Covering Treebanks with GLARF

Adam Meyers† and **Ralph Grishman**† and **Michiko Kosaka**‡ and **Shubin Zhao**†
†New York University, 719 Broadway, 7th Floor, NY, NY 10003 USA
‡Monmouth University, West Long Branch, N.J. 07764, USA
meyers/grishman/shubinz@cs.nyu.edu, kosaka@monmouth.edu

Abstract

This paper introduces GLARF, a framework for predicate argument structure. We report on converting the Penn Treebank II into GLARF by automatic methods that achieved about 90% precision/recall on test sentences from the Penn Treebank. Plans for a corpus of hand-corrected output, extensions of GLARF to Japanese and applications for MT are also discussed.

1 Introduction

Applications using annotated corpora are often, by design, limited by the information found in those corpora. Since most English treebanks provide limited predicate-argument (PRED-ARG) information, parsers based on these treebanks do not produce more detailed predicate argument structures (PRED-ARG structures). The Penn Treebank II (Marcus et al., 1994) marks subjects (SBJ), logical objects of passives (LGS), some reduced relative clauses (RRC), as well as other grammatical information, but does not mark each constituent with a grammatical role. In our view, a full PRED-ARG description of a sentence would do just that: assign each constituent a grammatical role that relates that constituent to one or more other constituents in the sentence. For example, the role HEAD relates a constituent to its parent and the role OBJ relates a constituent to the HEAD of its parent. We believe that the absence of this detail limits the range of applications for treebank-based parsers. In particular, they limit the extent to which it is possible to generalize, e.g., marking IND-OBJ and OBJ roles allows one to generalize a single pattern to cover two related examples ("John gave Mary a book" = "John gave a book to Mary"). Distinguishing complement PPs (COMP) from adjunct PPs (ADV) is useful because the former is likely to have an idiosyncratic interpretation, e.g., the object of "at" in "John is angry at Mary" is not a locative and should be distinguished from the locative case by many applications.

In an attempt to fill this gap, we have begun a project to add this information using both automatic procedures and hand-annotation. We are implementing automatic procedures for mapping the Penn Treebank II (PTB) into a PRED-ARG representation and then we are correcting the output of these procedures manually. In particular, we are hoping to encode information that will enable a greater level of regularization across linguistic structures than is possible with PTB.

This paper introduces GLARF, the Grammatical and Logical Argument Representation Framework. We designed GLARF with four objectives in mind: (1) capturing regularizations — noncanonical constructions (e.g., passives, filler-gap constructions, etc.) are represented in terms of their canonical counterparts (simple declarative clauses); (2) representing all phenomena using one simple data structure: the typed feature structure (3) consistently labeling all arguments and adjuncts for phrases with clear heads; and (4) producing clear and consistent PRED-ARGs for phrases that do not have heads, e.g., conjoined structures, named entities, etc. — rather than trying to squeeze these phrases into an X-bar mold, we customized our representations to reflect their head-less properties. We believe that a framework for PRED-ARG needs to satisfy these objectives to adequately cover a corpus like PTB.

We believe that GLARF, because of its uniform treatment of PRED-ARG relations, will be valuable for many applications, including question answering, information extraction, and machine translation. In particular, for MT, we ex-

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2001		**00-00-2001 to 00-00-2001**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Covering Treebanks with GLARF	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Computer Science ,New York University,715 Broadway,New York,NY,10003**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**8**	

pect it will benefit procedures which learn translation rules from syntactically analyzed parallel corpora, such as (Matsumoto et al., 1993; Meyers et al., 1996). Much closer alignments will be possible using GLARF, because of its multiple levels of representation, than would be possible with surface structure alone (An example is provided at the end of Section 2). For this reason, we are currently investigating the extension of our mapping procedure to treebanks of Japanese (the Kyoto Corpus) and Spanish (the UAM Treebank (Moreno et al., 2000)). Ultimately, we intend to create a parallel trilingual treebank using a combination of automatic methods and human correction. Such a treebank would be valuable resource for corpus-trained MT systems.

The primary goal of this paper is to discuss the considerations for adding PRED-ARG information to PTB, and to report on the performance of our mapping procedure. We intend to wait until these procedures are mature before beginning annotation on a larger scale. We also describe our initial research on covering the Kyoto Corpus of Japanese with GLARF.

2 Previous Treebanks

There are several corpora annotated with PRED-ARG information, but each encode some distinctions that are different. The Susanne Corpus (Sampson, 1995) consists of about 1/6 of the Brown Corpus annotated with detailed syntactic information. Unlike GLARF, the Susanne framework does not guarantee that each constituent be assigned a grammatical role. Some grammatical roles (e.g., subject, object) are marked explicitly, others are implied by phrasetags (Fr corresponds to the GLARF node label SBAR under a RELATIVE arc label) and other constituents are not assigned roles (e.g., constituents of NPs). Apart from this concern, it is reasonable to ask why we did not adapt this scheme for our use. Susanne's granularity surpasses PTB-based GLARF in many areas with about 350 wordtags (part of speech) and 100 phrasetags (phrase node labels). However, GLARF would express many of the details in other ways, using fewer node and part of speech (POS) labels and more attributes and role labels. In the feature structure tradition, GLARF can represent varying levels of detail by adding or subtracting attributes or defining subsumption hierarchies. Thus both Susanne's NP1p wordtag and Penn's NNP wordtag would correspond to GLARF's NNP POS tag. A GLARF-style Susanne analysis of "Ontario, Canada" is (NP (PROVINCE (NNP Ontario)) (PUNCTUATION (, ,)) (COUNTRY (NNP Canada)) (PATTERN NAME) (SEM-FEATURE LOC)). A GLARF-style PTB analysis uses the roles NAME1 and NAME2 instead of PROVINCE and COUNTRY, where name roles (NAME1, NAME2) are more general than PROVINCE and COUNTRY in a subsumption hierarchy. In contrast, attempts to convert PTB into Susanne would fail because detail would be unavailable. Similarly, attempts to convert Susanne into the PTB framework would lose information. In summary, GLARF's ability to represent varying levels of detail allows different types of treebank formats to be converted into GLARF, even if they cannot be converted into each other. Perhaps, GLARF can become a lingua franca among annotated treebanks.

The Negra Corpus (Brants et al., 1997) provides PRED-ARG information for German, similar in granularity to GLARF. The most significant difference is that GLARF regularizes some phenomena which a Negra version of English would probably not, e.g., control phenomena. Another novel feature of GLARF is the ability to represent paraphrases (in the Harrisian sense) that are not entirely syntactic, e.g., nominalizations as sentences. Other schemes seem to only regularize strictly syntactic phenomena.

3 The Structure of GLARF

In GLARF, each sentence is represented by a typed feature structure. As is standard, we model feature structures as single-rooted directed acyclic graphs (DAGs). Each nonterminal is labeled with a phrase category, and each leaf is labeled with either: (a) a (PTB) POS label and a word (*eat*, *fish*, etc.) or (b) an attribute value (e.g., singular, passive, etc.). Types are based on nonterminal node labels, POSs and other attributes (Carpenter, 1992). Each arc bears a feature label which represents either a grammatical role (SBJ, OBJ, etc.) or some attribute of a word or phrase (morphological features, tense, semantic features,

etc.).[1] For example, the subject of a sentence is the head of a SBJ arc, an attribute like SINGULAR is the head of a GRAM-NUMBER arc, etc. A constituent involved in multiple surface or logical relations may be at the head of multiple arcs. For example, the surface subject (S-SBJ) of a passive verb is also the logical object (L-OBJ). These two roles are represented as two arcs which share the same head. This sort of structure sharing analysis originates with Relational Grammar and related frameworks (Perlmutter, 1984; Johnson and Postal, 1980) and is common in Feature Structure frameworks (LFG, HPSG, etc.). Following (Johnson et al., 1993)[2], arcs are typed. There are five different types of role labels:

- Attribute roles: Gram-Number (grammatical number), Mood, Tense, Sem-Feature (semantic features like temporal/locative), etc.
- Surface-only relations (prefixed with S-), e.g., the surface subject (S-SBJ) of a passive.
- Logical-only Roles (prefixed with L-), e.g., the logical object (L-OBJ) of a passive.
- Intermediate roles (prefixed with I-) representing neither surface, nor logical positions. In "John seemed to be kidnapped by aliens", "John" is the surface subject of "seem", the logical object of "kidnapped", and the intermediate subject of "to be". Intermediate arcs capture are helpful for modeling the way sentences conform to constraints. The intermediate subject arc obeys lexical constraints and connect the surface subjects of "seem" (COMLEX Syntax class TO-INF-RS (Macleod et al., 1998a)) to the subject of the infinitive. However, the subject of the infinitive in this case is not a logical subject due to the passive. In some cases, intermediate arcs are subject to number agreement, e.g., in "Which aliens did you say were seen?", the I-SBJ of "were seen" agrees with "were".
- Combined surface/logical roles (unprefixed arcs, which we refer to as SL- arcs). For example, "John" in "John ate cheese" would be the target of a SBJ subject arc.

Logical relations, encoded with SL- and L-arcs, are defined more broadly in GLARF than in most frameworks. Any regularization from a non-canonical linguistic structure to a canonical one results in logical relations. Following (Harris, 1968) and others, our model of canonical linguistic structure is the tensed active indicative sentence with no missing arguments. The following argument types will be at the head of logical (L-) arcs based on counterparts in canonical sentences which are at the head of SL- arcs: logical arguments of passives, understood subjects of infinitives, understood fillers of gaps, and interpreted arguments of nominalizations (In "Rome's destruction of Carthage", "Rome" is the logical subject and "Carthage" is the logical object). While canonical sentence structure provides one level of regularization, canonical verb argument structures provide another. In the case of argument alternations (Levin, 1993), the same role marks an alternating argument regardless of where it occurs in a sentence. Thus "the man" is the indirect object (IND-OBJ) and "a dollar" is the direct object (OBJ) in both "She gave the man a dollar" and "She gave a dollar to the man" (the dative alternation). Similarly, "the people" is the logical object (L-OBJ) of both "The people evacuated from the town" and "The troops evacuated the people from the town", when we assume the appropriate regularization. Encoding this information allows applications to generalize. For example, a single Information Extraction pattern that recognizes the IND-OBJ/OBJ distinction would be able to handle these two examples. Without this distinction, 2 patterns would be needed.

Due to the diverse types of logical roles, we sub-type roles according to the type of regularization that they reflect. Depending on the application, one can apply different filters to a detailed GLARF representation, only looking at certain types of arcs. For example, one might choose all logical (L- and SL-) roles for an application that is trying to acquire selection restrictions, or all surface (S- and SL-) roles if one was interested in obtaining a surface parse. For other applications, one might want to choose between subtypes of logical arcs. Given

[1] A few grammatical roles are nonfunctional, e.g., a constituent can have multiple ADV constituents. We number these roles (ADV1, ADV2, . . .) to preserve functionality.

[2] That paper uses two arc types: category and relational.

```
(S (NP-SBJ (PRP they))
   (VP (VP (VBD spent)
       (NP-2 ($ $)
     (CD 325,000)
              (-NONE- *U*))
(PP-TMP-3 (IN in)
            (NP (CD 1989))))
   (CC and)
   (VP (NP=2 ($ $)
             (CD 340,000)
             (-NONE- *U*))
       (PP-TMP=3 (IN in)
           (NP (CD 1990))))))
```

Figure 1: Penn representation of gapping

a trilingual treebank, suppose that a Spanish treebank sentence corresponds to a Japanese nominalization phrase and an English nominalization phrase, e.g.,

Disney ha comprado Apple Computers
Disney's acquisition of Apple Computers
ディズニーのアップルコンピューターの買取。

Furthermore, suppose that the English treebank analyzes the nominalization phrase both as an NP (Disney = possessive, Apple Computers = object of preposition) and as a paraphrase of a sentence (Disney = subject, Apple Computers = object). For an MT system that aligns the Spanish and English graph representation, it may be useful to view the nominalization phrase in terms of the clausal arguments. However, in a Japanese/English system, we may only want to look at the structure of the English nominalization phrase as an NP.

4 GLARF and the Penn Treebank

This section focuses on some characteristics of English GLARF and how we map PTB into GLARF, as exemplified by mapping the PTB representation in Figure 1 to the GLARF representation in Figure 2. In the process, we will discuss how some of the more interesting linguistic phenomena are represented in GLARF.

4.1 Mapping into GLARF

Our procedure for mapping PTB into GLARF uses a sequence of transformations. The first transformation applies to PTB, and the output of each $transformation_n$ is the input of $transformation_{n+1}$. As many of these transformations are trivial, we focus on the most interesting set of problems. In addition, we explain how GLARF is used to represent some of the more difficult phenomena.

(Brants et al., 1997) describes an effort to minimize human effort in the annotation of raw text with comparable PRED-ARG information. In contrast, we are starting with annotated corpus and want to add as much detail as possible automatically. We are as much concerned with finding good procedures for PTB-based parser output as we are minimizing the effort of future human taggers. The procedures are designed to get the right answer most of the time. Human taggers will correct the results when they are wrong.

4.1.1 Conjunctions

The treatment of coordinate conjunction in PTB is not uniform. Words labeled CC and phrases labeled CONJP usually function as coordinate conjunctions in PTB. However, a number of problems arise when one attempts to unambiguously identify the phrases which are conjoined. Most significantly, given a phrase XP with conjunctions and commas and some set of other constituents $Y_1, \ldots, Y_n$, it is not always clear which Y_i are conjuncts and which are not, i.e., Penn does not explicitly mark items as conjuncts and one cannot assume that all Y_i are conjuncts. In GLARF, conjoined phrases are clearly identified and conjuncts in those phrases are distinguished from non-conjuncts. We will discuss each problematic case that we observed in turn.

Instances of words that are marked CC in Penn do not always function as conjunctions. They may play the role of a sentential adverb, a preposition or the head of a parenthetical constituents. In GLARF, conjoined phrases are explicitly marked with the attribute value (CONJOINED T). The mapping procedures recognize that phrases beginning with CCs, PRN phrases containing CCs, among others are not conjoined phrases.

A sister of a conjunction (other than a conjunction) need not be a conjunct. There are two cases. First of all, a sister of a conjunction can be a shared modifier, e.g., the right node raised

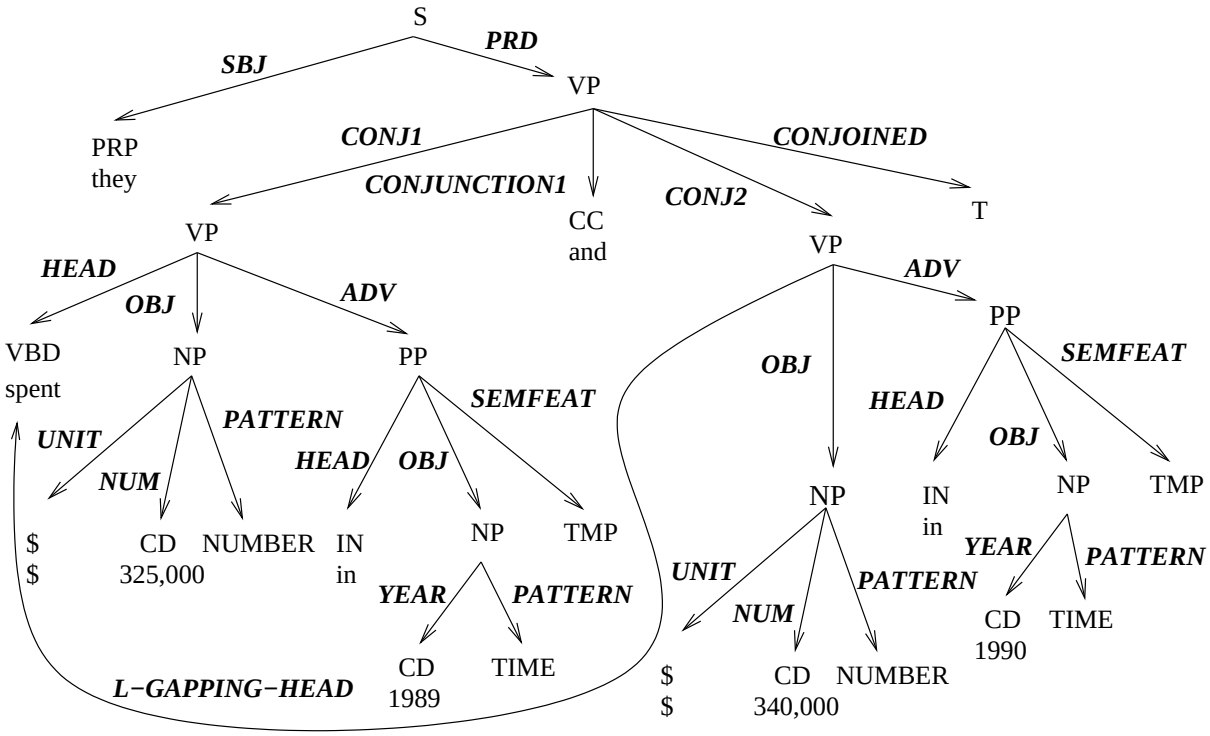

Figure 2: GLARF representation of gapping

PP modifier in "[NP senior vice president] and [NP general manager] [PP of this U.S. sales and marketing arm]"; and the locative "there" in "deterring U.S. high-technology firms from [investing or [marketing their best products] there]". In addition, the boundaries of the conjoined phrase and/or the conjuncts that they contain are omitted in some environments, particularly when single words are conjoined and/or when the phrases occur before the head of a noun phrase or quantifier phrase. Some phrases which are under a single nonterminal node in the treebank (and are not further broken down) include the following: "between $190 million and $195 million", "Hollingsworth & Vose Co.", "cotton and acetate fibers", "those workers and managers", "this U.S. sales and marketing arm", and "Messrs. Cray and Barnum". To overcome this sort of problem, procedures introduce brackets and mark constituents as conjuncts. Considerations included POS categories, similarity measures, construction type (e.g., & is typically part of a name), among other factors.

CONJPs have a different distribution than CCs. Different considerations are needed for identifying the conjuncts. CONJPs, unlike CCs, can occur initially, e.g., "[Not only] [was Fred a good doctor], [he was a good friend as well].). Secondly, they can be embedded in the first conjunct, e.g., "[Fred, not only, liked to play doctor], [he was good at it as well.]".

In Figure 2, the conjuncts are labeled explicitly with their roles CONJ1 and CONJ2, the conjunction is labeled as CONJUNCTION1 and the top-most VP is explicitly marked as a conjoined phrase with the attribute/value (CONJOINED T).

4.1.2 Applying Lexical Resources

We merged together two lexical resources NOMLEX (Macleod et al., 1998b) and COMLEX Syntax 3.1 (Macleod et al., 1998a), deriving PP complements of nouns from NOMLEX and using COMLEX for other types of lexical information.We use these resources to help add additional brackets, make additional role distinctions and fill a gap when its filler is not marked in PTB. Although Penn's -CLR tags are good indicators of complement-hood, they only apply to verbal complements. Thus procedures for making adjunct/complement distinctions benefited from the dictionary classes. Similarly, COMLEX's

NP-FOR-NP class helped identify those -BNF constituents which were indirect objects ("John baked Mary a cake", "John baked a cake [for Mary]"). The class PRE-ADJ identified those adverbial modifiers within NPs which really modify the adjective. Thus we could add the following brackets to the NP: "[even brief] exposures". NTITLE and NUNIT were useful for the analysis of pattern type noun phrases, e.g., "President Bill Clinton", "five million dollars". Our procedures for identifying the logical subjects of infinitives make extensive use of the control/raising properties of COMLEX classes. For example, X is the subject of the infinitives in "X appeared to leave" and "X was likely to bring attention to the problem".

4.1.3 NEs and Other Patterns

Over the past few years, there has been a lot of interest in automatically recognizing named entities, time phrases, quantities, among other special types of noun phrases. These phrases have a number of things in common including: (1) their internal structure can have idiosyncratic properties relative to other types of noun phrases, e.g., person names typically consist of optional titles plus one or more names (first, middle, last) plus an optional post-honorific; and (2) externally, they can occur wherever some more typical phrasal constituent (usually NP) occurs. Identifying these patterns makes it possible to describe these differences in structure, e.g., instead of identifying a head for "John Smith, Esq.", we identify two names and a posthonorific. If this named entity went unrecognized, we would incorrectly assume that "Esq." was the head. Currently, we merge the output of a named entity tagger to the Penn Treebank prior to processing. In addition to NE tagger output, we use procedures based on Penn's proper noun wordtags.

In Figure 2, there are four patterns: two NUMBER and two TIME patterns. The TIME patterns are very simple, each consisting just of YEAR elements, although MONTH, DAY, HOUR, MINUTE, etc. elements are possible. The NUMBER patterns each consist of a single NUMBER (although multiple NUMBER constituents are possible, e.g., "one thousand") and one UNIT constituent. The types of these patterns are indicated by the PATTERN attribute.

4.1.4 Gapping Constructions

Figures 1 and 2 are corresponding PTB and GLARF representations of gapping. Penn represents gapping via "parallel" indices for corresponding arguments. In GLARF, the shared verb is at the head of two HEAD arcs. GLARF overcomes some problems with structure sharing analyses of gapping constructions. The verb gap is a "sloppy" (Ross, 1967) copy of the original verb. Two separate spending events are represented by one verb. Intuitively, structure sharing implies token identity, whereas type identity would be more appropriate. In addition, the copied verb need not agree with the subject in the second conjunct, e.g., "was", not "were" would agree with the second conjunct in "the risks $were_i$ too high and the potential payoff e_i too far in the future". It is thus problematic to view the gap as identical in every way to the filler in this case. In GLARF, we can thus distinguish the gapping sort of logical arc (L-GAPPING-HEAD) from the other types of L-HEAD arcs. We can stipulate that a gapping logical arc represents an appropriately inflected copy of the phrase at the head of that arc.

In GLARF, the predicate is always explicit. However, Penn's representation (H. Koti, pc) provides an easy way to represent complex cases, e.g., "John wanted to buy gold, and Mary *gap* silver. In GLARF, the gap would be filled by the nonconstituent "wanted to buy". Unfortunately, we believe that this is a necessary burden. A goal of GLARF is to explicitly mark all PRED-ARG relations. Given parallel indices, the user must extract the predicate from the text by (imperfect) automatic means. The current solution for GLARF is to provide multiple gaps. The second conjunct of the example in question would have the following analysis: (S (SBJ $Mary_j$) (PRD (VP (HEAD GAP_i) (COMP (S NP_j (PRD (VP (HEAD GAP_k) (OBJ silver))))))))), where GAP_i is filled by "wanted", GAP_k is filled by "to buy" and NP_j is bound to Mary.

5 Japanese GLARF

Japanese GLARF will have many of the same specifications described above. To illustrate how we will extend GLARF to Japanese, we discuss

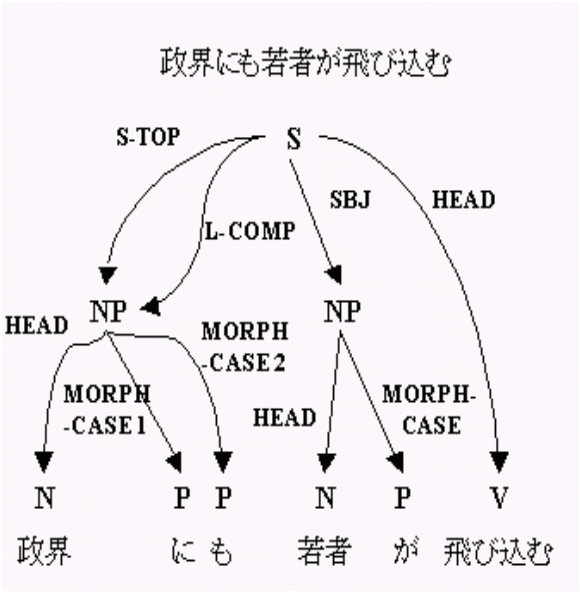

Figure 3: Stacked Postpositions in GLARF

two difficult-to-represent phenomena: elision and stacked postpositions.

Grammatical analyses of Japanese are often dependency trees which use postpositions as arc labels. Arguments, when elided, are omitted from the analysis. In GLARF, however, we use role labels like SBJ, OBJ, IND-OBJ and COMP and mark elided constituents as zeroed arguments. In the case of stacked postpositions, we represent the different roles via different arcs. We also reanalyze certain postpositions as being complementizers (subordinators) or adverbs, thus excluding them from canonical roles. By reanalyzing this way, we arrived at two types of true stacked postpositions: nominalization and topicalization. For example, in Figure 3, the topicalized NP is at the head of two arcs, labeled S-TOP and L-COMP and the associated postpositions are analyzed as morphological case attributes.

6 Testing the Procedures

To test our mapping procedures, we apply them to some PTB files and then correct the resulting representation using ANNOTATE (Brants and Plaehn, 2000), a program for annotating edge-labeled trees and DAGs, originally created for the NEGRA corpus. We chose both files that we have used extensively to tune the mapping procedures (training) and other files. We then convert the resulting GLARF Feature Structures into triples of the form {Role-Name Pivot Non-Pivot} for all logical arcs (cf. (Caroll et al., 1998)), using some automatic procedures. The "pivot" is the head of headed structures, but may be some other constituent in non-headed structures. For example, in a conjoined phrase, the pivot is the conjunction, and the head would be the list of heads of the conjuncts. Rather than listing the whole Pivot and non-pivot phrases in the triples, we simply list the heads of these phrases, which is usually a single word. Finally, we compute precision and recall by comparing the triples generated from our procedures to triples generated from the corrected GLARF.[3] An exact match is a correct answer and anything else is incorrect.[4]

6.1 The Test and the Results

We developed our mapping procedures in two stages. We implemented some mapping procedures based on PTB manuals, related papers and actual usage of labels in PTB. After our initial implementation, we tuned the procedures based on a training set of 64 sentences from two PTB files: wsj_0003 and wsj_0051, yielding 1285 + triples. Then we tested these procedures against a test set consisting of 65 sentences from wsj_0089 (1369 triples). Our results are provided in Figure 4. Precision and recall are calculated on a per sentence basis and then averaged. The precision for a sentence is the number of correct triples divided by the total number of triples generated. The recall is the total number of correct triples divided by the total number of triples in the answer key.

Out of 187 incorrect triples in the test corpus, 31 reflected the incorrect role being selected, e.g., the adjunct/complement distinction, 139 reflected errors or omissions in our procedures and 7 triples related to other factors. We expect a sizable improvement as we increase the size of our training corpus and expand the coverage of our pro-

[3] We admit a bias towards our output in a small number of cases (less than 1%). For example, it is unimportant whether "exposed to it" modifies "the group" or "workers" in "a group of workers exposed to it". The output will get full credit for this example regardless of where the reduced relative is attached.

[4] (Caroll et al., 1998) report about 88% precision and recall for similar triples derived from parser output. However, they allow triples to match in some cases when the roles are different and they do not mark modifier relations.

Data	Sentences	Recall	Precision
Training	64	94.4	94.3
Test	65	89.0	89.7

Figure 4: Results

cedures, particularly since one omission often resulted in several incorrect triples.

7 Concluding Remarks

We show that it is possible to automatically map PTB input into PRED-ARG structure with high accuracy. While our initial results are promising, mapping procedures are limited by available resources. To produce the best possible GLARF resource, hand correction will be necessary.

We are improving our mapping procedures and extending them to PTB-based parser output. We are creating mapping procedures for the Susanne corpus, the Kyoto Corpus and the UAM Treebank. This work is a precursor to the creation of a trilingual GLARF treebank.

We are currently defining the problem of mapping treebanks into GLARF. Subsequently, we intend to create standardized mapping rules which can be applied by any number of algorithms. The end result may be that detailed parsing can be carried out in two stages. In the first stage, one derives a parse at the level of detail of the Penn Treebank II. In the second stage, one derives a more detailed parse. The advantage of such division should be obvious: one is free to find the best procedures for each stage and combine them. These procedures could come from different sources and use totally different methods.

Acknowledgements

This research was supported by the Defense Advanced Research Projects Agency under Grant N66001-00-1-8917 from the Space and Naval Warfare Systems Center, San Diago and by the National Science Foundation under Grant IIS-0081962.

References

T. Brants and O. Plaehn. 2000. Interactive corpus annotation. *LREC 2000*, pages 453–459.

T. Brants, W. Skut, and B. Krenn. 1997. Tagging Grammatical Functions. In *EMNLP-2*.

J. Caroll, T. Briscoe, and A. Sanfillippo. 1998. Parse Evaluation: a Survey and a New Proposal. *LREC 1998*, pages 447–454.

B. Carpenter. 1992. *The Logic of Typed Features*. Cambridge University Press, New York.

Z. Harris. 1968. *Mathematical Structures of Language*. Wiley-Interscience, New York.

D. Johnson and P. Postal. 1980. *Arc Pair Grammar*. Princeton University Press, Princeton.

D. Johnson, A. Meyers, and L. Moss. 1993. A Unification-Based Parser for Relational Grammar. *ACL 1993*, pages 97–104.

B. Levin. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago.

C. Macleod, R. Grishman, and A. Meyers. 1998a. COMLEX Syntax. *Computers and the Humanities*, 31(6):459–481.

C. Macleod, R. Grishman, A. Meyers, L. Barrett, and R. Reeves. 1998b. Nomlex: A lexicon of nominalizations. *Euralex98*.

M. Marcus, G. Kim, M. A. Marcinkiewicz, R. MacIntyre, A. Bies, M. Ferguson, K. Katz, and B. Schasberger. 1994. The penn treebank: Annotating predicate argument structure. In *Proceedings of the 1994 ARPA Human Language Technology Workshop*.

Y. Matsumoto, H. Ishimoto, T. Utsuro, and M. Nagao. 1993. Structural Matching of Parallel Texts. In *ACL 1993*

A. Meyers, R. Yangarber, and R. Grishman. 1996. Alignment of Shared Forests for Bilingual Corpora. *Coling 1996*, pages 460–465.

A. Meyers, M. Kosaka, and R. Grishman. 2000. Chart-Based Transfer Rule Application in Machine Translation. *Coling 2000*, pages 537–543.

A. Moreno, R. Grishman, S. Lopez, F. Sanchez, and S. Sekine. 2000. A treebank of Spanish and its application to parsing. *LREC*, pages 107–111.

D. Perlmutter. 1984. *Studies in Relational Grammar 1*. University of Chicago Press, Chicago.

J. Ross. 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, MIT.

G. Sampson. 1995. *English for the Computer: The Susanne Corpus and Analytic Scheme*. Clarendon Press, Oxford.