SPECIFYING WHAT PEOPLE UNDERSTAND WITH MSFA

Kow KURODA
National Institute of Communications Technology (NICT), Japan
11/28/2005
Introducing MSFA, Multi-layered Semantic Frame Analysis (Kuroda and Isahara 2005)

(Briefly) comparing it with Berkeley FrameNet (BFN) (Fillmore, et al. 2003)

Presenting a sample MSFA of an English sentence

With ONE IMPORTANT CAVEAT:

So far, MSFA has been done for Japanese sentences: just a few sample analyses were attempted for English.

Note that this is kind of inevitable, because MSFA requires, by its very design, an annotator/analyst to specify a lot of knowledge hard to access for non-native speakers.

But we don’t have enough time to talk about FOCAL today.
OUTLINE OF TALK

- Presenting sample MSFAs
  - Explain how MSFA goes
  - Explain how MSFA is related to “ontologies”
- Giving some background
  - Especially why I deviated from Berkeley FrameNet (Fillmore et al. 2003)
- Summary
HOW MSFA GOES
—SAMPLE ANALYSIS—
OVERVIEW OF MSFA

- MSFA is a BFN-inspired framework for text analysis by linguists such that
  - it combines linguistic analysis with text annotation for “deeper” semantics
  - it makes linguistic analysis “database-ready”
- MSFA’s goal is NOT just a development of a language resource usable for NLP tasks only.
  - I’m rather a researcher in Cognitive Science, rather than being a linguist, or an NLP guy.
  - Rather, it aims at a versatile resource that enhances as many researches as possible in Cognitive Science/Psychology, as well as tasks in NLP.
**MSFA Procedure (Simplified)**

1. Segment a sentences $S$ into units $U_1, \ldots, U_n$.
   - Note incidentally that it’s better NOT to try to build up larger units from smaller units. This tends to lead annotators to a “false” analysis.
   - This is not independent from Step 2. So, you need to go cyclic.

2. For each $U_i$, find a set of frames $F_1, \ldots, F_m$ so that one of their “frame elements” is realized by $U_i$.
   - This is called “evocation” in the *Frame Semantics* literature.

3. Specify relationships among all the frames.
   1. Relevant relations are: “$F$ elaborates $G$” (deals with Inheritance), “$F$ constitutes $G$” (deals with part-of relations), “$F$ presupposes $G$” (deals with “logical implications”)

GUIDING PRINCIPLES

“Be meticulous”
- Every word (or morphome if morphological analysis is necessary) needs to realize at least one semantic role, i.e., “frame element” of a frame.
  - You are not allowed to ignore a minor element by saying “its meaning is uninteresting.” If this “excuse” is allowed, your analysis will get arbitrary very soon.

“Be greedy”
- To every word, you need to assign as many semantic roles as possible if they are not incompatible.
  - It is an open question how many frames you need specify: there is no a priori way to tell when an MSFA is “done.”
An English translation of a Japanese Newspaper article taken from Kyodai Corpus (Kurohashi and Nagao 1994):

2. The book will definitely be a much-talked-about, severely criticizing the past U.S. Presidents and their aides.
3. The title came as latest work of Ronald Kesler, an expert reporter and investigator at the “Washington Post” and other media.
4. The book, for instance, reveals the following episodes.
5. ...
The following is the original Japanese text:

1. 「ホワイトハウスの内側」という本が十四日，米国で発売される。
2. 歴代大統領と関係者をこきおろしており，話題になるのは間違いいない。
3. 「ワシントン・ポスト」紙などで長年，調査報道をしてきたロナルド・ケストラー氏の新著。
4. 例えば次のような内容だ。
5. ...
<table>
<thead>
<tr>
<th>Frame ID</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
<th>F8</th>
<th>F9</th>
<th>F10</th>
<th>F11</th>
<th>F12</th>
<th>F13</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>F-to-F relations</td>
<td>elaborates F2; constitutes F3</td>
<td>constitutes F5; presumes F5; elaborates F4</td>
<td>presupposes F6; elaborates F4</td>
<td>presupposes F6; elaborates F5; presumes F7</td>
<td>presupposes F7</td>
<td>presupposes F7</td>
<td>presupposes F8</td>
<td>presupposes F9</td>
<td>presupposes F9</td>
<td>presupposes F9</td>
<td>presupposes F9</td>
<td>constitutes F3,F5</td>
</tr>
<tr>
<td>3</td>
<td>Frame</td>
<td>Title Giving</td>
<td>Name Giving</td>
<td>Writing</td>
<td>Authoring</td>
<td>Publishing</td>
<td>Selling</td>
<td>Purchasing</td>
<td>Consuming</td>
<td>Reading</td>
<td>Having Fun</td>
<td>Presidential Government in the U.S.</td>
<td>Disclosure</td>
</tr>
<tr>
<td>4</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>*</td>
<td>Purpose</td>
<td>GOVERNOR</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>*</td>
<td>Purpose2</td>
<td>GOVERNOR[+ composite]</td>
<td>GOVERNOR[+ composite]</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>*</td>
<td>Purpose2</td>
<td>GOVERNOR[+ composite]</td>
<td>GOVERNOR[+ composite]</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>*</td>
<td>Purpose2</td>
<td>GOVERNOR[+ composite]</td>
<td>GOVERNOR[+ composite]</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>*</td>
<td>Purpose2</td>
<td>GOVERNOR[+ composite]</td>
<td>GOVERNOR[+ composite]</td>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Instantiation Network of Semantic Frames, Specifying "Ontological Hierarchies"
The hierarchy of frames, especially the hierarchy of frame elements, expresses conceptual hierarchies you usually find in thesauri, e.g., *WordNet* synset hierarchies.

Why?

A possible — and very reasonable — answer is

- Instantiation links express “ontological hierarchies,”
- Part — and a probably substantial body — of human conceptual system is an organization of semantic “roles” rather than one of semantic “types”
What MSFA is meant to do is to list up all the relevant situations in text understanding in terms of frames, assuming that:

- Frames are organizations of frame elements, i.e., situation-specific “semantic roles”
  - *Author*, as a concept, names an Agent-class semantic role specific to the “Authoring” situation.
  - *Writer*, as a concept, names an Agent-class semantic role specific to the “Writing” situation, a subclass of “Authoring.”
- Frames are organized in principled ways.
  - So-called “thematic roles”, or “deep cases” are most abstract semantic roles.
WSD NEEDS TO BE FRAME-WISE

“Entities” in the understood content of a text may—and tend to—realize multiple roles/frame elements simultaneously.

- For example, *book* realizes such roles as:
  - *<Information Carrier>* in *<Reading>* frame
  - *<Good>* in *<Selling/Buying>* frame
  - *<Piece of Work>* in *<Writing>* frame
  - *<Publication>* in *<Publishing>* frame

This means that *Word Sense Disambiguation (WSD)* needs to be done frame-wise, explaining why WSD isn’t enough for text understanding, at least for simplex one.
Current Status

- MSFA was done to a tiny portion of Kyodai Corpus texts (3 articles, 63 sentences)
  - Kyodai Corpus is a collection of Japanese newspaper articles: its English translation is complete at NICT.
- Characteristics
  - No full evaluation yet
    - We need feedback from limited users, but publication is not unrestricted.
  - But, on average, a sentence has nearly 60 frames, showing that MSFA provides much deeper, ontology-based semantics than BFN.
**Really Need a Frame Database?**

- Unlike BFN, frames are identified and defined in an *ad hoc* manner, which is a method based on a deliberate decision.
  - MSFA does NOT make wide-coverage a priority.
  - Basically, the way MSFA works is *exploratory*, and it MAY not assume a pre-existing database of frames.
    - So, we *may* be faced with the “standardization” issue.
- Why? — Nobody knows the *optimal* granularity in semantic description even in terms of frames.
  - This means that a large-scale development of a frame database *can* be premature (but who knows?)
In principle, frames used in MSFA are defined independently of BFN frames.

- We DO NOT assume that BFN frames for (U.S.) English are applicable to Japanese without modification.

- Kanamaru, et al. (2005) examined the correspondence between the MSFA and BFN frames, showing that BFN frames are coarse-grained than MSFA frames.

- To get a more precise assessment for compatibility, we expect much to text annotation in Japanese FrameNet (Ohara, et al. 2003, 2004), but nothing has come out (yet).

- It’s vital to know how it will look like when BFN frames are applied to the analysis of Japanese texts.
WHY MSFA, NOT BFN?
——A BACKGROUND——
Beyond WSD

Text understanding is NOT simply a task of *Word Sense Disambiguation* (WSD). Clearly, a lot more is needed.

(Too) many researchers in NLP, and even in Linguistics and Psychology, believe that semantic analysis reduces to the WSD problem.

The real question is,

- What is WSD needed for?
- *Exactly what else* is needed in addition to WSD?

To this question, *Frame Semantics* (Fillmore 1985; Fillmore and Atkins 1994) comes to rescue.
MSFA is a derivative of Frame Semantics (FS), addressing the following two questions:

For a given sentence $S$,

A. How to specify what people understand when they hear or read $S$? — Call this the “Specification” Problem

B. How to represent what people understand when they hear or read $S$? — Call this the “Representation” Problem

MSFA is NOT concerned with the “truth” of $S$.

As FS says, knowing “what to do with $S$” is crucial. Knowing “when an $S$ is true” is subsidiary.
The “Representation” Problem makes sense only when the “Specification” Problem is properly treated.

But, the question is, *Is the “Specification” Problem properly treated?*

The answer is, *No,* obviously.

But why? — Linguists, at least in the Post-Chomskian linguistics, are in a “vicious circle.”
**Before you try to explain anything ...**

- **Why?**
  - Linguists have always tried to “explain” why people interpret such and such things, in such and such ways, without meeting the “Specification” Problem.
  - So, Linguistics is too immature a science even now: virtually any explanation in linguistics is arbitrary.

- **So what?**
  - We need to specify what people understands in sentences before explaining why people do so.
  - Linguists, too, need to be checked if their “interpretations” are the same as the real hearer/reader’s performances in some way.
  - FOCAL provides such opportunities.
Important fact:

- There is no guarantee that frames provided by BFN have an optimal semantic granularity.

This means that you need to check the *psychological reality* of descriptive devices, i.e., frames, used to specify the meaning of sentences.

- You can’t trust on linguists too much, as you already know.
- If you are too candid to believe BFN frames as such, your analysis will soon get arbitrary.
Test Case: “Attack” Frame

- (Some of) BFN frames can’t account for some cases of selectional restrictions: For example, <Attack> frame with core FEs <Assailant> and <Victim> can’t fully explain the following patterns:

  1. The lion attacked {a. the flock of impalas; b. ???the bank branch; c. ??innocent people on street}
  2. The robbers attacked {a. ???the flock of impalas; b. the bank branch; c. ?innocent people on street}.
  3. The random killer attacked {a. ???the flock of impalas; b. ??the bank branch; c. innocent people on street}.

- More granularity, which differentiates the <Purpose> of an <Assailant>, is clearly needed to account for this sort of selectional restrictions.
The optimality of semantic analysis/annotation in terms of granularity is task-dependent.

- There is NO optimal level for semantic analysis without specifying what you want to do with it.

The best way is

- NOT to disguise yourself as defining semantic frames at the optimal level of granularity.
- to assign a granularity index to each frame, ranging from a shallow to a very deep level one.
Given a frame for a verb XXY (e.g., X attack Y), you have a set of semantic co-variations between X and Y in terms of finer-grained semantic types.

- Selectional restrictions clearly correlate with units of such co-variations. For example, a <Predator> only attacks a <Prey> living in the same environment. This explains why the following contrasts:
  - The {a. tuna; b. ???wolf} attacked the sardins.
  - The {a. ???tuna; b. wolf} attacked the sheep.

- Usually, BFN frames have a number of subclasses, which serves as “units” of selectional restrictions.
For the case of “X-ga Y-wo osou” (“X attacks Y”, “X hits Y” in English), 15 different situations F01, F02, ..., F15, were identified by FOCAL and were shown to make sense to non-linguists through experiments.
SUMMARY

- MSFA tries to overcome some weaknesses of BFN by providing much finer-grained semantic analysis than BFN, to fully account for most cases of selectional restrictions.
- MSFA is not as useful as BFN for NLP: it doesn’t try to provide a wide-coverage database of frames.
- My tentative evaluation:
  - MSFA would be more preferable for researches in Cognitive Science/Psychology than linguistic resource developments in NLP.
  - But NLP will require semantic descriptions at this level of finer-granularity sooner or later.
My research is indebted to contribution by the following people:

Toshi-yuki KANAMRU, Kyoto University

Keiko NAKAMOTO, Kyoto University

Jea-Ho LEE, NICT

Hajime NOZAWA, NICT

Masao UTIYAMA, NICT