28
<timex>	112	6—	5	0	0—	1	107	0—	4	83	96	17	96	0
type	112	6—	5	0	0—	1	107	0—	4	83	96	17	96	0
text	112	6—	4	0	1—	1	107	0—	4	67	96	17	96	20
subtotals	224	12—	9	0	1—	2	214	0—	4	75	96	17	96	10
<numex>	93	5—	5	0	0—	0	88	0—	5	100	95	0	95	0
type	93	5—	5	0	0—	0	88	0—	5	100	95	0	95	0
text	93	5—	3	0	2—	0	88	0—	3	60	95	0	97	40
subtotals	186	10—	8	0	2—	0	176	0—	4	80	95	0	96	20
ALL OBJECTS	2228	40—	27	0	7—	6	2194	0—	1	68	98	15	99	20
MATCHED ONLY	34	34—	27	0	7—	0	0	0—	79	79	0	0	20	20

	P&R	2P&R	P&2R
F-MEASURES	2.38	5.65	1.51

* * * DOCUMENT SECTION SCORES * * *

SLOT	POS	ACT—	COR	PAR	INC—	SPU	MIS	NON—	REC	PRE	UND	OVG	ERR	SUB
HL	128	0—	0	0	0—	0	128	0—	0	*	100	*	100	*
DD	60	4—	4	0	0—	0	56	0—	7	100	93	0	93	0
DATELINE	52	0—	0	0	0—	0	52	0—	0	*	100	*	100	*
TXT	1988	36—	23	0	7—	6	1958	0—	1	64	98	17	99	23

UNOPPICIAL SCORE OF EDITED RESULTS

* * * TOTAL SLOT SCORES * * *

SLOT	POS	ACT—	COR	PAR	INC—	SPU	MIS	NON—	REC	PRE	UND	OVG	ERR	SUB
<enamex>	912	422—	254	0	0—	168	658	0—	28	60	72	40	76	0
type	912	422—	230	0	24—	168	658	0—	25	54	72	40	79	9
text	912	422—	204	0	50—	168	658	0—	22	48	72	40	81	20
subtotals	1824	844—	434	0	74—	336	1316	0—	24	51	72	40	80	14
<timex>	112	178—	87	0	0—	91	25	0—	78	49	22	51	57	0
type	112	178—	86	0	1—	91	25	0—	77	48	22	51	58	1
text	112	178—	79	0	8—	91	25	0—	70	44	22	51	61	9
subtotals	224	356—	165	0	9—	182	50	0—	74	46	22	51	59	5
<numex>	93	105—	70	0	0—	35	23	0—	75	67	25	33	45	0
type	93	105—	70	0	0—	35	23	0—	75	67	25	33	45	0
text	93	105—	59	0	11—	35	23	0—	63	56	25	33	54	16
subtotals	186	210—	129	0	11—	70	46	0—	69	61	25	33	50	8
ALL OBJECTS	2234	1410—	728	0	94—	588	1412	0—	32	52	63	42	74	11
MATCHED ONLY	822	822—	728	0	94—	0	0	0—	88	88	0	0	11	11

```
-----------+-------+-----+---------
P&R     2P&R     P&2R
F-MEASURES       39.96    46.23    35.18

* * * DOCUMENT SECTION SCORES * * *
-----------+-------+-----+---------
SLOT          POS  ACT-  COR  PAR  INC- SPU MIS NON- REC PRE UND OVG ERR SUB
-----------+-------+-----+---------
HL        128    0-    0    0    0-    0 128   0-    0  * 100  * 100  *
DD         60   60-   60    0    0-    0   0   0- 100 100   0   0   0   0
DATELINE        52    0-    0    0    0-    0  52   0-    0  * 100  * 100  *
TXT       1994 1350-  668    0   94-  588 1232   0-  34  49  62  44  74  12
-----------+-------+-----+---------
```

Edits

Below, we explain our edits and enclose the answer key for the walkthrough article with the hand-written markings showing the expressions we recognized as well as our edited markings.

In the first 6 test articles, there were 16, or 8% problematic markings, and 186, or 92% of unproblematic markings, with bits of text that was not there appearing and portions of teh existing text disappearing:

1. Double-closed marks

```
Example:     Ms. <ENAMEX TYPE="PERSON">Washington</ENAMEX>,</ENAMEX>,
Edit:        Ms. <ENAMEX TYPE="PERSON">Washington</ENAMEX>,
Nr of edits: 4
```

2. Simple unclosed/unopen marks

```
Example:     of News Corp.'s News <ENAMEX TYPE="LOCATION">America Publishing unit
Edit:        of News Corp.'s News <ENAMEX TYPE="LOCATION">America </ENAMEX>
Nr of edits: 10
```

3. Really scrambled:

```
Example:     <TIMEX TYPE="TIME">$<NUMEX TYPE="MONEY">$725,000 last year</TIMEX>
Edit:        <NUMEX TYPE="MONEY">$725,000</NUMEX> <TIMEX TYPE="TIME"> last year</TIMEX>
Nr of edits: 2
```

4. Nothing was done about the introduced text and the markings that were not printed at all, but expressions clearly recognized, as evidenced by the trace.

Qualitative Indication of Our Performance for the Walkthrough Article

```
LEGEND:
 + correctly identified expression
 +- correctly identified expression, but too short or too long
 - not done

<DOCID> wsj94'026.0231 </DOCID>
<DOCNO> 940224-0133. </DOCNO>
<HL>    Marketing & Media - Advertising:
@ <ENAMEX TYPE="PERSON">John Dooner</ENAMEX> Will Succeed <ENAMEX TYPE="PERSON">James</ENAMEX>
@ At Helm of <ENAMEX TYPE="ORGANIZATION">McCann-Erickson</ENAMEX>
@ ---
@ By <ENAMEX TYPE="PERSON">Kevin Goldman</ENAMEX> </HL>
<DD> <TIMEX TYPE="DATE">02/24/94</TIMEX> </DD>
<SO> WALL STREET JOURNAL (J), PAGE B8 </SO>
<CO>    IPG K </CO>
<IN> ADVERTISING (ADV), ALL ENTERTAINMENT & LEISURE (ENT),
  FOOD PRODUCTS (FOD), FOOD PRODUCERS, EXCLUDING FISHING (OFP),
  RECREATIONAL PRODUCTS & SERVICES (REC), TOYS (TMF) </IN>
<TXT>
<p>
  One of the many differences between <ENAMEX TYPE="PERSON">Robert L. James</ENAMEX>, chairman and
chief executive officer of <ENAMEX TYPE="ORGANIZATION">McCann-Erickson</ENAMEX>, and
<ENAMEX TYPE="PERSON">John J. Dooner Jr.</ENAMEX>,
the agency's president and chief operating officer, is quite
telling: Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> enjoys sailboating, while Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>
owns a powerboat.
</p>
<p>
  Now, Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> is preparing to sail into the sunset, and Mr.
<ENAMEX TYPE="PERSON">Dooner</ENAMEX> is poised to rev up the engines to guide
<ENAMEX TYPE="ORGANIZATION">Interpublic Group</ENAMEX>'s
<ENAMEX TYPE="ORGANIZATION">McCann-Erickson</ENAMEX> into <TIMEX TYPE="DATE">the 21st century</TIMEX>. Yesterday,
<ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX> made
official what had been widely anticipated: Mr. <ENAMEX TYPE="PERSON">James</ENAMEX>, 57 years old,
is stepping down as chief executive officer on <TIMEX TYPE="DATE">July 1</TIMEX> and will
```

retiring as chairman at the end of the year. He will be succeeded by
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, 45.
</p>
<p>
It promises to be a smooth process, which is unusual given the
volatile atmosphere of the advertising business. But Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> has
a big challenge that will be his top priority. "I'm going to focus
on strengthening the creative work," he says. "There is room to
grow. We can make further improvements in terms of the perception of our creative work."
</p>
<p>
Even <ENAMEX TYPE="PERSON">Alan Gottesman</ENAMEX>, an analyst with <ENAMEX TYPE="ORGANIZATION">PaineWebber</ENAMEX>,
who believes <ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX> is filled with "vitality" and is in "great shape," says that
from a creative standpoint, "You wouldn't pay to see their reel" of commercials.
</p>
<p>
While <ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX>'s world-wide billings rose <NUMEX TYPE="PERCENT">12%</NUMEX> to
<NUMEX TYPE="MONEY">$6.4 billion</NUMEX> last
year from <NUMEX TYPE="MONEY">$5.7 billion</NUMEX> in <TIMEX TYPE="DATE">1992</TIMEX>, the agency still is dogged by the
loss of the key creative assignment for the prestigious Coca-Cola
Classic account. "I would be less than honest to say I'm not
disappointed not to be able to claim creative leadership for <ENAMEX TYPE="ORGANIZATION" STATUS="OPT">Coke</ENAMEX>,"
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> says.
</p>
<p>
<ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX> still handles promotions and media buying for
<ENAMEX TYPE="ORGANIZATION">Coke</ENAMEX>. But
the bragging rights to <ENAMEX TYPE="ORGANIZATION" STATUS="OPT">Coke</ENAMEX>'s ubiquitous advertising belongs to
<ENAMEX TYPE="ORGANIZATION">Creative Artists Agency</ENAMEX>, the big <ENAMEX TYPE="LOCATION">Hollywood</ENAMEX>
talent agency. "We are
striving to have a strong renewed creative partnership with
<ENAMEX TYPE="ORGANIZATION">Coca-Cola</ENAMEX>," Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> says. However, odds of that
happening are
slim since word from <ENAMEX TYPE="ORGANIZATION">Coke</ENAMEX> headquarters in <ENAMEX TYPE="LOCATION">Atlanta</ENAMEX>
is that <ENAMEX TYPE="ORGANIZATION">CAA</ENAMEX> and
other ad agencies, such as <ENAMEX TYPE="ORGANIZATION">Fallon McElligott</ENAMEX>, will continue to
handle <ENAMEX TYPE="ORGANIZATION">Coke</ENAMEX> advertising.
</p>
<p>
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, who recently lost 60 pounds over three-and-a-half
months, says now that he has "reinvented" himself, he wants to do
the same for the agency. For Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, it means maintaining his
running and exercise schedule, and for the agency, it means
developing more global campaigns that nonetheless reflect local
cultures. One <ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX> account, "I Can't Believe It's Not Butter," a
butter substitute, is in 11 countries, for example.
</p>
<p>
<ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX> has initiated a new so-called global collaborative system,
composed of world-wide account directors paired with creative
partners. In addition, <ENAMEX TYPE="PERSON">Peter Kim</ENAMEX> was hired from
<ENAMEX TYPE="ORGANIZATION">WPP Group</ENAMEX>'s <ENAMEX TYPE="ORGANIZATION">J.
Walter Thompson</ENAMEX> last <TIMEX TYPE="DATE">September</TIMEX> as vice chairman, chief strategy
officer, world-wide.
</p>
<p>
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> doesn't see a creative malaise permeating the agency.
He points to several campaigns with pride, including the Taster's
Choice commercials that are like a running soap opera. "It's a <NUMEX TYPE="MONEY">$19
million</NUMEX> campaign with the recognition of a <NUMEX TYPE="MONEY">$200 million</NUMEX> campaign,"
he says of the commercials that feature a couple that must hold a record for the length of time dating before kissing.
</p>
<p>
Even so, Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> is on the prowl for more creative talent and
is interested in acquiring a hot agency. He says he would like to
finalize an acquisition "yesterday. I'm not known for patience."
</p>
<p>
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> met with <ENAMEX TYPE="PERSON">Martin Puris</ENAMEX>,
president and chief executive
officer of <ENAMEX TYPE="ORGANIZATION">Ammirati & Puris</ENAMEX>, about <ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX>'s
acquiring the agency
with billings of <NUMEX TYPE="MONEY">$400 million</NUMEX>, but nothing has materialized. "There
is no question," says Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, "that we are looking for quality
acquisitions and <ENAMEX TYPE="ORGANIZATION">Ammirati & Puris</ENAME (> is a quality operation. There are
some people and entire agencies that I would love to see be part of
the <ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX> family." Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> declines to
identify possible acquisitions.
</p>
<p>
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> is just gearing up for the headaches of running one of
the largest world-wide agencies. (There are no immediate plans to
replace Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> as president; Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> operated
as chairman,
chief executive officer and president for a period of time.) Mr.
<ENAMEX TYPE="PERSON">James</ENAMEX> is filled with thoughts of enjoying his three hobbies: sailing, skiing and hunting.
</p>
<p>
Asked why he would choose to voluntarily exit while he still is
so young, Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> says it is time to be a tad selfish about how he
spends his days. Mr. <ENAMEX TYPE="PERSON">James</ENAMEX>, who has a reputation as an
extraordinarily tough taskmaster, says that because he "had a great
time" in advertising," he doesn't want to "talk about the
disappointments." In fact, when he is asked his opinion of the new
batch of <ENAMEX TYPE="ORGANIZATION" STATUS="OPT">Coke</ENAMEX> ads from <ENAMEX TYPE="ORGANIZATION">CAA</ENAMEX>,
Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> places his hands over his
mouth. He shrugs. He doesn't utter a word. He has, he says, fond
memories of working with <ENAMEX TYPE="ORGANIZATION">Coke</ENAMEX> executives. "<ENAMEX TYPE="ORGANIZATION">Coke</ENAMEX>
has given us great
highs," says Mr. <ENAMEX TYPE="PERSON">James</ENAMEX>, sitting in his plush office, filled with
photographs of sailing as well as huge models of, among other things, a Dutch tugboat.
</p>
<p>
He says he feels a "great sense of accomplishment." In 36
countries, <ENAMEX TYPE="ORGANIZATION">McCann</ENAMEX> is ranked in the top three; in 75 countries, it is in the top 10.
</p>
<p>

Soon, Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> will be able to compete in as many sailing races
as he chooses and concentrate on his duties as rear commodore at
the <ENAMEX TYPE="ORGANIZATION">New York Yacht Club</ENAMEX>.
</p>
<p>
Maybe he'll even leave something from his office for Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>.
Perhaps a framed page from the New York Times, dated <TIMEX TYPE="DATE">Dec. 8, 1987</TIMEX>,
showing a year-end chart of the stock market crash earlier that
year. Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> says he framed it and kept it by his desk as a
"personal reminder. It can all be gone like that."
</p>
</TXT>
</DOC>

Edited Markings for the Walkthrough Article

<DOC>
<DOCID> wsj94'026.0231 </DOCID>
<DOCNO> 940224-0133. </DOCNO>
<HL> Marketing & Media – Advertising:
@ John Dooner Will Succeed James
@ At Helm of McCann-Erickson
@ —-
@ By Kevin Goldman </HL>
<DD> <TIMEX TYPE="DATE">02/24/94</TIMEX> </DD>
<SO> WALL STREET JOURNAL (J), PAGE B8 </SO>
<CO> IPG K </CO>
<IN> ADVERTISING (ADV), ALL ENTERTAINMENT & LEISURE (ENT),
FOOD PRODUCTS (FOD), FOOD PRODUCERS, EXCLUDING FISHING (OFP),
RECREATIONAL PRODUCTS & SERVICES (REC), TOYS (TMF) </IN>
<TXT>
<p>
One of the many differences between <ENAMEX TYPE="PERSON">Robert L. James</ENAMEX>, chairman and
chief executive officer of McCann-Erickson, and <ENAMEX TYPE="PERSON">John J. Dooner Jr</ENAMEX>.,
the agency's president and chief operating officer, is quite
telling: Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> enjoys sailboating, while Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>
owns a
powerboat.
</p>
<p>
Now, Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> is preparing to sail into the sunset, and Mr.
<ENAMEX TYPE="PERSON">Dooner</ENAMEX> is poised to rev up the engines to guide Interpublic Group's
McCann-Erickson into the <TIMEX TYPE="TIME">21st century</TIMEX>. Yesterday, McCann made
official what had been widely anticipated: Mr. <ENAMEX TYPE="PERSON">James</ENAMEX>, <TIMEX TYPE="TIME">57 years</TIMEX>
old,
is stepping down as chief executive officer on <TIMEX TYPE="DATE">July 1</TIMEX> and will
retire as chairman at the end of the year. He will be succeeded by
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, 45.
</p>
<p>
It promises to be a smooth process, which is unusual given the
volatile atmosphere of the advertising business. But Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> has
a big challenge that will be his top priority. "I'm going to focus
on strengthening the creative work," he says. "There is room to
grow. We can make further improvements in terms of the perception of
our creative work."
</p>
<p>
<ENAMEX TYPE="PERSON">Even Alan Gottesman</ENAMEX>, an analyst with PaineWebber, who believes
McCann is filled with "vitality" and is in "great shape," says that
from a creative standpoint, "You wouldn't pay to see their reel" of
commercials.
</p>
<p>
While McCann's world-wide billings rose <NUMEX TYPE="PERCENT">12%</NUMEX> to
<TIMEX TYPE="TIME"><NUMEX TYPE="MONEY">$6.4 billion</NUMEX> last
year</TIMEX> from <NUMEX TYPE="MONEY">$5.7 billion</NUMEX> in <TIMEX TYPE="DATE">1992</TIMEX>,
loss of the key creative assignment for the prestigious Coca-Cola
Classic account. "I would be less than honest to say I'm not
disappointed not to be able to claim creative leadership for Coke,"
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> says.
</p>
<p>
McCann still handles promotions and media buying for Coke. But
the bragging rights to Coke's ubiquitous advertising belongs to
Creative Artists Agency, the big <ENAMEX TYPE="LOCATION">Hollywood</ENAMEX> talent agency. "We are
striving to have a strong renewed creative partnership with
Coca-Cola," Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> says. However, odds of that happening are
slim since word from Coke headquarters in <ENAMEX TYPE="LOCATION">Atlanta</ENAMEX> is that CAA and
other ad agencies, such as Fallon McElligott, will continue to
handle Coke advertising.
</p>
<p>
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, who recently lost <NUMEX TYPE="MONEY">60 pounds</NUMEX> over
three-and-a-<TIMEX TYPE="TIME">half
months</TIMEX>, says now that he has "reinvented" himself, he wants to do
the same for the agency. For Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, it means maintaining his
running and exercise schedule, and for the agency, it means
developing more global campaigns that nonetheless reflect local
cultures. One McCann account, "I Can't Believe It's Not <ENAMEX TYPE="PERSON">Butter</ENAMEX>," a
butter substitute, is in 11 countries, for example.
</p>
<p>
McCann has initiated a new so-called global collaborative system,
composed of world-wide account directors paired with creative
partners. In addition, <ENAMEX TYPE="PERSON">Peter Kim</ENAMEX> was hired from WPP Group's J.
<ENAMEX TYPE="PERSON">Walter Thompson</ENAMEX> last <TIMEX TYPE="DATE">September</TIMEX> as vice chairman, chief strategy
officer, world-wide.
</p>
<p>
Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> doesn't see a creative malaise permeating the agency.
He points to several campaigns with pride, including the Taster's
Choice commercials that are like a running soap opera. "It's a <NUMEX TYPE="MONEY">$19</NUMEX>
million campaign with the recognition of a <NUMEX TYPE="MONEY">$200 million</NUMEX> campaign,"
he says of the commercials that feature a couple that must hold a

```
record for the length of time dating before kissing.
</p>
<p>
  Even so, Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> is on the prowl for more creative talent and
is interested in acquiring a hot agency. He says he would like to finalise an acquisition "yesterday. I'm not known for patience."
</p>
<p>
  Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> met with <ENAMEX TYPE="PERSON">Martin Puris</ENAMEX>, president and
chief executive
officer of Ammirati & Puris, about McCann's acquiring the agency
with billings of <NUMEX TYPE="MONEY">$400 million</NUMEX>, but nothing has materialised. "There
is no question," says Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>, "that we are looking for quality
acquisitions and Ammirati & Puris is a quality operation. There are
some people and entire agencies that I would love to see be part of
the McCann family." Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> declines to identify possible
acquisitions.
</p>
<p>
  Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> is just gearing up for the headaches of running one of
the largest world-wide agencies. (There are no immediate plans to
replace Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX> as president; Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> operated
as chairman, chief executive officer and president for a period of time.) Mr.
<ENAMEX TYPE="PERSON">James</ENAMEX> is filled with thoughts of enjoying his three hobbies: sailing, skiing and hunting.
</p>
<p>
  Asked why he would choose to voluntarily exit while he still is
so young, Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> says it is time to be a tad selfish about how he
spends his days. Mr. <ENAMEX TYPE="PERSON">James</ENAMEX>, who has a reputation as an
extraordinarily tough taskmaster, says that because he "had a great
time" in advertising," he doesn't want to "talk about the
disappointments." In fact, when he is asked his opinion of the new
batch of Coke ads from CAA, Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> places his hands over his
mouth. He shrugs. He doesn't utter a word. He has, he says, fond
memories of working with Coke executives. "Coke has given us great
highs," says Mr. <ENAMEX TYPE="PERSON">James</ENAMEX>, sitting in his plush office, filled with
photographs of sailing as well as huge models of, among other things, a Dutch tugboat.
</p>
<p>
  He says he feels a "great sense of accomplishment." In 36
countries, McCann is ranked in the top three; in 75 countries, it is
in the top 10.
</p>
<p>
  Soon, Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> will be able to compete in as many sailing races
as he chooses. And concentrate on his duties as rear commodore at
the <ENAMEX TYPE="LOCATION">New York</ENAMEX> Yacht Club.
</p>
<p>
  Maybe he'll even leave something from his office for Mr. <ENAMEX TYPE="PERSON">Dooner</ENAMEX>.
Perhaps a framed page from the <ENAMEX TYPE="LOCATION">New York</ENAMEX> Times, dated <TIMEX TYPE="DATE">Dec. 8</TIMEX>,
<TIMEX TYPE="DATE">1987</TIMEX>,
showing a year-end chart of the stock market crash earlier that
year. Mr. <ENAMEX TYPE="PERSON">James</ENAMEX> says he framed it and kept it by his desk as a
"personal reminder. It can all be gone like that."
</p>
</TXT>
</DOC>
```

Discussion of Results

For each subtask, we show the D/T (developed/tested) ratio, the ratio of the number of articles used for development to the number of articles used for testing. For errors, we show a brief explanation of the cause.

NUMEXES

D/T = 100/600 + many convoluted made up examples

Our performance on numexes is very good– it is not an extremely difficult task and we strove and achieved a near-completeness. We correctly identified interesting cases:

"*C$7.4 million*"

"*135 million Canadian dollars*"

"*$414,167*"

"*87.5 cents*"

"*18 Canadian cents*"

"*three cents*"

"*five to 10 cents*"

Occasionally the system makes an error. For example, "*though it is assumed that dozens of them won't be*", "*Co., to succeed Mark*", "*60 pounds*" are identified as money (the system knows pound, mark and won to be British, German and Korean monetary units, respectively; the culprit is too local interpretation.)

TIMEXES

D/T = 100/600

Our very good performance on the task of identifying temporal expressions was sligthly improved with handling numbers. Previously we were missing "bare" numeric years such as "*the 1980 election*", "*1980s*", "*pre-1970*", and dates such as "*4.12*" (as referring to "*April 12-th*").

Our errors included: "*57 years old*" (probably misinterpreted task definition), "*last year*" (deliberately kept) and "*three-and-a-half months*" (only some dashes handled).

LOCATIONS

D/T = 100/600 (100/3000)

Good performance on identifying locations stems from the combination of our rather complete critical knowledge bases with the major named types and geographical information for all countries and automatic interpretation of locative expressions with a known geographical type such as "*the city of Farmington Hills*" and "*The Isle of Man*", and expressions of the form: "Smaller Region, Larger Region" such as "*Poland, New York*".

PERSONAL NAMES

D/T = 100/600 (100/3000)

Unsurprisingly good performance of our system in processing personal names stems from our fairly complete knowledge bases with reasonable and not very strange first names along with the gender and nicknames as aliases information. Processing personal names is also supported by an extensive, 300-entry dictionary of titles and professional degrees. Some of our more interesting titles include:

"*Lieutenant Junior Grade*" abbreviated as "*Lt. jg*", "*Knight Commander of the Order of the British Empire*" and "*Your Serene Highness*".

We correctly identified the following interesting cases:

"*Ms. Washington*", "*Mr. York*" and "*Ms. Lansing*" were not confused with locations

"*John J. Dooner Jr*"

"*Steve*" as first name reference

"*Kenneth H. Thorn*" no title

"*Peter A. Leff*" was not confused with the direction as in "*go to the left*"

"*Howard Dean*" was not confused with the title

Our errors included:

"*Robert S. "Steve" Miller*" as two persons "*Robert S.*" and "*Steve Miller*" (doublequotes basically ignored)

"*CFO Paul Rizzo*" as a personal name (no real correction of misspels)

"*Thomas H. O'Brien Jr.*" (forgot about the singlequote in names, single letter cannot be a last name)

"*Dawn Capital*" as a personal name (organization names suppressed)

"*Sun Chief Executive Scott McNealy*" with "*Sun*" as TIMEX because abbreviates "*Sunday*", "*Chief*" as title because of a short cut, "*Executive Scott McNealy*" as a first-middle-last name triple

(organization names not marked and interpretation too local)

"*does not consider himself a Butt-Head Astronomer*" as a person with the "*Head*" title and the last name "*Astronomer*" (interpretation too local)

Surprises

We were surprised by two things. First, that there is so much (depth to) abbreviations. And second, that personal pronouns do not obey expected distance constraints, contrary to the claims in the existing literature that 96% of pronominal referents are no further than three sentences away.

Greatest Limiting Factors

For us, the greatest limiting factors were time, time, and time again. We also found it very hard to create and debug large, circa 20,000 entry, typed knowledge bases. Our knowledge acquisition and debugging techniques involved a combination of hand-crafting (geographical knowledge base), scanning (abbreviations), material that others shared with us (for example, our department gave us an on-line, 800-or-so list of universities), and semi-automatic acquisition (our mentioned earlier newly developed technique based on definite anaphora).

References

[exp, 1996] (1996). *Special Issue of the International Journal of Expert Systems on Knowledge Representation and Inference for Natural Language Processing.* JAI Press, Inc.

[Aït-Kaci, 1986] Aït-Kaci, H. (1986). An algebraic semantics approach to the effective resolution of type equations. *Journal of Theoretical Computer Science*, 45:293–251.

[Aït-Kaci and Richard Meyer and Peter Van Roy, 1993] Aït-Kaci and Richard Meyer and Peter Van Roy, H. (1993). Wild LIFE: A User Manual (Preliminary version). Prl research report, Digital Equipment Corporation, Paris Research Laboratory.

[Barwise and Cooper, 1981] Barwise, J. and Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159-219.

[Dowty et al., 1981] Dowty, D. R., Wall, R. E., and Peters, S. (1981). *Introduction to Montague semantics.* D. Reidel Publ. Co.

[Earley, 1985] Earley, J. (1985). An efficient context-free parsing algorithm. In *Readings in Natural Language Processing*, pages 25–33. Morgan Kaufmann Publishers, Inc.

[Hamm, 1989] Hamm, F. (1989). *Naturlich-sprachliche quantoren.* Max Niemeyer Verlag.

[Horn, 1972] Horn, L. R. (1972). *On the Semantic Properties of Logical Operators in English.* IULC.

[Horn, 1989] Horn, L. R. (1989). *A Natural History of Negation.* The University of Chicago Press.

[Iwańska, 1989] Iwańska, L. (1989). Automated Processing of Narratives Written by 6-12 Grade Students: The BILING Program. Technical Report UIUCDCS-R-89-1508, Dept. of Computer Science, University of Illinois at Urbana-Champaign.

[Iwańska, 1992] Iwańska, L. (1992). A General Semantic Model of Negation in Natural Language: Representation and Inference. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR92)*, pages 357–368.

[Iwańska, 1993] Iwańska, L. (1993). Logical Reasoning in Natural Language: It Is All About Knowledge. *International Journal of Minds and Machines, Special Issue on Knowledge Representation for Natural Language*, 3(4):475–510.

[Iwańska, 1994] Iwańska, L. (1994). Talking about Time: Temporal Reasoning as A Problem of Natural Language. In *Working Notes of the AAAI Fall Symposium on* Knowledge Representation for Natural Language Processing in Implemented Systems, pages 70–81. Also appeared in Proceedings of the Third International Workshop on Intelligent Information Systems, Springer Verlag.

[Iwańska, 1996a] Iwańska, L. (1996a). Definite Anaphora for Knowledge Acquisition. In preparation.

[Iwańska, 1996b] Iwańska, L. (1996b). Natural (Language) Temporal Logic: Reasoning about Absolute and Relative Time. *International Journal of Expert Systems (to appear).*

[Keenan and Faltz, 1985] Keenan, E. L. and Faltz, L. M. (1985). *Boolean Algebra Semantics Of Natural Language.* D. Reidel Publ. Comp.

[Montague, 1973] Montague, R. (1973). The Proper Treatment of Quantification in Ordinary English. In Hintikka, J., Moravcik, J., and Suppes, P., editors, *Approaches to Natural Language.* D. Reidel Publ. Co.

[Palmer et al., 1993] Palmer, M., Passonneau, R., Weir, C., and Finin, T. (1993). The KERNEL Text Understanding System. *Artificial Intelligence*, 63:17–69.

[Pereira, 1987] Pereira, F. (1987). Grammars and logics of partial information. In *Proceedings of of the International Conference on Logic Programming*, volume 2, pages 989–1013.

[Winograd, 1983] Winograd, T. (1983). *Language as a Cognitive Process.* Addison-Wesley Publ. Comp.