University of Massachusetts: Description of the CIRCUS System as Used for MUC-4

W. Lehnert, C. Cardie, D. Fisher, J. McCarthy, E. Riloff, & S. Soderland

University of Massachusetts
Department of Computer Science
Amherst, MA 01003
lehnert@cs.umass.edu

THE CIRCUS SENTENCE ANALYZER

CIRCUS is a conceptual analyzer that produces semantic case frame representations for input sentences. Although space does not permit us to give a full technical description of CIRCUS, we will attempt to convey some sense of sentence analysis via CIRCUS. For more details, please consult [2] and [1].

CIRCUS uses no syntactic grammar and produces no parse tree as it analyzes a sentence. Rather, it uses lexically-indexed syntactic knowledge to segment incoming text into noun phrases, prepositional phrases, and verb phrases. These constituents are stored in global buffers that track the subjects, verbs, direct objects, and prepositional phrases of a sentence. Because we restrict the buffer contents to simple constituents with a highly local sense of the sentence, larger constituents like clauses are not explicitly stored by the syntactic component of CIRCUS.

While syntactic buffers are being bound to sentence fragments, a mechanism for handling predictive semantics is responsible for establishing case role assignments. Semantic case frames are activated by concept node (CN) definitions, and each CN definition can be triggered by one or more lexical items. Associated with each slot in a CN are both hard and soft constraints. A hard constraint is a predicate that *must* be satisfied, while a soft constraint defines a *preference* rather than an absolute requirement. When a CN instantiation meets certain criteria established by the CN definition, CIRCUS freezes that case frame and passes it along as output from the sentence analyzer. A single sentence can generate an arbitrary number of case frame instantiations depending on the conceptual complexity of the sentence and the availability of relevant CN definitions in the dictionary.

Because CIRCUS is designed to generate case frame representations in response to sentence fragments, ungrammatical sentences or sentences with highly complicated syntactic structures are often navigated without difficulty. CIRCUS was designed to maximize robust processing in the face of incomplete knowledge. It does not require complete dictionary coverage with respect to CN definitions or even part-of-speech recognition, so CIRCUS is especially well-suited for text extraction applications from unconstrained text. The first serious evaluation of CIRCUS took place with MUC-3, where CIRCUS posted the highest combined scores for recall and precision of all the participating sites [3, 4, 5].

MEMORY-BASED CONSOLIDATION

Consolidation refers to the problem of mapping CN instantiations produced by CIRCUS into event descriptions appropriate for target template instantiations. Since information pertaining to a single event can be distributed across multiple sentences, problems associated with consolidation are challenging, especially when a text describes multiple events. It is necessary to know when different noun phrases point to the same referent and when the topic shifts from one event to another.

The UMass/MUC-3 system used a rule based consolidation module which was largely dominated by rules designed to merge appropriate structures. Because the rule base was large (168 rules), it was difficult to pinpoint weak spots in the rule base and it became increasingly difficult to make reliable adjustments as needed. Because of our dissatisfaction with last year's approach, we decided to design a new consolidation module for MUC-4.

Our new consolidation module is "memory-based" in the sense that it assumes a specific memory organization strategy, and all processing is motivated by a small number of memory manipulations. The basic structure of memory-based consolidation (MBC) is a simple stack of incident structures, along with two associated

stacks that track human targets and physical targets. At the end of each consolidation run, the number of incident structures on the incident stack usually corresponds to the number of templates we will instantiate, with each incident structure containing all the information needed to fill at least one template.

The incident-structure serves as the basic data type inside MBC as well as the data type that is output from MBC. An incident structure is a frame consisting of slots for a date, location, perpetrators, and subevents. Each subevent consists of a specific incident type (murder, bombing, robbery, etc.) along with victims, physical targets, instruments, and effects. Although multiple subevents are permitted in an incident-structure to handle combined events like a arson/robbery combination, most incident structures contain only one subevent. When a new incident-structure is input to MBC, it will either merge with an existing incident structure already on the incident stack, or it will be added to the incident stack as a separate incident. When target templates are eventually generated from incident structures on the incident stack, each subevent within an incident structure will spawn its own template instantiation.

In comparing MBC with the rule-based consolidation module in UMass/MUC-3, we find that MBC tends to generate fewer spurious templates without sacrificing significant recall. However, we have seen test sets where MBC does lag behind in recall. In general, the two modules seem quite comparable in terms of overall performance, although MBC is easier to understand, maintain, and scale-up. Most of the merging rules used by rule-based consolidation were incorporated into MBC, so it makes sense that the two modules exhibit similar behavior. Our decision to run MBC for MUC-4 was largely motivated by use of the All Templates metric as the official scoring metric for MUC-4. Because All Templates is maximally sensitive to all types of precision loss, it is generally advantageous to minimize spurious templates for this metric. MBC seemed consistently better at eliminating spurious templates, so we decided to risk a possible loss of some recall for the sake of maximizing our precision.

A SHORT WALK THROUGH TST2-MUC4-0048

In order to illustrate the behavior of UMass/MUC-4 in operation, we will trace the processing of a sample text that contains two separate bombing incidents. In general, CIRCUS generates multiple CN instantiations in response to each sentence, while Memory-Based Consolidation (MBC) extracts information from the CNs and organizes it within incident structures. CIRCUS and MBC work in a serial fashion: CIRCUS analyzes the entire text first, and then MBC works on the resulting concept node instantiations. But for the sake of this presentation, we will examine the effects of CIRCUS and MBC working together on a sentence-by-sentence basis.

Because our CN definitions extract information on the basis of phrase fragments, we will underline those portions of the input sentences that are important to relevant CN's. Any remaining segments of the input sentences that are not underlined are effectively ignored during semantic processing by CIRCUS. We will also show the preprocessed version of each input sentence, to indicate which items have been recognized by the phrasal lexicon (these will be catenated), and other minor transformations to the original source text. Abbreviations preceded by ">" represent punctuation marks. For example, >CO is a comma.

The first job of MBC is to partition multiple CNs into event structures which are then restructured into incident structures. As a rule, all CNs generated from a single sentence tend to fall into the same partition, so we will omit any detailed discussion of this preliminary conversion. But it is important to note that essential merging operations can take place during the creation of initial incident structures. For example, S1 illustrates how an accused perpetrator is linked to a murder because their associated CNs fall into a single partition:

S1: (SALVADORAN PRESIDENT-ELECT ALFREDO CRISTIANI CONDEMNED THE TERRORIST KILLING OF ATTORNEY GENERAL ROBERTO GARCIA ALVARADO AND ACCUSED THE FARABUNDO_MARTI_NATIONAL_LIBERATION_FRONT (FMLN) OF THE CRIME >PE)

CIRCUS triggers a murder CN from "KILLING" which picks up a target = "ATTORNEY GENERAL ROBERTO GARCIA ALVARADO." The subject of the sentence has been recognized as such but does not enter into the murder CN. When CIRCUS encounters the verb "ACCUSED", a clause boundary is recognized. This allows CIRCUS to reset syntactic buffers and pick up "ACCUSED" as a new verb while retaining the previous subject buffer. "ACCUSED" triggers a perpetrator CN with confidence = SUSPECTED_OR_ACCUSED, accuser = "SALVADORAN PRESIDENT-ELECT ALFREDO CRISTIANI", and perpetrator = FARABUNDO_MARTI_NATIONAL_LIBERATION_FRONT. Note that the FMLN is recognized as a terrorist

organization, thereby satisfying a soft constraint in the perpetrator CN. "ACCUSED" tells us to assume a less than factual confidence level within the perpetrator CN, but CIRCUS does not connect the perpetrator CN with any event description. In particular, no attempt is made by CIRCUS to resolve a referent for "the crime." The two resulting CN instantiations look like:

```
-----------------------------------------------------------------------------------------------
TYPE = MURDER
VICTIM = WS-GOVT-OFFICIAL,...     noun group = (ATTORNEY GENERAL ROBERTO GARCIA ALVARADO)

-----------------------------------------------------------------------------------------------
TYPE = PERPETRATOR
CONFIDENCE = SUSPECTED_OR_ACCUSED_BY_AUTHORITIES
ACCUSER = WS-GOVT-OFFICIAL          noun group = (PRESIDENT-ELECT ALFREDO CRISTIANI)
                                    predicates = (SALVADORAN)
PERPETRATOR = WS-ORGANIZATION,...   noun group = (FARABUNDO_MARTI_NATIONAL_LIBERATION_FRONT)
-----------------------------------------------------------------------------------------------
```

MBC's preprocessing and partitioning merges these two CNs into a single event structure before any high-level memory integration is attempted. Incident structures are designed to collapse multiple events (subevents) associated with a single perpetrator into a single structure. The incident structure for S1 looks like:

```
INCIDENT
     DATE = NIL
     LOCATION = NIL
     PERPS = (#S(PERPETRATOR
                 ID NIL
                 ORG (FARABUNDO_MARTI_NATIONAL_LIBERATION_FRONT)
                 WORD-SENSES (WS-TERRORIST WS-ORGANIZATION)
                 CONFIDENCE (SUSPECTED_OR_ACCUSED_BY_AUTHORITIES)
                 NEW-INFO NIL
                 SENTENCE 1))
     NEW = NIL
     PLURAL = NIL
     DISCOURSE-MODE = NIL
     SUBEVENT: NIL
          TARGETS: NIL      EFFECTS: NIL      INSTRUMENT: NIL
          VICTIMS: (#s(VICTIM
                       ID (ROBERTO GARCIA ALVARADO)
                       TITLE (ATTORNEY GENERAL)
                       NATIONALITY NIL
                       NUM 1
                       TYPE (WS-GOVT-OFFICIAL WS-LEGAL-OR-JUDICIAL
                            WS-PROPER-NAME)
                       EFFECTS (DEATH)
                       SENTENCE 1))
```

Because MBC has no incident structures on its incident stack, this new incident structure is added to the stack, and the victim description is added to the victim stack.

S2: (we omit this sentence from the discussion - no alterations to memory are made)

S3: (GARCIA_ALVARADO >CO &&56 >CO WAS KILLED WHEN A BOMB PLACED BY URBAN GUERRILLAS ON HIS VEHICLE EXPLODED AS IT CAME TO A HALT AT AN INTERSECTION IN DOWNTOWN SAN_SALVADOR >PE)

CIRCUS generates 5 CNs in response to this sentence. A simple CN describing a weapon is generated by "BOMB." More complicated CNs are triggered by "KILLED," "PLACED," and "EXPLODED."

The trigger "KILLED" creates a murder CN with victim = "GARCIA ALVARADO."

The trigger "PLACED" creates a location CN with instrument = "BOMB," and actor = "URBAN GUERRILLAS." This same CN also looks for a physical target inside a prepositional phrase, but it misses "ON HIS VEHICLE" because "on" is not one of the prepositions that it predicts. If the sentence had said "outside", "inside", "by", "near", "in", "under", "opposite", "across_from", or "in_front_of", instead of "on", we would have

picked up this physical target. The omission of "on" was a simple oversight in an otherwise legitimate CN definition. This particular CN is specifically predicting a bomb since bombs are frequently the object of the verb "to place" in this domain.

The trigger "EXPLODED" creates a bombing CN with instrument = "A BOMB."

Note that we miss the location San Salvador in S3. Although we have a bottom-up mechanism designed to find dates and locations, it doesn't always work. All 5 CNs are placed in a single partition which generates a new incident structure containing a single subevent:

SUBEVENT: BOMBING
      TARGETS: NIL   EFFECTS: NIL   INSTRUMENT: (#S(INSTRUMENT
      VICTIMS: (#S(VICTIM   ID (BOMB)
              ID (GARCIA ALVARADO)   TYPE WS-BOMB))
              TITLE NIL
              NATIONALITY NIL
              NUM 1
              TYPE (WS-GOVT-OFFICIAL WS-LEGAL-OR-JUDICIAL
                    WS-PROPER-NAME)
              EFFECTS (DEATH)
              SENTENCE 3))

When MBC receives this new incident structure, it runs a memory integration test for compatible target/victim descriptions, and determines that this new subevent is compatible with the incident structure already in memory. MBC therefore merges the two incidents, and memory acquires the fact that Alvarado was killed by a bomb.

S4-7: (we omit these sentences from the discussion - no alterations to memory are made)

S8: (VICE PRESIDENT-ELECT FRANCISCO MERINO SAID THAT WHEN THE ATTORNEY @GENERAL@S CAR STOPPED AT A LIGHT ON A STREET IN DOWNTOWN SAN SALVADOR >CO <u>AN INDIVIDUAL PLACED A BOMB ON THE ROOF OF THE ARMORED VEHICLE</u> >PE)

CIRCUS generates two CNs here. One fairly complicated CN is triggered by "PLACED." This CN picks up not just the bomb as a weapon, but also the individual as the responsible party, and the vehicle as a target. The second CN describes the bomb as a weapon and its link to the targeted vehicle (as before). These two CNs are largely redundant, and they are merged into a single incident structure because they share the same partition. This incident structure contains a perpetrator id = "AN INDIVIDUAL" along with the following subevent:

SUBEVENT: BOMBING
      TARGETS: (#S(PHYS-OBJ   VICTIMS: NIL
              ID (ARMORED VEHICLE)   EFFECTS: NIL
              NUM 1   INSTRUMENT: (#S(INSTRUMENT
              TYPE (WS-TRANSPORT-VEHICLE)   ID (BOMB)
              EFFECTS NIL   TYPE WS-BOMB))
              SENTENCE 8))

MBC checks this incident structure against the incident structure already in memory and determines that they should be merged, thereby picking up a physical target for the first time. Had we picked up this physical target from S3 as well, the target integration test would have merged the two vehicle descriptions at this point as well. Note that MBC merges the description of the perpetrator as "an individual" with the previously encountered descriptor "urban guerrillas" because the earlier description is recognized to be more specific.

S9-10: (we omit these sentences from the discussion - no alterations to memory are made)

S11: (<u>GUERRILLAS ATTACKED @MERINO@S HOME IN SAN_SALVADOR ON APR_14_89</u> >CO &&5 DAYS AGO >CO <u>WITH EXPLOSIVES</u> >PE)

CIRCUS generates 7 highly redundant CNs in response to S11. The most comprehensive CN instantiates an attack with actor = "GUERRILLAS," target = "MERINO'S HOME," and instrument = "EXPLOSIVES." This same

CN also picks up the location (San Salvador) and date (April 14) by the bottom-up attachment mechanism. Locations and dates are normally not predicted by CN definitions, but they can be inserted into available CNs via bottom-up attachment. All of this information is incorporated into a single incident structure containing a bombing subevent (an attack using explosives is understood to be a bombing). The resulting incident structure is then passed to the memory integration portion of MBC.

Just as before, MBC checks to see if the new incident can be merged into the lone incident structure currently stored in memory. But this time the new structure fails to match the existing structure because of incompatible targets. MBC cannot merge a home with a vehicle. When MBC fails to merge the new bombing incident with the old bombing incident, it moves down the target stack to see if there is another incident structure that might merge, but there are no other physical targets in memory. MBC adds the new incident to the top of the incident stack, and memory now contains two bombing incidents.

S12: (THERE WERE &&7 CHILDREN >CO INCLUDING &&4 OF THE VICE @PRESIDENT@S CHILDREN >CO IN THE HOME AT THE TIME >PE)

CIRCUS produces no output for this sentence because no CN triggers are encountered. We sometimes miss information in sentences where the only verb is a form of "to be."

S13: (A 15-YEAR-OLD NIECE OF @MERINO@S WAS INJURED >PE)

CIRCUS generates an injury CN with victim = "A 15-YEAR-OLD NIECE." This results in a subevent of unknown type with a victim id = "A 15-YEAR-OLD NIECE." When MBC receives this incident, it examines the first incident on the top of its stack to see if a merge is possible. Since no incompatible victims are found in memory for this incident (the latest bombing incident specifies no victims), a merging occurs.

S14-S17: [we omit these sentences from our discussion - no alterations are made to memory.]

S18: (RICARDO VALDIVIESO >CO PRESIDENT OF THE LEGISLATIVE ASSEMBLY AND AN ARENA LEADER >CO SAID THE FMLN AND ITS FRONT GROUPS ARE RESPONSIBLE FOR THE "IRRATIONAL VIOLENCE THAT KILLED ATTORNEY GENERAL GARCIA >DQ >PE)

CIRCUS produces a murder CN with victim = "Attorney General Garcia" and actor = "irrational violence." This CN has a soft constraint on the actor slot which specifies a human or organization, but the CN survives the CN filter because its other variable slot has a filler that does meet the required soft constraints (the filter errs on the side of spurious information if one slot looks good and the other slot looks bad). MBC is careful to check available soft constraints when it integrates information into its preliminary incident structures. Any slot fill that violates a soft constraint is discarded at that time.

When MBC attempts to integrate this incident into memory, it locates a compatible target in the victim stack, and merges the new incident structure with the existing structure that describes Garcia as a victim. Because we have now merged new information into an incident that was not at the top of the incident stack, we have to reorder the incident stack by moving the most recently referenced incident to the top of the stack. This effectively identifies the first incident as the current topic once again. Ideally, this would set us up to correctly integrate information contained later in S21 and S22 where new information is presented about the vehicle bombing, but CIRCUS fails to pick up the additional human targets from those sentences, so the topic shift that we've successfully recognized at S18 goes unrewarded.

When MBC completes its analysis, the two bombing incident structures are converted into two template instantiations, along with a third threat incident picked up from additional sentences near the end of the text. In order to instantiate the final templates, we rely on semantic features in our dictionary to recognize a home as a civilian residence and an armored vehicle as a transport vehicle. When instantiating response templates, we attempt to fill all slots with the exception of phys-tgt-total-num and hum-tgt-total-num.

We did fairly well on the first template (see Figure 1). We missed San Salvador as the location within El Salvador, we said the vehicle was destroyed instead of damaged, and we missed 3 human targets (the driver who was not hurt, and the 2 bodyguards, one of whom was injured). All the other slots were correctly filled. On the second template, we fail in three places. We have no perpetrator organization, we miss the physical target type for Merino's

```
0. MESSAGE: ID                      TST2-MUC4-0048
1. MESSAGE: TEMPLATE                 1                                          ;correct
2. INCIDENT: DATE                    - 19 APR 89                                ;correct
3. INCIDENT: LOCATION                EL SALVADOR                                :partial
4. INCIDENT: TYPE                    BOMBING                                    ;correct
5. INCIDENT: STAGE OF EXEC.          ACCOMPLISHED                               ;correct
6. INCIDENT: INSTRUMENT ID           "BOMB"                                     ;correct
7. INCIDENT: INSTRUMENT TYPE         BOMB: "BOMB"                               ;correct
8. PERP: INCIDENT CATEGORY           TERRORIST ACT                              ;correct
9. PERP: INDIVIDUAL ID               "URBAN GUERRILLAS"                         ;correct
10: PERP: ORGANIZATION ID            "FARABUNDO MARTI NATIONAL LIBERATION       ;correct
                                        FRONT"
11: PERP: ORG CONFIDENCE             SUSPECTED OR ACCUSED BY AUTHORITIES:       ;correct
                                        "FARABUNDO MARTI NATIONAL
                                        LIBERATION FRONT"
12: PHYS TGT: ID                     "ARMORED VEHICLE"                          ;correct
13: PHYS TGT: TYPE                   TRANSPORT VEHICLE: "ARMORED VEHICLE"       ;correct
14: PHYS TGT: NUMBER                 1: "ARMORED VEHICLE"                       ;correct
15: PHYS TGT: FOREIGN NATION         -                                          ;N/A
16. PHYS TGT: EFFECT                 DESTROYED: "ARMORED VEHICLE"               ;partial
17: PHYS TGT: TOTAL NUMBER           -                                          ;N/A
18: HUM TGT: NAME                    "ROBERTO GARCIA ALVARADO"                  ;correct
19: HUM TGT: DESCRIPTION             "ATTORNEY GENERAL": "ROBERTO GARCIA        ;correct/missing
                                        ALVARADO"
20: HUM TGT: TYPE                    GOVERNMENT OFFICIAL: "ROBERTO GARCIA       ;correct/missing
                                        ALVARADO"
21: HUM TGT: NUMBER                  1: "ROBERTO GARCIA ALVARADO"               ;correct/missing
22. HUM TGT: FOREIGN NATION          -                                          ;N/A
23: HUM TGT: EFFECT                  DEATH: "ROBERTO GARCIA ALVARADO"           ;correct/missing
24: HUM TGT: TOTAL NUMBER            -                                          ;N/A
```

Figure 1: Our response template for the first bombing incident

home (it should have been GOVERNMENT OFFICE OR RESIDENCE), and we are missing the 7 children that were human targets (this is one of the few texts where a hum-tgt-total-num slot should receive a value).

Overall, TST2-MUC4-0048 showed the UMass/MUC-4 system working fairly well and not making any major errors. Most of our recall loss resulted from a failure to recognize relevant information in S12 (the 7 children), S21 and S22 (the driver and 2 bodyguards). As we saw in this message, we can recover from some failures in sentence analysis when a text provides redundant descriptions (e.g. we missed the physical target in S3, but picked it up correctly in S8). When memory-based consolidation responds correctly to topic transitions, the output that CIRCUS generates usually makes it into the correct places in the response templates. TST2-MUC4-0048 shows how MBC was able to correctly recognize two topic transitions: first from an old incident to a new incident, and then back again to the earlier incident. Given that the errors encountered for TST2-MUC4-0048 were relatively minor (one could even argue that the third template was valid and should have been covered by an optional key template), there is nothing here that illustrates the more damaging problems that impacted our TST3 and TST4 score reports.

Figure 2 shows score reports for the two templates that mapped to TST2-MUC4-0048 answer keys, along with the score report for the entire message which averages in the spurious template that we generated for the threat. This final score report for the whole message illustrates how much negative impact spurious templates have on precision if a system is generating one spurious template for every two good templates. If we had generated a summary score report based on only two templates instead of three, our All Templates precision would have been 94. With the third template averaged in, our All Templates precision drops to 76.

Vehicle Bombing Template

	POS	ACT	COR	PAR	INC	ACR	IPA	SPU	MIS	NON	REC	PRE	OVG
inc-total	6	6	5	1	0	0	0	0	0	0	92	92	0
perp-total	4	4	4	0	0	0	0	0	0	0	100	100	0
phys-tgt-total	4	4	3	1	0	0	0	0	0	2	88	88	0
hum-tgt-total	14	5	5	0	0	0	0	0	9	2	36	100	0
TOTAL	28	19	17	2	0	0	0	0	9	4	64	95	0

Home Bombing Template

	POS	ACT	COR	PAR	INC	ACR	IPA	SPU	MIS	NON	REC	PRE	OVG
inc-total	6	6	6	0	0	0	0	0	0	0	100	100	0
perp-total	4	2	2	0	0	0	0	0	2	0	50	100	0
phys-tgt-total	3	3	2	0	1	0	0	0	0	3	67	67	0
hum-tgt-total	11	4	4	0	0	0	0	0	7	2	36	100	0
TOTAL	24	15	14	0	1	0	0	0	9	5	58	93	0

Total Scores for TST2-MUC4-0048

	POS	ACT	COR	PAR	INC	ACR	IPA	SPU	MIS	NON	REC	PRE	OVG
inc-total	12	16	11	1	0	0	0	4	0	2	96	72	25
perp-total	8	7	6	0	0	0	0	1	2	3	75	86	14
phys-tgt-total	7	7	5	1	1	0	0	0	0	11	78	78	0
hum-tgt-total	25	12	9	0	0	0	0	3	16	8	36	75	25
MATCHED/MISSING	52	34	31	2	1	0	0	0	18	9	62	94	0
MATCHED/SPURIOUS	52	42	31	2	1	0	0	8	18	24	62	76	19
MATCHED ONLY	52	34	31	2	1	0	0	0	18	9	62	94	0
ALL TEMPLATES	52	42	31	2	1	0	0	8	18	24	62	76	19
SET FILLS ONLY	23	16	14	1	1	0	0	0	7	5	63	91	0
STRING FILLS ONLY	15	10	10	0	0	0	0	0	5	1	67	100	0

P&R	2P&R	P&2R
68.29	72.72	64.37

Figure 2: Partial and Overall Scores for TST2-MUC4-0048

In a domain that is characterized by complicated domain guidelines, and lots of grey areas, answer keys cannot be trusted to give encodings that are necessarily superior to the output of a high performance extraction system. If this is the case, it may be very difficult to attain 85% precision under all templates, and optimal precision levels may be closer to the 70-80% range.

BIBLIOGRAPHY

[1] Cardie, C. and Lehnert, W. (1991) "A Cognitively Plausible Approach to Understanding Complex Syntax" in *Proceedings of the Ninth National Conference on Artificial Intelligence*. Anaheim, CA. pp. 117-124.

[2] Lehnert, W. (1991) "Symbolic/Subsymbolic Sentence Analysis: Exploiting the Best of Two Worlds" in *Advances in Connectionist and Neural Computation Theory*. Vol 1. (eds: J. Pollack and J. Barnden). pp. 135-164. Ablex Publishing. Norwood, NJ. pp. 135-164.

[3] Lehnert, W., Cardie, C., Fisher, D., Riloff, E., Williams, R. (1991a) "University of Massachusetts: MUC-3 Test Results and Analysis" in *Proceedings of the Third Message Understanding Conference*. Morgan Kaufman. San Mateo, CA. pp. 116-119.

[4] Lehnert, W., Cardie, C., Fisher, D., Riloff, E., Williams, R. (1991b) "University of Massachusetts: Description of the CIRCUS System as Used for MUC-3" in *Proceedings of the Third Message Understanding Conference*. Morgan Kaufman. San Mateo, CA. pp. 223-233.

[5] Lehnert, W., Williams, R., Cardie, C, Riloff, E., and Fisher, D.(1991c) "The CIRCUS System as Used in MUC-3," Technical Report No. 91-59, Department of Computer and Information Science, University of Massachusetts. 1991.