

Proposing the *Multilayered Semantic Frame Analysis* of Text

As an Effective Framework to Reveal What You Need to Know
Before Defining Entries for a (Generative) Lexicon

Kow KURODA

kuroda@nict.go.jp

Hitoshi ISAHARA

isahara@nict.go.jp

National Institute of Information and Communications Technology (NiCT), Japan

Abstract

This paper introduces a framework for both semantic analysis and annotation, called *Multilayered Semantic Frame Analysis (MSFA)* of text, inspired by the Berkeley FrameNet approach to semantic analysis of natural language text [8, 13]. MSFA is a work in progress, yet to be completed.

MSFA is so called because it describes the semantic specification of a sentence as an “integration of multiple semantic frames,” with each being represented as a distinct “layer.” MSFA defines a “high-precision,” “database-ready” encoding scheme for semantic entities that appear in a real text. It is useful to reveal how words and morphemes are linked to encyclopedic knowledge. This way, MSFA will help discover what knowledge is needed to enrich the “qualia structure” [16] for a given lexical item in a systematic way.

MSFA, if correct, implies theoretically that word sense disambiguation needs to be done multidimensionally, in such a way that each sense is recognized relative to a semantic frame comprising the semantics of a given sentence s , rather than to a sense of the predicates in s . This is an implication, we suggest, that can affect the definition for the word sense disambiguation task.

1 Introduction

1.1 What is MSFA? and Why?

It is generally agreed that lexical semantic analysis constitutes a “bottleneck” of effective Natural Language Processing (NLP). This is in part why recent NLP community is eager to build high-quality language resources annotated for semantic information, and there is always a need for a better framework for insightful and coherent lexical semantic analysis.

Even supported with a good theory of lexical semantics like Generative Lexicon [16], developing linguistic resources is not an easy task if it is not guided by a coherent framework for semantic

analysis. While we have several such frameworks recently, one of the most promising approaches is Berkeley FrameNet (BFN) [5, 8, 13], along with PropBank [9]. Even an attempt to automate semantic role tagging was pioneered by [7]. This is followed by the Session 5 of SENSEVAL-3 [12], a competitive workshop hosting for FrameNet-based sense disambiguation systems.

But the BFN framework, however, turned out to be not really satisfactory for our purpose of developing a semantically tagged corpus for Japanese. We needed to extend the BFN in the way specified in what follows. This is why we developed a framework called MSFA to be presented in this paper.

One thing needs to be noted explicitly at the beginning: we are **not** proposing an alternative to pre-existing processing models for lexicon building: we are just proposing a “preprocessing task” that supplements many of them, and a framework useful to achieve it.

1.2 Assumptions that Guide MSFA

MSFA assumes the following:

- (1) Human understanding in general is situation-driven,¹⁾ and so is linguistic understanding. To be more specific, situations serves as “units of knowledge organization,” at least in the sense that they are best characterized as internal “cognitive models.”
- (2) Such cognitive models can be specified in terms of semantic frames [4, 5] in the sense

¹⁾What we intend by this statement is **not** that the basis of semantics is situation-based in the sense of *Situation Semantics* [1]. While the notion of “situations” we assume in this paper is not explicit enough, it is sure that our perspective is broader than that of Situation Semantics/Theory.

that each situation is an organization of semantic roles²).

- (3) More explicitly, “(parameterized) states of affairs” are recognizable as “situations” (or whatever) by (more or less) humans because they have apt mental structures, finite in number, that recognize them: such mental structures/models are called (semantic) frames³.
- (4) While frames, specifying what situations the interpretation $I(s)$ of a given sentence s is liked to be, give a very rich and detailed semantic description to $I(s)$, frames can be successfully specified with minimum reference to syntactic structure of s .

Let us explain each of them in turn.

1.2.1 Understanding is situationally driven (Assumption 1)

The first assumption can be paraphrased into this: “Situations are units of human general understanding.” More specifically, this hypothesis says:

- (5) There exist certain “units” in human understanding in general. Linguistic understanding is just a special case of such general understanding. So, it is situation-based, too.
- (6) “Situations,” at least idealized ones, are one distinguished class of units of general understanding, and they stay so in linguistic understanding.

1.2.2 A set of semantic roles defines a situation (Assumption 2)

The second assumption can be paraphrased into this: “A situation is an organization of semantic roles.” More specifically, it says:

- (7) An idealized situation is an organization, or “gestalt,” of **situational (semantic) roles**.
- (8) Fillmore’s semantic frames, or at least one important subclass of them, are an adequate device to describe the idealized situations in the way defined in (9) below, adopting the format developed by Berkeley FrameNet:

- (9) { [\langle EFFECTIVE \rangle]⁴: what], [\langle GOVERNOR \rangle]: do what], [\langle OBJECT \rangle]: to what], [\langle MANNER \rangle]: how], [\langle PURPOSE \rangle]: for what], [\langle LOCATION \rangle]: where], [\langle TIME \rangle]: when], ... }

Admittedly, (9) is a general scheme, or “template,” of a situation. It is sure that important details are missed, but some of them will be clarified in the following discussion. Specifically, a situation, in most cases, are made of a number of subevents, each of which can be described in terms of frame.

MSFA distinguished **semantic roles** from **semantic types**, which we find different in kinds.⁵ Roughly, semantic types specify “natural kinds,” whereas semantic roles specify elements of “cognitive models” that need not have objective realities. Thus, semantic roles are susceptible to cultural differences, whereas semantic types are not.

Adopting the Berkeley FrameNet terminology, we often use “frame elements” and “semantic roles” interchangeably. “Thematic roles” in the generative literature, are a very special case of semantic roles in this sense. So, please be careful about what semantic roles denote in this paper. What we call semantic roles are not abstract entities like { AGENTS, PATIENT, INSTRUMENT, ... }, but roles or rather “role names” like { ATTACKER, VICTIM, WEAPON, ... }, { ROBBER, BANK, WEAPON, ... } that are particular to a situation (e.g., of ATTACKING, BANK ROBBING).

Also, the problem of what “names” are most suitable for semantic roles at this generic level like (9) is an unimportant one, theoretically or practically. The most important level is the “level of situation” which shapes human understanding.

1.2.3 \langle BUYING \rangle situation (example)

One of such interesting situations is the following \langle BUYING \rangle ,⁶ which is now given a description in

⁴We cannot find a good name for this semantic role. \langle AGENT(IVE) \rangle is too strong. The role need not be animate. The sense of “agent” in *chemical agent* is preferable, but this is not a typical sense of the term, unfortunately. \langle AFFECTIVE \rangle is pretty good, but it has a somewhat misleading connotation related to \langle LOVE \rangle ... We chose \langle EFFECTIVE \rangle , admitting that it is somewhat unusual, but terminology is not crucial.

⁵This distinction may look unusual, and even arbitrary. The first author has written a detailed article on this subject, but it is in Japanese and not included in references.

⁶Berkeley FrameNet has \langle COMMERCE.BUY \rangle for this. The frame consists of the following frame elements: { \langle BUYER, GOODS, MANNER, MEANS, MONEY, PLACE, RATE, RECIP-

²The sense of “semantic roles” is different from that of theoretical linguistics literature. We equate semantic roles with “frame elements” in the FrameNet terminology [6].

³This is close to the “classical” definition of frames by [15]

terms of semantic frame in (11):

- (10) *John bought a reference book for \$200 bucks at a local bookstore nearby for the coming exam on chemistry the other day, without hesitation.*
- (11) { [\langle BUYER \rangle : “John”], [\langle GOVERNOR \rangle : “bought”], [\langle GOODS \rangle : “a reference book”], [\langle PRICE = MANNER(OBJECT) \rangle : “for \$200 bucks”], [\langle LOCATION & SELLER \rangle : “at a local bookstore nearby”], [\langle PURPOSE \rangle : “for the coming exam on chemistry”], [\langle TIME \rangle : “the other day”], [\langle MANNER(AGENT) \rangle : “without hesitation”], ... }

It is necessary to recognize that \langle MANNER \rangle has two distinct components: one of them, \langle MANNER(EFFECTIVE) \rangle specifies the way \langle EFFECTIVE \rangle is doing something, the other, \langle MANNER(OBJECT) \rangle , specifies the way \langle OBJECT \rangle is characterized in a given situation.

1.3 Additional Assumptions to Extend the BFN framework

The Berkeley FrameNet (BFN) framework is interesting, pioneering, and very suggestive, but we find it somewhat unsatisfactory, at least for the following reasons:

- (12) So far, BFN annotation for semantic roles isn't quite useful to reveal what is really understood when people understand a sentence (or an utterance, if you like), because the current annotation for semantic role tagging is highly “selective,” and it doesn't specify the method to give an “full” annotation to a given sentence.
- (13) So far, BFN avoids annotation or analysis of “troublesome” cases including metaphor. While it is a reasonable strategy for building a frame lexicon rapidly, it is fatal for a project that aims at providing a “comprehensive” semantic analysis of a given sentence, because

AGENT, SELLER, TIME, UNIT } . Internal hierarchical organization of frame elements isn't assumed (so far). For instance, \langle MONEY \rangle , \langle RATE \rangle and \langle UNIT \rangle clearly specify the \langle MANNER \rangle component.

Also, frame element identification in BFN suffers from an inconsistency: it's better not to treat \langle MONEY \rangle as a semantic role: it's just a typical **value** for \langle PRICE \rangle , which is clearly a semantic role. This motivates to the aforementioned distinction between the semantic types and semantic roles.

it is an impoverishment of the sense specification/disambiguation problem, and the complexity of the lexical sense disambiguation problem is somewhat trivialized.

- (14) So far, BFN doesn't (seem to) consider the possibility of “multiple semantic role assignment” to a word, with each semantic role defined relative to a distinct frame, or at least it is not implemented yet. There is no guarantee that a sentence, or even a predicate, has just one frame.

For whatever reason, BFN ignores the very rich and complex structuring of semantic representation in many “real” sentences. If it is not shown how multiple frames are “integrated” into the semantics of a sentence, its annotation is basically useless. Inheritance in the frame hierarchy is not the only possibility for a sentence to have multiple frames linked to it. Semantics of a given sentence is susceptible to the “blending” effect [3].

1.3.1 Frame evocation and integration

To make the semantic analysis more satisfactory and comprehensive, MSFA extends the BFN framework, and assumes the following, relating to the “principles” for how to link frames to language:

- (15) **Frame-evocation by a linguistic unit** (Definition):
A linguistic unit u “evokes” a situation σ if and only if u “realizes” or “instantiates” a semantic role r of σ , sometimes denoted by $\sigma.r$.⁷⁾

Remark: While frames and situations are different in kinds, we (loosely and inadequately) equate “frames” with idealized situations hereafter, for terminological compatibility with BFN.

- (16) **Frame-evocation in a sentence** (Definition):
For a given sentence $s = m_1 m_2 \dots m_n$,

- a. frame-evocation takes place for every possible segmentation of s , including discontinuous ones⁸⁾; thus, frame-evocation by morphemes $M(s) = \{ m_1, \dots, m_n \}$ is just a special case of it.

⁷⁾This effect of evocation is probably association-based, and has an important link to “pattern-completion” in Hopfield nets, we suppose.

⁸⁾One of anonymous reviewers pointed out that it is not clear if this much degree of freedom is not too much to be

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Frame ID	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13
2	F-to-F relations	elaborates F2; constitutes F3		constitutes F5; presumes F5; elaborates F4		presupposes F3	presupposes F4; constitutes F5; presumes F7	presupposes F6; elaborates F9		presupposes F5	presupposes F9	Presidential Government in the U.S.	constitutes F3,F5	
3	Frame	Title Giving	Name Giving	Writing	Authoring	Publishing	Selling	Purchasing	Consuming	Reading	Having Fun		Disclosure	Reporting
4	*													Reporter
5	*					Purpose		GOVERNOR	GOVERNOR	Means				Report[start 1,end]
6	*							Purpose	Means	GOVERNOR	Means			
7	*							Purpose	Purpose	GOVERNOR				
8	*					Retailer Customer	Seller Customer	Seller Customer	Provider3 Consumer	Reader	Enjoyer			
9	*													
10	*	Title Giver[secondary]	Name Giver[2]		Supporter	Publisher	Provider		Provider2					
11	*	Title Giver[primary]	Name Giver[1]	Writer	Author	Supporter2			Provider1					Revealer
12	*			Purpose1							Domain=Topic		GOVERNOR	
13	A	Work	Object	Book	Work[+Piece]	Publication	Goods	Goods	Commodity	Book	Fun Source			Report[start 2,end]
14	book													
15	titled	GOVERNOR	GOVERNOR	Book.attribute	Work.attribute	Publication.attribute	Goods.attributes	Goods.attributes	Commodity.attribute	Book.attribute	Fun Source.attribute			
16	*	MARKER[1,2]	MARKER[1,2]											
17	The	Title	Name											Secrets: EVOKER
18	Inside													
19	White											Presidential Office: EVOKER	Target	
20	House													
21		MARKER[2,2]	MARKER[2,2]											
22	will					EXTENDER2	EXTENDER2							
23	go					EXTENDER1	EXTENDER1							
24	on			Purpose2		GOVERNOR[+composite]	GOVERNOR[+composite]							Means
25	sale													
26	in					MARKER	MARKER							
27	the					Place	Place							
28	U.S.													
29	on					MARKER	MARKER							
30	January					Time: Date	Time: Date							
31	14													
32	.													

Figure 1: MSFA of (18)

- b. At any level, frame-evocation takes place for each segmentation.⁹⁾
- c. The frame-evocation by m_i is independent from the frame-evocation by m_j if $m_i \neq m_j$.
- d. The number of frames linked to s is not limited, as least theoretically, as far as they are consistent.

(17) **Criteria for convergence and optimization** (Definition):

- a. “Be parsimonious for cost (i.e., memory)” (Criterion 1): For a given sentence s , the fewer the total number of the frames evoked is, the cheaper its semantic specification is, and the better it is.
- b. “Be greedy for richness (i.e., expressiveness)” (Criterion 2): For each morpheme m_i in s , the more frames m_i

computationally tractable. It is a reasonable concern, but we are not really concerned with computational implementation for the moment, while we are pretty sure that a certain kind of PDP-style, “parallel, distributed” computation should implement the task —because we believe human brain is implementing it anyway—. While we do not have a concrete computational model yet, we are, in a sense, at a stage of trying to determine what properties need to be included in such modeling as specifications.

⁹⁾This would explain why idioms, jargons, collocations, and styles, all varieties of so-called “multi-word expressions,” exist in every natural language.

“participates” (by realizing their frame elements), the richer the semantic specification of s is, and the better it is.

Put together, these two contradicting criteria lead to the integration and optimization of frame-evocation in a given sentence.

1.3.2 Separating (methodologically) semantic descriptions from syntactic ones

This is a provocative assumption, but we decide, at least methodologically, not to rely on detailed syntactic analysis. Thus, tree parsing is not a prerequisite for semantic analysis. MSFA assumes very “shallow” syntactic description, which are not hierarchicalized themselves. Admittedly, this decision/specification is open to criticism.

Above the definitions so far, let us give a few examples of the proposed framework.

2 Sample Analyses

2.1 Data from Newspaper article

(18)–(22) are the English translation of the Japanese newspaper article, (23)–(27) that appeared in the Japanese newspaper corpus, called *Kyodai Corpus* [10].

For illustration, let us perform a MSFA to (18), English, and (18), Japanese: (18) is the English translation of (23).

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	F-ID	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13
2	F-to-F relations	elaborates F2; constitutes F3		constitutes F5; presumes F5; elaborates F4		presupposes F3	presupposes F5; constitutes F5; presumes F7	presupposes F6; constitutes F8; elaborates F9		presupposes F5	presupposes F9	Presidential Government in the U.S.	constitutes F3,F5	
3	Frame identifier	Title Giving	Name Giving	Writing	Authoring	Publishing	Selling	Purchasing	Consuming	Reading	Having Fun		Disclosure	Reporting
4	*													
5	*					Purpose		GOVERNOR	GOVERNOR	Means				Report[start1,end]
6	*							Purpose	Means	GOVERNOR	Means			
7	*								Purpose	Purpose	GOVERNOR			
8	*					Retailer	Seller	Provider	Provider[tern]	Reader	Enjoyer			
9	*					Customer	Customer	Purchaser	Customer					
10	*	Title Giver[secondary]	Name Giver[secondary]			Publisher			Provider[secondary]					
11	*	Title Giver[primary]	Name Giver[primary]	Writer	Author	Supporter?			Provider[primary]				Revealer	
12	*												GOVERNOR	Report[start2,end]
13	「	MARKER[1,2]	MARKER[1,2]	Book.attribute	Work.attribute	Publication.attribute	Goods.attribute	Goods.attribute	Commodity.attribute	Book.attribute	Fun.attribute			
14	ホワイトハウス	Title	Name									Presidential Office: EVOKER	Target	
15	の												MARKER	
16	内側												Secrets: EVOKER	
17	」	MARKER[2,2]	MARKER[2,2]											
18	と	EVOKER1	EVOKER1											
19	題	GOVERNOR	GOVERNOR		EVOKER1									
20	する	EXTENDER	EXTENDER		EVOKER2									
21	本	A Piece of Work	Object	Book (as a Piece of Work)	Work	Publication	Goods	Goods	Commodity	Book (as Information Carrier)	Fun Source			
22	が					MARKER	MARKER							
23	十四					Time: Date	Time: Date							
24	日													
25	,													
26	米国					Place	Place							
27	で					MARKER	MARKER							
28	発売					GOVERNOR	GOVERNOR						Means	
29	され					EXTENDER1	EXTENDER1							
30	る					EXTENDER2	EXTENDER2							
31	,													

Figure 2: MSFA of (23)

- (18) *A book* titled “The Inside White House” will go on sale in the U.S. on January 14.
- (19) *The book* will definitely be a much-talked-about, severely criticizing the past U.S. Presidents and their aides.
- (20) *The title* came as the latest work of Ronald Kesler, an expert reporter and investigator at the “Washington Post” and other media.
- (21) *The book*, for instance, reveals the following episodes. [skipped]
- (22) Americans are very curious about the Presidential couple’s response to *the book*.
- (23)–(27) are the original Japanese version:
- (23) 「ホワイトハウスの内側」と題する**本**が十四日、米国で発売される。
- (24) 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ない。
- (25) 「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・ケスラー氏の**新著**。
- (26) 例えば次のような**内容**だ。[skipped]
- (27) 夫妻の反応が見ものだ。

Both in English and Japanese, boldfaced elements identify morphemes related to the *book*-concept.

2.2 MSFAs of (18) and (23)

Figure 1 gives the MSFA for (18). Figure 2 gives the MSFA for (23), which is Japanese.

2.2.1 MSFA terminology and notation

In each figure, each column corresponds to a frame, and provided with (i) a frame index (i.e., F_i), (ii) specification for the F-to-F relation; and (ii) frame name/identifier (e.g., “Title Giving”).

F-to-F relation means the “frame-to-frame relation”. Currently, implicational relations such as “ F presupposes G ,” “ F constitutes G ,” “ F elaborates G ” are recognized, though they are not exhaustive. Some of those relations are borrowed from BFN.

“Governors,” or “frame-governors,” are the term borrowed from the BFN framework. They name frames, and are typically predicates like verbs and prepositions.

“Evokers” do not appear in the BFN framework. They explicitly indicate, when possible and adequate, nominal (and sometimes adjectival) frame-evoking elements that are not frame-governors.¹⁰⁾

“Markers” and “extenders” do not appear in the BFN, either. Unlike BFN, **MSFA treats prepositions as not parts of semantic roles**: prepositions are explicitly distinguished as markers. This is not an arbitrary decision, and has nontrivial consequences, but we will not look into them here.

Extenders are somewhat similar to the BFN notion of “supporting verbs,” but are regarded as a

¹⁰⁾Our conception of frames is more conservative than BFN’s. We are cautious not to recognize too many items as frame-governors. From our perspective, most adjectives are not governors but evokers strongly linked to certain frames.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Frame ID	F1	F1	F2	F3	F4	F6	F7	F8	F9	F10	F11	F12	F13
2	F-to-F relations		presumes F2	presupposes F1						constitutes F1; elaborates F10				
3	Frame identifier	Reading for Fun	Writing	Publishing	Predicting	Predicting*	Categorizing*	Chatting*	<Conjunction[clausal]>	Criticizing	Judgment_communication*	<Conjunction[nominal]>	<Anaphoric Relation>	Assisting
4	*				Predictor	Speaker*	Cognizer*							
5	*				GOVERNOR	GOVERNOR*								
6	*	Reader	Audience	Customer	Happening[1, 2]	Eventuality*		Interlocutors*	Event1[Specification of Result]					
7	*	Publisher	Supporter	Publisher				Topic3*						
8	*	Author	Writer	Supporter				Topic2*		Criticizer	Communication*			
9	The	BOOK: EVOKER	WORK: EVOKER	PUBLICATION: EVOKER			Item*	Topic1*						
10	book													
11	will				EXTENDER: EVOKER									
12	definitely				Degree of Certainty									
13	be	Effect: Reaction	Response	Aftermath	Happening[2, 2]			EVOKER*						
14	a							Category*						
15	much-							Degree*						
16	talked							GOVERNOR*						
17	-about							Topic1*						
18	,								GOVERNOR					
19	severely	Book.Content	Purpose[1,2]		Reason[1,2]	(Reason*)			Event2[Specification of Cause]	Degree				
20	criticizing									GOVERNOR EXTENDER	GOVERNOR*			
21	ing									Target	Evaluate*	Item1	Target=Antecedent	Principal
22	the		Purpose[2,2]		Reason[2,2]									
23	past													
24	U.S.													
25	Presidents													
26	and													
27	their													
28	aid													
29	-es													
30	.													

Figure 3: MSFA of (19)

special case of a more general class of “supporters,” which are a special case of markers. Extenders are elements that extend the function of “governors”, and sometimes behave like “deputies” of governors.

In English and Japanese, markers and extenders specify their arguments in opposite directions: markers and extenders are prepositional in English, whereas they are postpositional in Japanese.

* indicates a NULL instantiation of a semantic role. This does not mean there is a “trace” where * occurs. **No syntactic operation is assumed as to the occurrence of *.** It just means “such and such semantic role has no overt surface manifestation”; that’s all. Generally, you can put * wherever you want, and its position (usually) does not affect the analysis, at least MSFA is so designed.

Discontinuous units are easily handled with $F.R[i, n]$, which encodes the i th segment of the role R for F , with R having n segments in total.

Dubious role occurrences are indicated by bracketing their names. Morphological analysis of a word w is indicated by inserting “-” into w .

In those MSFA’s and others, **semantic role specification is usually partial.** Only overtly expressed roles or “salient” implicit roles are indicated in MSFA for their “informativeness.” It is an open question if semantic roles can be specified exhaustively. Assuming a version of Frame Semantics [4, 5], we believe it’s possible, at least at a certain level of abstraction. The hardest thing to do is

to tell where it is.

2.2.2 How frame-evocation converges

It should be noted that frame-evocations are best characterized as “pattern recognition” processes that run strictly in parallel, in the sense that all frames “recognize” their elements without knowing what other frames are doing. The condition for convergence is a “winner-and-his-friends-take-all” style competition among all the frames evoked. MSFA assumes that human linguistic understanding builds on a parallel, distributed computation.

2.3 MSFA of (19)

For comparison with the BFN analysis, we provide the MSFA of (19) in Figure 3, where BFN frame definitions (in dark green) are also included.

Clearly, BFN frames are not detailed enough to reveal the rich semantics of the sentence, but this is kind of unavoidable, considering that the rapid development of frame lexicon for English comes at the top of BFN’s research agenda.

While BFN aims at providing a “bridge” between the syntactic and semantic information, MSFA doesn’t. We do not assume that tree parses provide proper descriptions of syntactic structures of sentences. Dependency parsing would be much better than tree parsing, but substantial enrichment will be needed to make it meet the high demands of semantic description.

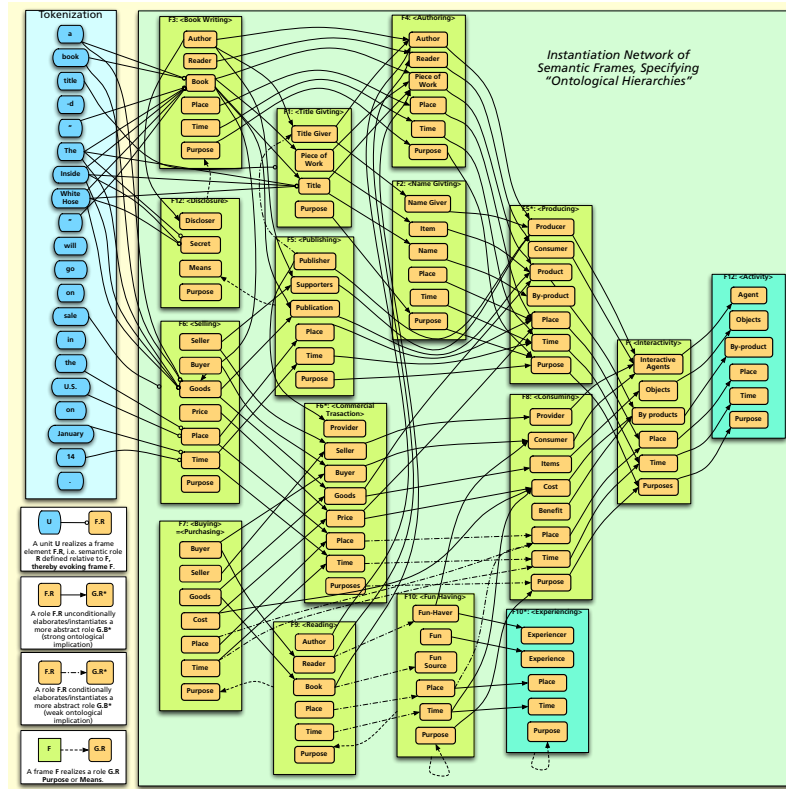


Figure 4: An “ontology-like” specification based on MSFA of (18)

2.4 Points that MSFA makes

As demonstrated in the sample analyses in Figures 1 and 2, MSFA makes the following points, related to the rich lexical semantic description of a given word or morpheme.

2.4.1 Ability to reveal links from language to knowledge

As demonstrated in Figures 1 and 2, MSFA is a powerful method to reveal the links from language to “world knowledge.” This is the very feature that is demanded by most NLP tasks.

Readers may wonder how MSFA performs this, partially doing a task of knowledge representation. Actually, an anonymous reviewer commented on our submitted paper as follows: “The paper claims that MSFA ‘reveals the link from language to world knowledge’, but unfortunately, it is not clear how this can be achieved by the method, unless one considers the frames as encyclopedic, a tall order.” Although enough space is not allowed to go into relevant details here, one of what we will try to do is to extend frames to be encyclopedic—at least as much encyclopedic as realistic as a re-

search in Cognitive Science. We are not so sure how it is realistic as an NLP task. We are aware that this is a quite controversial point. Let us mention a few points briefly.

Some of the links from lexical items to pieces of world knowledge are diagrammed in Figure 4, which shows, in an abbreviated fashion, how MSFA provides an “interface” to the ontological specifications of what is understood when (18) is read by an ideal “average” reader.

The diagram in Figure 4 is manually crafted, based on the information that MSFA of (18), in Figure 1, provides. No processing technique has been developed to automate this task, but we already have a visualization tool that converts an MSFA into a simplified diagram, which helps validate F-to-F relations. Some sample results can be seen at <http://61.115.230.87/~mutiyama/cgi-bin/hiki/hiki.cgi?FrontPage>.

On the right-hand side of the diagram in Figure 4, semantic frames and their frame elements, i.e., semantic roles, are networked in terms of class/instance hierarchy. So-called “type hierarchies” are partial descriptions of the network of semantic frames.

As the diagram reveals, some frames are **evoked lexically**, and linked to the tokenization of a sentence directly. All others frames are **evoked inferentially**, and linked to it indirectly. For (18), F1: ⟨TITLE GIVING⟩ is evoked by the sequence of words [*titled*, “, *The, Inside, the, White, House,* ”], F3: ⟨BOOK WRITING⟩ by [*a, book*], F6: ⟨SELLING⟩ by [*(go), on, sale*], and F12: ⟨DISCLOSURE⟩ by [*“, The, Inside, . . . , ”*]. Again, frame-evokers need not be continuous, though many of them are continuous. Recognizing this is important to allow for “distributed” evocation, which serves as a basis for multi-word expressions.

The distinction between lexical and inferential evocations is not clearly encoded in MSFA’s in Figures 1, 2 and 3, and this may invite confusions.

F9: ⟨READING⟩, for example, is not evoked lexically in (18). It is evoked, or rather “activated,” as a result of “spreading activation” over the network of semantic frames and semantic roles. There are two routes of such activation:

- (28) a. F6: ⟨SELLING⟩ ⇒ F6*:
 ⟨COMMERCIAL TRANSACTION⟩ ⇒
 ⟨BUYING⟩ ⇒ ⟨READING⟩
- b. F3: ⟨BOOK WRITING⟩ ⇒ F4:
 ⟨AUTHORING⟩ ⇒ F9: ⟨READING⟩

Some links are conditional. For example, the instantiation link from F9: ⟨READING⟩ to F10: ⟨FUN HAVING⟩ is conditional. Actually, all readings are not for fun having: consulting a reference book usually gives you no fun.

All frames are organized in a certain systematic way. Part of such organization is what we call “(lexical) knowledge.” Partial, and usually incomplete, description of it is so-called qualia structure, we suggest. One of such organizations is that, as the comparison of MSFA’s for (18) and (19) shows, certain frames —such as ⟨WRITING⟩, ⟨SELLING⟩, ⟨PURCHASING⟩, and probably ⟨PRINTING⟩ not included in the MSFA’s— “cluster” to constitute ⟨PUBLICATION⟩ as a (social) “(inter)activity”. Part of such information is encoded by the Frame-to-Frame relations at the second row of each MSFA.

A final note on the diagram in Figure 4: this is not intended as an exhaustive specification of world knowledge. Vast information, which provides symbol grounding, is missing. What we are trying to suggest is just this: **MSFA can be a useful tool to link natural language expressions to a fully specified (ontological) knowledge**

base without too much messing up entries of the lexicon. In this specific sense, we suggest that MSFA serves as a useful and powerful “preprocessing” before researchers in (computational) lexical semantics determine what properties need to be included into the definitions of lexical items — especially into their qualia structures. As far as we know, there seems to be no heuristics to find out the qualia structure of a *given* lexical item.

Thus, MSFA has a dual function. First, it helps to “detect,” for a given sentence, what lexical items serve as “entry points” into an ontological knowledge base. Second, it helps to “discover” what knowledge, in terms of semantic frames, are accessed to get a full interpretation of a given sentence. With this, it is expected that MSFA reduces the complexity of the lexicon building task.

It needs to be emphasized that MSFA doesn’t replace lexicon building task, whether it be a generative lexicon or not. It would be best understood as a powerful preprocessing technique to prepare a (generative) lexicon. It would be especially useful to determine what information is specified where.

2.4.2 Ability to integrated lexical semantic analysis and semantic annotation

Viewed as a preprocessing procedure, MSFA provides another important feature: lexical semantic analysis and semantic annotation are achieved at the same time: they are not separated. This makes MSFA of a given text “database-ready.”

2.4.3 Ability to provide cross-linguistically compatible description

While MSFA doesn’t assume “happy-go-lucky universalism” as to semantic entities, the comparison between (18) and (23) is fairly straightforward. Virtually, the same set of frames is used in this English/Japanese pair, though it is not always true.

2.4.4 Ability to encode many kinds of lexical semantic phenomena

MSFA provides an “automated detection,” if not “automatic discovery,” of a variety of metonymic effects. For example, **simultaneous type coercion effect**¹¹⁾ can be easily detected as to the interpretation of *the book* in (18). This phrase, in this specific

¹¹⁾It is somewhat unclear if this effect is independent of selective binding.

context, receives the following different semantic roles, some of them correspond to **agentive**, **telic** roles of the qualia structure of “book”:

- (29) *(the) book* in (18) realizes such roles as:
- a. ⟨PIECE OF WORK⟩ in ⟨TITLE GIVING⟩
 - b. ⟨BOOK⟩ (as a ⟨PIECE OF WORK⟩) in ⟨BOOK WRITING⟩
 - c. ⟨BOOK⟩ (as ⟨INFORMATION CARRIER⟩) in ⟨READING⟩
 - d. ⟨PUBLICATION⟩ in ⟨PUBLISHING⟩
 - e. ⟨GOODS⟩ in ⟨SELLING⟩ and ⟨BUYING⟩
 - f. ⟨FUN SOURCE⟩ in ⟨FUN HAVING⟩

While *title* selects (29a) and *(go) on sale* selects (29d) and (29e), all of these semantic roles are latent in the meaning of *book*, and always there in its lexical meaning. ⟨WORK⟩, ⟨PUBLICATION⟩ and ⟨GOODS⟩ constitute the **agentive** role of qualia structure, and ⟨BOOK (AS ⟨INFORMATION CARRIER⟩)⟩ and ⟨FUN SOURCE⟩ constitute the **telic** role. MSFA defines any of these situational roles relative to general notion of understandable “situations,” specified in terms of semantic frames, and tells when they occur and where in a text.

While it is not demonstrated in the sample analyses, MSFA treatment of metaphor is also straightforward. For a given metaphor, MSFA can specify the “source” and “target” meanings in the sense of Lakoff and Johnson [11], indicating the link, or “transfer” from the source to the target. Interpretation of “books” in phrases like *cook the books* requires this kind of treatment. MSFA doesn’t automate the analysis; yet it would reduce the complexity in the task effectively.

3 Concluding Remarks

If the framework of MSFA, proposed and outlined in this paper, is correct, it has certain implications of theoretical importance, one of which is this:

- (30) Word sense disambiguation needs to be done **multi-dimensionally** in such a way that each sense is recognized relative to a semantic frame that constitutes the semantics of a given sentence *s*, rather than to a sense of the predicates in *s*.

This is an implication, we suggest, that can change the “definition” for the word sense disambiguation.

3.1 How MSFA helps to deal with polysemy

We are making a strong claim— we are aware of it. Moreover, we didn’t provide enough evidence to validate it, unfortunately. This is why an anonymous reviewer rightly remarked: if MSFA “aims to help solve the problem of polysemy. In that case, at least a few other sentences containing the same words but with other senses (*a book of stamps, accounting books, a book as a chapter of a larger book*, phrases such as *cook the books, throw the book at someone*, etc.) should have been analyzed as well.” We would be happy if enough space and time were allowed to demonstrate missing details.

For space, 10 pages is just too short. For time, we are not really ready to present English MSFA’s in as much detail as we hope. So far, MSFA has been being developed and elaborated for Japanese text analysis. Its application to English is far from satisfactory for the moment, let alone complete.

Under this caveat, we would like to add some relevant details needed to disambiguate the meaning of *book* in the phrase *a book of stamps*. Its interpretation is done against the ⟨COLLECTING⟩ frame, which comprises such roles as: ⟨⟨COLLECTOR⟩, ⟨TARGET⟩, ⟨COLLECTION⟩, ⟨MEANS⟩, ⟨PURPOSE⟩, . . .⟩. ⟨COLLECTION⟩ has a physical entity, and need to be ⟨MAINTAIN⟩ed. This motivates a “unit” and a “mode” of a ⟨COLLECTION⟩’s existence, which is also useful to ⟨MEASURE⟩ it. This unit is —more or less accidentally— conventionally termed as a “book” for ⟨STAMP COLLECTION⟩. So, in sentences like:

- (31) *He’s collected stamps for many years to have thousands books of them, now occupying a room.*

book(s) (of stamps) appears to refer to the single entity, but its aspects (or “facets” in Cruse’s [2] term) selected by predicates “ x_1 collected y_1 ,” “ x_2 have y_2 ,” and “ x_3 occupy y_3 ” are different. For the first predicate, *book* denotes a value for y_1 , a conventional unit of stamp collection, with x being a value for ⟨COLLECTOR⟩; for the second, *book* denotes a value for y_2 , a unit of ⟨MAINTENANCE⟩ and ⟨POSSESSION⟩, with x_2 being a value for ⟨OWNER⟩ and ⟨POSSESSOR⟩; for the third, *book* specifies a value for x_3 , a unit of space-occupation, with y_3 being a ⟨SPACE⟩ to be occupied.

3.2 Shortcomings of MSFA?

3.2.1 MSFA is intuition-demanding

MSFA has at least one shortcoming: **it is intuition-demanding**. MSFA asks analysts for very sharp intuitions about their language that would be impossible for non-native speakers to have. But this is a common feature of what is called “common-sense,” and this only shows that expertise would be indispensable for its description, we presume.

3.2.2 MSFA blurs the syntax-semantics mapping?

Another potential shortcoming is this: In MSFA, **the linking mechanism from semantic roles to argument structure is more or less blurred**; at least there is no straightforward mapping from one to the other. This is, however, a natural outcome of our decision not to rely lexical semantic description too much heavily on syntactic description. By this, we’d rather suggest that the notion of argument structure needs to be redefined, reconsidering what a given argument is an argument of. MSFA suggests that many of the so-called arguments need not be arguments of the predicates definable relative to a syntactic structure.

But this may not, we beg, be interpreted that MSFA fails to provide the coherent interface between syntax and semantics: quite the contrary. MSFA *does* provide a coherent syntax-semantics interface, but only in a novel way that is just rarely heard of in the generative tradition. The design feature embodied in MSFA is the **Parallel Distributed Processing (PDP)** architecture [17, 14], which is widely accepted in Cognitive Science.

References

- [1] J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, 1983.
- [2] D. A. Cruse. Polysemy and related phenomena from a cognitive linguistic viewpoint. In P. Saint-Dizier and E. Viegas, editors, *Computational Lexical Semantics*, pages 33–49. Cambridge University Press, Cambridge, UK, 1995.
- [3] G. R. Fauconnier and M. Turner. Conceptual integration networks. *Cognitive Science*, 22:133–187, 1998.
- [4] C. J. Fillmore and B. T. S. Atkins. Starting where the dictionaries stop: The challenge for computational lexicography. In B. T. S. Atkins and A. Zampoli, editors, *Computational Approaches to the Lexicon*, pages 349–393. Clarendon Press, Oxford, UK, 1994.
- [5] C. J. Fillmore and C. F. Baker. Frame Semantics for text understanding. In *Proceedings of WordNet and Other Lexical Resources Workshop, NAACL, Pittsburgh*. 2001.
- [6] T. Fontenelle, editor. *International Journal of Lexicography — Special Issue: FrameNet and Frame Semantics*, volume 16 (5). Oxford University Press, 2003.
- [7] D. Gildea and D. Jurafsky. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288, 2002.
- [8] C. R. Johnson and C. J. Fillmore. The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure. In *Proceedings of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics (ANLP-NAACL 2000)*, pages 56–62, 2000.
- [9] P. Kingsbury and M. Palmer. From TreeBank to PropBank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002)*, 2002.
- [10] S. Kurohashi and M. Nagao. Building a Japanese parsed corpus: While improving the parsing system. In A. Abeillé, editor, *Treebanks: Building and Using Parsed Corpora*, pages 249–260. Kluwer Academic, Dordrecht/Boston/London, 2003.
- [11] G. Lakoff and M. Johnson. *The Philosophy in the Flesh*. Basic Books, 1999.
- [12] K. C. Litkowski. SENSEVAL-3 task: Automatic labeling of semantic roles. In *SENSEVAL-3: Third International Workshop on the Evaluation of Systems (ACL for the Semantic Analysis of Text, Barcelona, Spain, July 2004)*, pages 9–12, 2004.
- [13] J. B. Lowe, C. F. Baker, and C. J. Fillmore. A frame-semantic approach to semantic annotation. In *Proceedings of the SIGLEX Workshop on Tagging Text with Lexical Semantics: Why, What, and How?* 1997.
- [14] J. McClelland, D. Rumelhart, and The PDP Research Group. *Parallel Distributed Processing, Vol. 2*. MIT Press, 1986.
- [15] M. L. Minsky. A framework for representing knowledge. In P. H. Winston, editor, *The Psychology of Computer Vision*, pages 211–277. McGraw-Hill, 1975.
- [16] J. Pustejovsky. *The Generative Lexicon*. MIT Press, 1995.
- [17] D. Rumelhart, J. McClelland, and The PDP Research Group. *Parallel Distributed Processing, Vol. 1*. MIT Press, 1986.