Specifying deeper semantics of a text using MSFA

Yoshikata SHIBUYA†, Jae-Ho LEE†, Kow KURODA†, and Hitoshi ISAHARA†

† National Institute of Information and Communications Technology (NICT) Knowledge Creation Research Center
3-5 Hikaridai, Seika-cho, Souraku-gun, Kyoto, 619–0289 Japan
E-mail: †{yshibuya,jhlee,kuroda,isahara}@nict.go.jp

Abstract In this article, we introduce the JCASR Project (in progress at NICT), which aims to develop a relatively small Japanese corpus of texts annotated for “semantic frames” and their “frame elements” (aka “semantic roles”), using Multilayered/Multidimensional Semantic Frame Analysis (MSFA) [3]. MSFA is a framework of semantic annotation/analysis compatible with the Berkeley FrameNet project [1], [6] that provides a multidimensional description of “contextualized” meanings of words and phrases. After outlining the project, we provide case studies of two sentences (one from a newspaper article and the other of a discourse).

Key words contextualized meanings of words and phrases, FrameNet, the JCASR project, Multilayered/dimensional Semantic Frame Analysis (MSFA), semantic role tagging

1. Introduction

The past 20 years have seen the maturation of surface-true, distributionally and statistically based knowledge acquisition techniques. The development has offered us a parse with a good coverage, free from inconsistencies immanent in manual analysis. It was like a new age of NLP. It turns out, however, that such methods clearly have limitations. One of the critical problems is that these techniques do not always meet our need for processing “deeper” semantics (or “shallow pragmatics”) in and of the language people use everyday. What we mean by “deeper semantics” is a class of semantic specifications that goes beyond simple “word senses/meanings” including the so-called “inferences” (yet applying this term would be no help here, because nobody knows exactly what “inferences” are, after all). This is the classical problem of Knowledge Representation (KR). Suppose that the KR problem is resolved totally after we have finished building all the relevant ontologies and inference engines running over them. Did we get to our ultimate goal at last? We suspect the answer is no, because we still need to find out how pieces of knowledge are linked to pieces of language. This is what Kuroda and Isahara (2005) called the problem of Knowledge-Language (K-L) Linking.

If our suspicion is true, we then need to construct a wide-coverage database that specifies what kind of linguistic units are linked to what pieces of world knowledge, or “ontologies” without (too simplistically) assuming that words denote “concepts” (i.e., building blocks of ontologies). This assumption is far from true, however: virtually every word of a given language is ambiguous, and it is far from well-understood how ambiguity comes into play, although there are good theories of word sense disambiguation/creation like Generative Lexicon (Pustejovsky [7]). This reality makes it very ineffective to state that words “denote” concepts. In fact, this situation demands us to build lexical resources dedicated to the identification and specification of “units” of the K-L Linking.

1.1 The JCASR project

Development of a Japanese Corpus Annotated for Semantic Roles (JCASR) is being attempted as one of the research projects at the National Institute of Information and Communications Technology (NICT), Japan. The project is proceeded with a crucial assumption that “units” of the K-L linking are “idealized/(stereo)typical situations” which can be identified as (semantic) frames in the same sense as the Berkeley FrameNet (BFN) project (Baker et al. [1]; Fillmore et al. [2]). The goal of the JCASR project is to construct a (relatively small) corpus of Japanese texts annotated for (seman-
tic) frames and their frame elements (aka “(situationally-defined) semantic roles”). The goal is to establish a set of (ontological) links from “pieces of world knowledge” to text segments in terms of semantic role tagging.

1.2 What we mean by “semantic role tagging”

In our approach, a strong emphasis is given to the identification and specification of finer-grained, situation-specific roles at concrete levels (e.g., ⟨Robbers⟩, ⟨Victim⟩, ⟨Valuables⟩; hereafter, ⟨⟩ represents a role) rather than those of coarse-grained, general-purpose roles at abstract levels (e.g., ⟨Agent⟩, ⟨Patient⟩, ⟨Theme⟩). Section 3 is devoted to illustration of this approach. Our approach is theoretically motivated by the hypothesis we assume that deeper, and “better,” understandings are achieved at more concrete levels, rather than at more abstract levels. This hypothesis is one of the points that differentiate MSFA from other (usually more “formally oriented”) approaches to semantic annotation/analysis which tend to assume that the deepest semantic analysis is the most abstract semantic analysis. More formally, we posit the following:

Concreteness bias on semantic interpretation: the more specific and concrete your understanding is, the better it is (as long as it is not obviously wrong).

A lot of phenomena suggest that “deep enough” semantic analysis of a text demands effective specifications of what questions people make, as well as of semantic types of text segments. What is suggested is that it does not really matter whether people’s understandings are semantically based or pragmatically based as far as our goal is to illustrate people’s text understanding: to specify what people understand is the point at issue, while how they do so is not. The semantics/pragmatics distinction would make sense if the issue is how people understand (after we have clarified what they understand). This reasoning would be both good news and bad news, depending on your perspective. It is good news if you feel that routes to deeper semantics are promised. It can also be considered bad news if you feel that you cannot be so optimistic as to say “Leave it all to pragmatics” anymore, because what is at issue now is what pragmatics does and how it works: you need to specify it.

2. Development Scheme

2.1 Status of the project

The JCASR project officially began two years ago. It is (still) at a preliminary, “exploratory” stage. At the moment, we are trying to find out what kinds of frames/situations are found at what granularity levels without assuming a pre-existing, “ready-to-use” database of semantic frames and their frame elements. Some preliminary results have been reported in Kuroda et al. (2006), for instance. We have not started serious development of a semantically tagged corpus yet, but annotation samples are available freely or privately at the web sites (contact us for more details). It should probably be noted that we are currently working independently of the Japanese FrameNet (JFN) project (Obara et al. 2003). But we are also negotiating with the BFN staff to make the MSFA-based annotation scheme shared with BFN.

2.2 Overview

Currently, we are following the “incremental” development scheme as per the following: (1) select a Japanese text T from a text database; (2) have each sentence of T segmented into text segments by the staff at NICT (each result of segmentation always needs to be checked manually, because the standard outputs of the so-called “morphological analyzers” like KNP and ChaSen are sometimes inappropriate for our purposes; this issue to be mentioned again later below); (3) ask “external” annotators to annotate the segmented texts by making reference to databases D1 and D2 of “sample annotations” hosted at the web sites (available both publicly and privately); (4) collect the annotations conducted by annotators as “drafts,” and check and edit the results if necessary, which is very often the case (this is conducted by the staff of the project group at NICT); (5) add the edited results to the databases D1 and D2, and “sanitize” the databases when needed.

T is always chosen from Japanese texts which are aligned with English texts. This is because we expect that future comparisons against other annotations (using the BFN database, for example) can be facilitated. So far, all texts have been taken from the following text bases:

D1: Sample annotations for texts from a collection of English-Japanese alignments of copyright-free texts like Aesop’s Fables are hosted at [http://www.kotonoba.net/~mutiyama/cgi-bin/hiki2/hiki.cgi?FrontPage]. No access restriction applies.

D2: Samples for texts from Kyoto University Corpus are hosted at [http://www.kotonoba.net/~mutiyama/cgi-bin/hiki2/hiki.cgi?FrontPage]. Access restriction applies.

The original texts for D1 are provided at [http://www2.nict.go.jp/x/x161/members/~mutiyama/align/index.html]. Tentatively, we separate the procedures to identify (a) frames for event conceptualizations (e.g., ⟨Robbery⟩, ⟨Predation⟩) and (b) frames for social interactions (e.g., speech acts like ⟨Claiming⟩, ⟨Criticizing⟩, ⟨Doubting⟩, ⟨Protesting⟩, ⟨Warning⟩). This is because the second type of frames are more complex, more data selective, and harder to specify. Currently, only Hajime Nozawa is working on the second type (see §3.3). Nozawa’s work has not yet been integrated into the results of the first type worked on by Kuroda, Lee and Shibuya.

3. Case Studies

3.1 The procedure

In conducting an MSFA, one employs a table of T m + 3 rows and n columns: m is the number of text segments (including “null instantiations”), and n is the number of frames identified as comprising the “understood content” of a sentence s. The cell at (i, j) of T specifies the semantic role r of the jth frame f_j. The value for f_j.r includes “null,” which means “non-realized role for f_j.” As illustrated in the following sections, the MSFA’s analytical scheme includes three header lines: “Frame ID” (row 1), “Frame-to-Frame relations” (row 2), and “Frame Name” (row 3). After the completion of the segmentation task, one turns to fill in the cells rightward, specifying (or identifying) Frame Names (together with the names of Frame Elements) and Frame-to-Frame relations among these frames. Frame IDs are local variables used to specify Frame-to-Frame Relations, whereas Frame Names and Frame-to-Frame Relations are global variables. Each sentence of a text T is segmented into text segments before one starts annotating them. The sentence segmentation process is conducted by using morphological analyzers such as ChaSen.
<table>
<thead>
<tr>
<th>Frame ID</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>E5</th>
<th>E6</th>
<th>E7</th>
<th>E8</th>
<th>E9</th>
<th>E10</th>
<th>E11</th>
<th>E12</th>
<th>E13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Koudo/Last year</td>
<td>Taiji</td>
<td>Taisai/Parabatai1</td>
<td>Taisai/Parabatai2</td>
<td>Taisai/Parabatai3</td>
<td>Taisai/Parabatai4</td>
<td>Taisai/Parabatai5</td>
<td>Taisai/Parabatai6</td>
<td>Taisai/Parabatai7</td>
<td>Taisai/Parabatai8</td>
<td>Taisai/Parabatai9</td>
<td>Taisai/Parabatai10</td>
<td>Taisai/Parabatai11</td>
</tr>
<tr>
<td>Location</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
<td>Shizuoka</td>
</tr>
</tbody>
</table>

図1 An MSFA of (1)

図2 An SFNA of (1)
Consider the sentence in (1) taken from Kyoto University Corpus (Kurohashi and Nagao [5]). (2) gives the English translation.

Table 1 provides an MSFA of (1):

(1) XXXXXXXXXXXXXXXXXX [S-ID:950101075-033]

(2) Last year, Habu challenged Yonenaga for the title.

In Table 1, the annotations have been converted from Japanese to English (the frame names are given tentatively in English; for the original MSFA in Japanese, see http://www.kotonoba.net/"mutiyama/cgi-bin/hiki2/hiki.cgi?c=view&p=msfa-round2-s33). All but the first column present the frames identified, or "discovered" in the analysis. As shown in the table, 14 frames were identified in (1) (but note that the number is subject to further modifications; see the later discussion). Each of the frames consists of some semantic roles. The Match frame (=XX) in the original Japanese analysis; the URL given above, for example, includes the roles such as (Player) (=XXXX), (Place for match) (=XX[XXXX]), (Time of match) (=XX[XXXX]). The colored cells indicate the semantic roles that are considered to be realizing their values in the relevant frames. The uncolored (or "empty") ones, on the other hand, represent those that do not seem to have any specific roles in the frames.

In conducting an MSFA, it is the semantic roles not the semantic types that are annotated in analyzing a text. Semantic types are roughly equated with natural kinds. In contrast, semantic roles are "situation-specific concepts" which are considered highly culture-particular and hence are taken to play a more crucially important role than semantic types in one's understanding of a text.

Having provided a brief illustration of how one analyzes a newspaper article sentence of (1) with MSFA, let us now turn to the prospects of this framework for cross-linguistic semantic annotation research. Consider (4), which is the Korean translation of (1) (translation provided by Jae-Ho Lee):

(4) Jagnyeon-e yonenaga-e dojeonha n salam-i habu.

An MSFA of (4) (here omitted for lack of space) reveals the similarities and differences in Japanese and Korean in terms of the availability of the frames (in 1 and 4). Below is the list of the additional frames that were identified with the Korean sentence (4):

(5) (i) (Specification[of being Japanese] (~)=(-XX[XXXX]~)) (ii) (Specification[of Time] (~)=(-XX[XXXX]~)) (iii) (Specification[of difference in ability] (~)=(-XX[XXXX]~)) (iv) (Transmission[to person] (~)=(-XX[XXXX]~)) (v) (Modification (~)=(-XX~))
(vi) ⟨~Specification[of the fact that it is an interpersonal event~⟩)
(=~⟨XX[XXXX]~⟩)

The involvement of these Korean-specific frames is considered to be due to the syntactic characteristics of the Korean language. It is interesting to see that Korean and Japanese (two typologically close languages) differ in the availability of the frame types in understanding an equivalent sentence. The results of an MSFA of (4) suggest an interesting prospect for cross-linguistic (semantic annotation) research, because it is suggested that a careful semantic analysis using MSFA would make a contribution to clarifying the universals and particulars of languages with respect to how people understand a sentence.

3.3 Sample 2: A segment of a prose

Let us briefly discuss how one analyzes a sentence from a discourse with MSFA. Consider the sentence in (6), taken from Aesop’s Fables (The Ass and the Grasshopper):

(6) XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

(7) They replied, “The dew.” [http://www2.nict.go.jp/x/x161members/mutiyama/align/htmPages/aesop-0.htm]

An MSFA of (6) is given in Table 3 and an SFNA in Figure 2. Similar to Table 1 and Figure 2, all the annotations of Table 3 and Figure 4 have been converted from Japanese to English (for the original analyses, see http://www.kotonoba.net/~mutiyama/cgi-bin/hiki/hiki.cgi?c=view&p=msfa-aesop01-s03-s04). A total of 25 frames were identified with (6). A remarkable fact with (6) is the involvement of the frames such as (Narration) (=XX), ⟨Telling a story⟩ (=XX), ⟨Creative work⟩ (=XX), ⟨Allegory creation⟩ (=XXXXX), ⟨Imagination⟩ (=XX), which are identified because of the nature of the text (i.e., a discourse-style fable). Of the frames found in the MSFA of (6), the ⟨Having curiosity⟩ frame (=XXXXXXXX) is perhaps among those that significantly differentiate MSFA from other semantic annotation frameworks for its elaborate semantic specification. In MSFA, one recognizes the ⟨Having curiosity⟩ frame in a reply sentence like (6), because it is a sentence with which the grasshoppers (specified by they) answer the preceding question that was posed by the ass who was “curious” to find out what type of food the grasshoppers live on so that they could possess such beautiful voices (see the preceding context given below; also notice the description given in Table 2 that says “G6 motivates F9”):

(8) XXXXXXX XXXXXXX X X XXX XXX XXXXXXXX

(9) AN ASS having heard some Grasshoppers chirping, was highly enchanted; and, desiring to possess the same charms of melody, demanded what sort of food they lived on to give them such beautiful voices.

In this section, we have briefly introduced how one analyzes a discourse segment, showing that MSFA provides a framework for describing discourse understanding as well. There is, however, a caveat to note before closing this section. The caveat to bear in mind is that the current version of MSFA scheme has not yet been equipped with a descriptive tool to link a reply to the descriptive contents of the preceding sentences effectively (one possible scheme is being developed by H. Nozawa; see his paper in this volume).

4. Conclusion

In this article, we introduced the JCASR Project, which aims to develop a relatively small Japanese corpus of texts annotated for semantic frames and their frame elements (aka semantic roles) in the same sense as the Berkeley FrameNet. After outlining the project, including its status, methodological procedures, strategies, and so forth, we provided case studies of two sentences, suggesting that MSFA provides a framework for deep semantic description.

参考文献

図 3 An MSFA of (6)

図 4 An SFNA of (6)