Developing a Japanese Corpus Annotated for "Frames" and their "Elements" —Specifying What People Understand with MSFA—

Kow KURODA
National Institute of Communications Technology (NICT), Japan
09/06/2006

Today's Topic

- Introducing *Multi-layered/dimensional Semantic Frame Analysis* (MSFA: Kuroda and Isahara 2005)
  - It was developed as an annotation scheme hopefully compatible with *Berkeley FrameNet* (henceforth, BFN) (Baker, et al. 1998; Fillmore, et al. 2003; Johnson and Fillmore 2001; Lowe, et al. 1997)
- Caveats:
  - So far, MSFA has been done for Japanese: just a few sample analyses were attempted for English.
  - MSFA requires, *by its very design*, an annotator/analyst to specify *a lot of knowledge hard to access for non-native speakers*.

Omitted Topics

- MSFA is coupled with a theoretical framework called *Frame-Oriented Concept Analysis of Language* (FOCAL: Kuroda, et al. 2005; Nakamoto, et al. 2005).
  - Competitive Theory of Frame Selection (Kuroda et al. 2006, presented at DGfS) is a product of FOCAL
- But we don't have enough time to talk about FOCAL today.

Overview

- Giving some background
  - Especially why frame definitions and the annotation scheme of *Berkeley FrameNet* (BFN: Fillmore et al. 2003) were not used so far
  - Supplement it with a competitive theory of frame selection
- Presenting sample MSFAs
  - Explain how MSFA goes
  - Try to show what issues BFN will face when full-text analysis/annotation is seriously attempted.
- Summary

How to Annotate Japanese Texts for Semantic Roles with MSFA

Where Does MSFA Come from?

- When Kow Kuroda came to know about the BFN approach to semantic annotation (Johnson and Fillmore 2001; Pinkal et al. 2003) at ACL 2003, he found it really exciting, and wanted to try out the same thing for Japanese text analysis.
- But he faced some difficulties

Major Obstacles

- The following are major obstacles:
  1. At that time, the *coverage* of BFN database wasn't broad enough, and as much fine-grained as he needed.
  2. What's worse, if we decide to go with BFN frames,
     A. good understanding of English is required (for both the staff and annotators); this is too selective.
     B. There will be little chance to link annotation to entries in Japanese thesauri (e.g., *Nihongo Goi-taikei* (A (Comprehensive) Japanese Lexicon)
- While issue 1 is improved greatly in last two years, issue 2 is still a problem.

Decisions Made Two Years Ago

- To annotate Japanese texts for *deep enough* semantics is our objective. So, we decided
  - not to go along with BFN frames
  - to develop our own *scheme* for semantic annotation/ analysis such that
    - it is applicable to *given* Japanese sentences
    - it provides *deep enough* semantic analysis useful for research in *Cognitive Science of Meaning*
- The following are not our goals *per se*:
  - develop NLP applications like MT, QA system, IR
  - construct a frame database as an "extended lexicon"

Why Semantic Annotation?

- We need annotation/analysis of deep enough semantics to provide the "infrastructure" for *cognitive science of meaning*.
  - Our research is not seriously oriented for NLP tasks, even if there should be no incompatibilities.
- Rather, we addressed the following explorations:
  - Given a sentence *s*, *what kinds of frames* are needed if we wanted to achieve a "psychologically real" description *d* of what (average) people understand when *s* is heard or read?
  - and *how frames* are "put together" in description *d*?

Far from Trivial Matters

- How to deal with metaphor, metonymy, and other sorts of "figures of speech"
- How to break a sentence into "meaningful units"
  - How to deal with frame-evocation by complex, often discontinuous units?
- How to deal with anaphora?
- How to treat topic marker: is it part of a FE or not?

Short (and Hopefully Gentle) Introduction to FrameNet

What Is a (Semantic) Frame?

- A (semantic) "frame" is
  - an organization of "frame elements" (FEs), i.e., situation-specific "semantic roles" in human mind/brain
  - that represents a schematization of situation, or a generalization over "events" (or "states")
- Caveat:
  - Don't confuse semantic roles in this sense with so-called "thematic roles," or "deep cases" in the sense of Case Grammar (Fillmore 1968)

Examples

1. <Predation>=<<Predator>, <Prey>, ...>
   i. [Predator *A group of killer whales*]
   ii. [Predate.GOVENOR *attacked*]
   iii. [Prey *a humpback whale*].
2. <Bank Robbery> = <<Robber>, <Bank>, <Weapons>>
   i. [Robber *A group of masked men*]
   ii. [Rob.GOVERNOR *attacked*]
   iii. [Target *a bank branch in L.A.*].

Hierarchies of Frames and Frame Elements

- <Predation> IS-A <Harm-Causation> = <<Harm-cauaser>, <Victim>>
  - <Predator> IS-A <Harm-Causer>
  - <Prey> IS-A <Victim>
- <Bank Robbery> IS-A <Robbery> IS-A <Attacking> IS-A <Harm-Causation>
  - <Attack> = <Assailant>, <Victim>
- <Attacking> IS-A <Intentionally_act> (IS-A <Agentive_act>)
  - <Bank Robber> IS-A <Robber> IS-A <Agent>

Role-denoting Names

- Note
  - "robber," "victim," "predator," and "prey" are all *role-denoting/specifying nouns*.
- But
  - "(a group of) killer whales," "a humpback whale," "(a group of) masked men," and "a bank branch" are not; they are *entity-denoting/specifying nouns*.
- See Kuroda, Nakamoto and Isahara (2006), Gentner (2005), Gentner and Kurtz (2005) for relevant details.

But the Story Isn't Over Yet

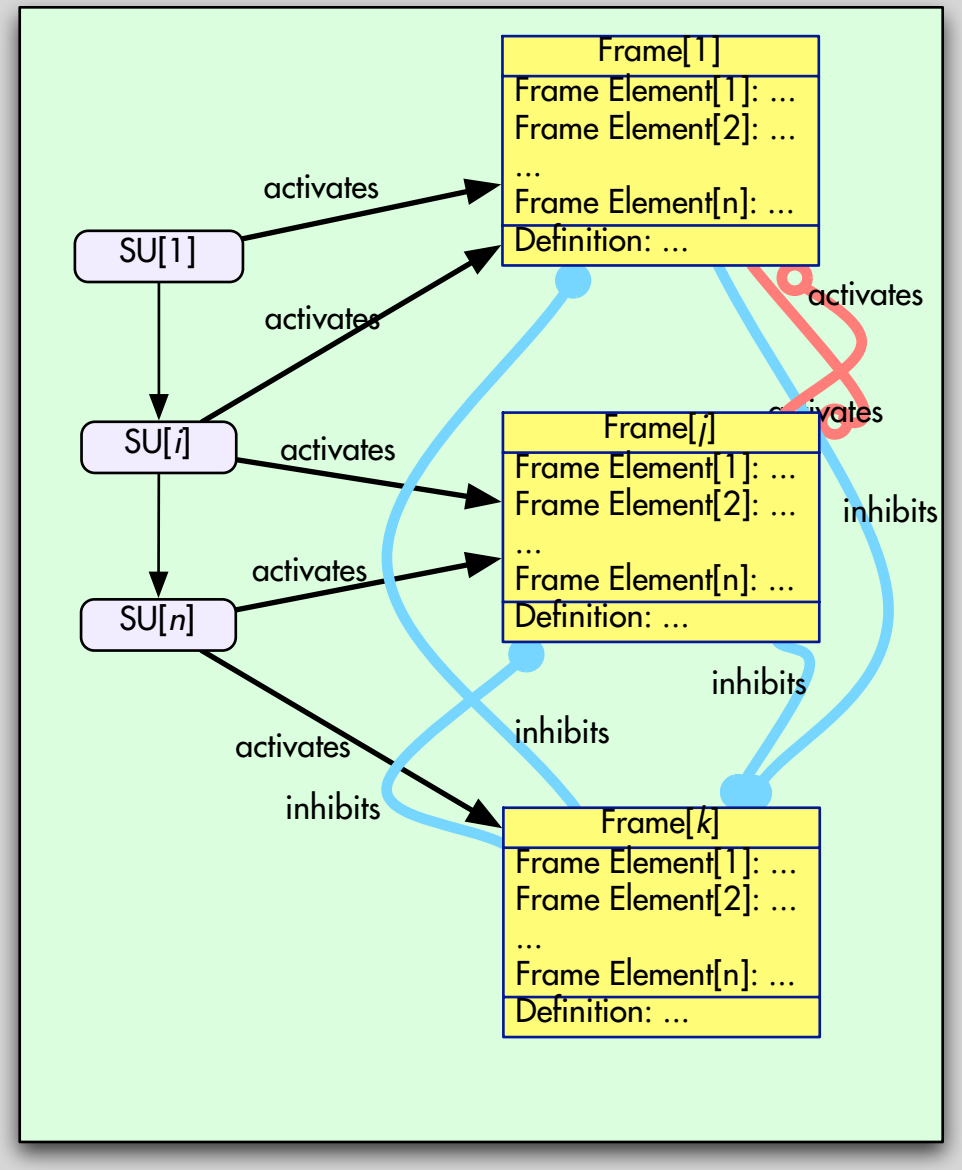
- Why is it that sentences 1 and 2 make sense, and 3 and 4 don't, unless they are "adjusted" (metaphorically or metonymically)?
  1. A group of *killer whales* attacked a humpback whale. [<Predation> situation]
  2. A group of *masked men* attacked a bank branch in L.A.. [<Bank Robbery> situation]
  3. A group of *killer whales* attacked a bank branch in L.A.. [<??> situation]
  4. A group of *masked men* attacked a humpback whale. [<??> situation]

Samples of "Adjustment" by Semantic Accommodation

- Sentence 3 can mean a <Bank Robbery> if "killer whales" are understood as nicknames for robbers
- Sentence 3 can mean a <Predation> if "masked men" are understood, somehow, to mean a group of <Predator> (e.g., killer whales, sharks,)
- Sentence 4 can mean a <Predation> if "a bank branch" is understood, somehow, to be a nickname for a whale or something of being a <Prey>
- Sentence 4 can mean a <Bank Robbery> if "a humpback whale and her baby" are understood, somehow, to mean a <Bank>

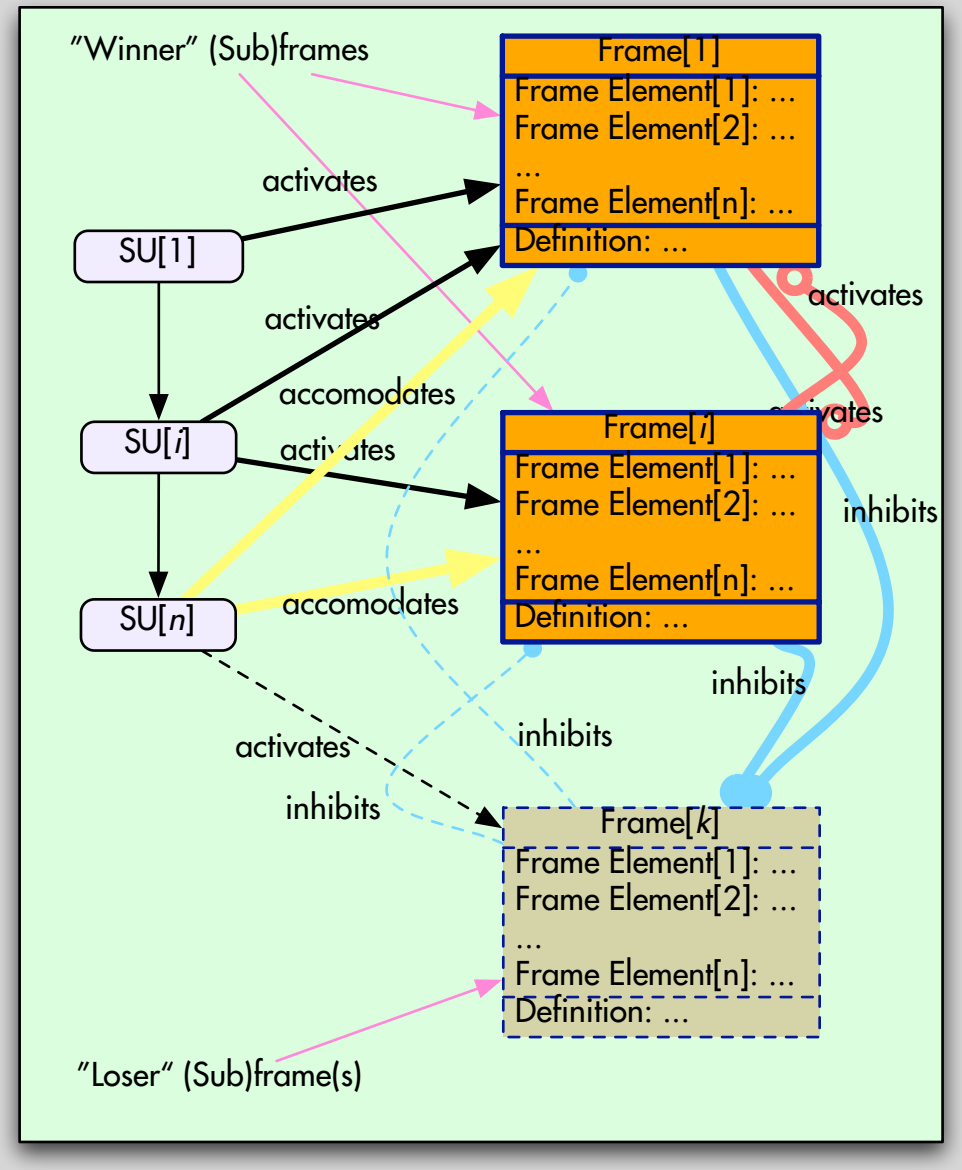
But How Come?

Supplementing Frame Semantics with a Theory of "Frame Selection"



- Each semantic unit SU "activates" a set of frames independently.
- Evoked frames "compete" each other either by mutual "activation" or lateral "inhibition"
- Once competition settles down, the (meaning of) SUs of the "loser" frames "accommodate" to the (meanings of) "winner" frames

Supplementing Frame Semantics with a Theory of "Frame Selection"



- Each semantic unit SU "activates" a set of frames independently.
- Evoked frames "compete" each other either by mutual "activation" or lateral "inhibition"
- Once competition settles down, the (meaning of) SUs of the "loser" frames "accommodate" to the (meanings of) "winner" frames

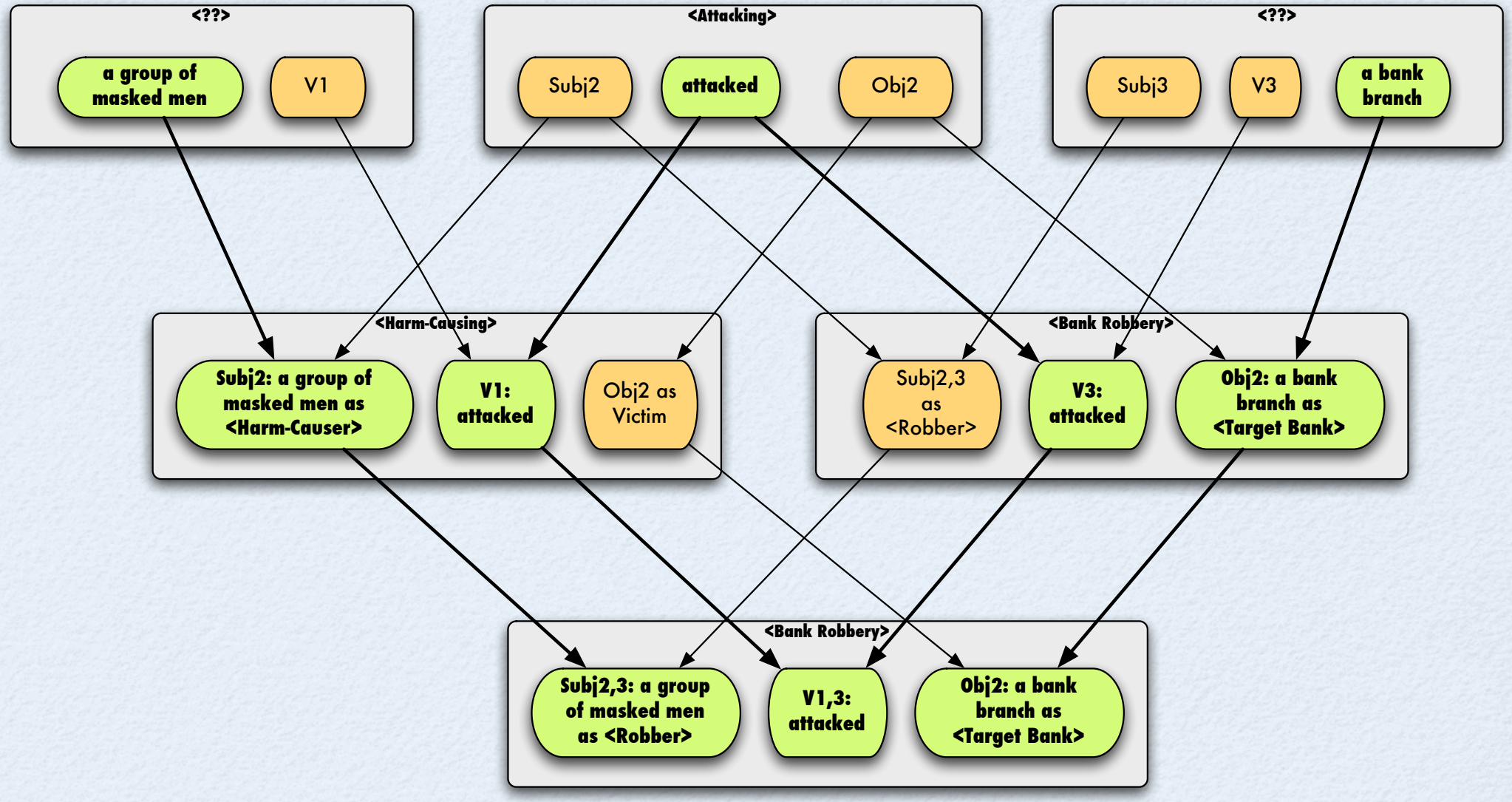
Supplementing Frame Semantics with a Theory of "Frame Selection"

- Given a sentence $s = w_1\ w_2 \cdots w_n$,
  - Each word w_i (or (possibly discontinuous) substring "$\ldots w_i \ldots w_j \ldots$") "evokes" a frame F_i independently each other, and "strengthen" or "suppress" each other.
- Competition among frames evoked takes place, and it "converges" when
  - the specifications of "loser" frames are adjusted to the specifications of the "winner" frames.
- This way, the set $\mathscr{F} = \{F_1, F_2, \ldots, F_n\}$ of evoked frames reduces into a smaller set of frames $\mathscr{F}'$.

Sample Analysis

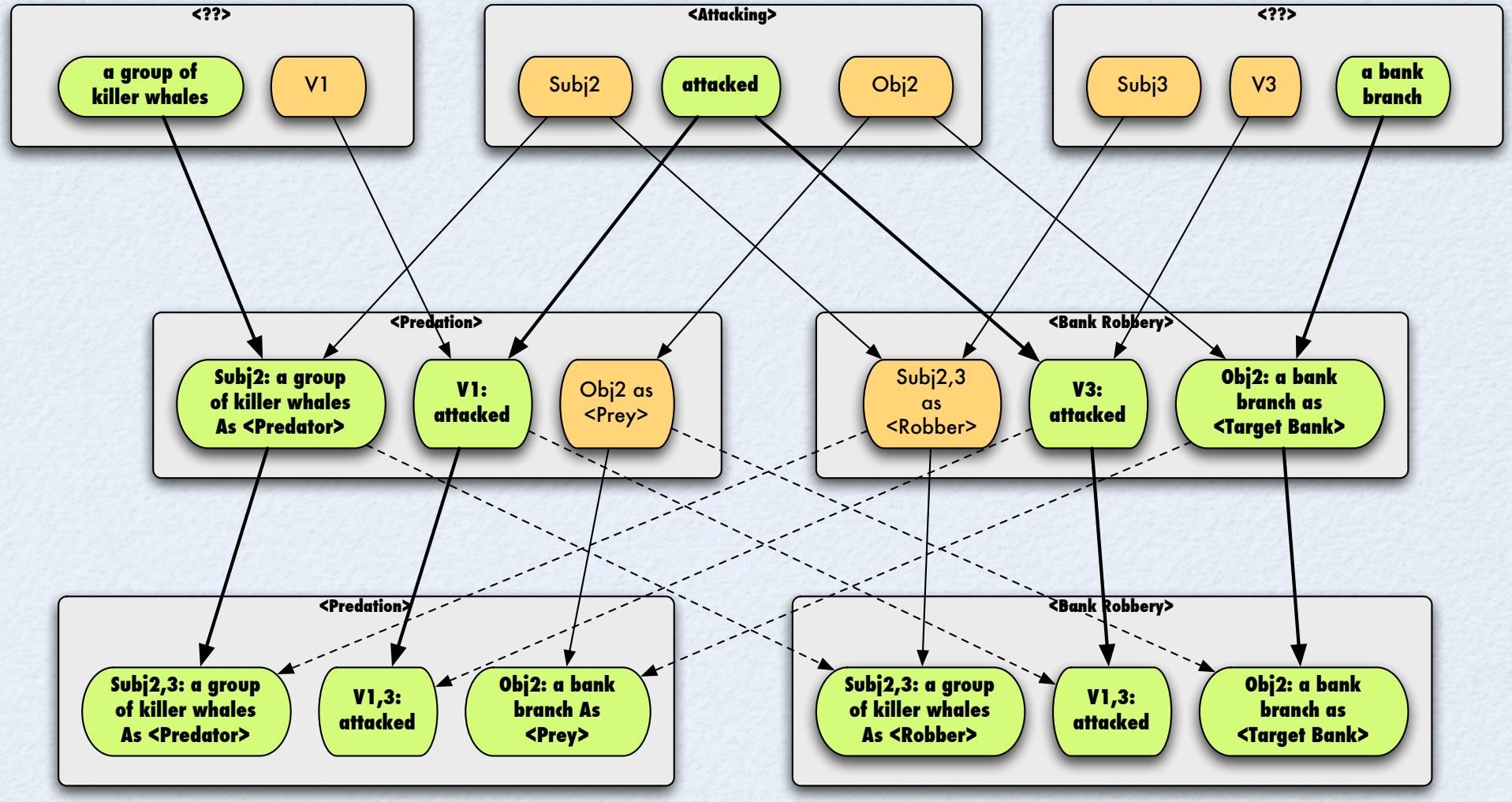
Frame Specification Flow of (2)

Feature Specification Flow



Frame Specification Flow of (3)

Feature Specification Flow (Adjustments Required)



Remarks

- Constituency plays virtually no role.
  - No effect from [NP [V NP]
- Frame specification flows, with and without bifurcation, should explain the "selectional restrictions" imposed on 1, 2, 3 and 4.
  - The origin of such restrictions are not really lexical one, as suggested by Fillmore in 70's.

Suggestion

- Complex units like the following are able to "evoke" finer-grained, specific frames:
  - "(a group of) killer whales attack ..."
  - "(a group of) masked men attacked ..."
  - "... attacked a bank branch ..."
  - "... attacked a humpback whale ..."
- Typically, this takes place when a role-denoting noun is combined with a verb.
- This fact needs to be considered in annotation tasks. MSFA does it.

MSFA Procedure (Simplified)

1. Segment a sentences S into units U_1, …, U_n.
   - This is not independent from Step 2. So, you need to go cyclic.
   - 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.
2. For each U_i, find a set of frames F_1, …, F_m so that one of their "frame elements" is *realized* by U_i.
3. Specify relationships among all the frames.

Guiding Principles of MSFA

- "Be meticulous"
  - Every word (or morpheme if morphological analysis is necessary) needs to realize at least one "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

How MSFA Goes
-Sample Analysis-

Sample MSFA

- The following is a text taken from *Kyoto University Corpus* (Kurohashi and Nagao 1994):

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. ...

Sample MSFA

- The English translations of the text:
  1. *A book titled "Inside the White House" will go on sale in the U.S. on January 14.*
  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. *...*

Sample MSFA (for Japanese)

	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; constitues F8; elaborates F9		presupposes F5	presupposes F9		constitutes F3,F5	
3	Frame idenfitier	作品への題名つけ	命名	執筆	制作	出版	商品販売	購入	消費	読書	楽しみ	米国での大統領による為政	秘密の暴露	報告
4	*													報告者
5	*					目的		GOVERNOR	GOVERNOR	手段				報告内容 [start1,end]
6	*							目的	手段	GOVERNOR	手段			
7	*								目的	目的	GOVERNOR			
8	*					小売り業者	販売者	販売者	提供者 [ternary]					
9	*					客	客	購入者	消費者	読者	楽しむ者			
10	*	題名付与者 [secondary]	命名者[+aux]			出版社			提供者 [secondary]					
11	*	題名付与者 [primary]	命名者	執筆者	作者	協力者?			提供者 [primary]				暴露者	
12	*			目的1									GOVERNOR	
13	「	MARKER[1,2]	MARKER[1,2]	特徴		出版物.特徴	商品.特徴	商品.特徴	消費物.特徴	読み物.特徴	楽しみ.特徴			報告内容 [start2,end]
14	ホワイトハウス	題名	名称									大統領官邸	標的	
15	の												MARKER	
16	内側												秘密: EVOKED	
17	」	MARKER[2,2]	MARKER[2,2]											
18	と	EVOKER1	EVOKER1											
19	題	GOVERNOR	GOVERNOR	EVOKER1										
20	する	EXTENDER	EXTENDER	EVOKER2										
21	本	作品	命名対象	書物: EVOKER3	作品	出版物	商品	商品	消費物	読み物	楽しみ.源			
22	が					MARKER	MARKER							
23	十四					日時:日付	日時:日付							
24	日													
25	、													
26	米国					場所	場所							
27	で					MARKER	MARKER							
28	発売			目的2		GOVERNOR	GOVERNOR						手段	
29	され					EXTENDER1	EXTENDER1							
30	る					EXTENDER2	EXTENDER2							
31	。													

Sample MSFA (for English)

	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		constittues F3,F5	
3	Frame	Title Giving	Name Giving	Writing	Authoring	Publishing	Selling	Purchasing	Consuming	Reading	Having Fun	Presidential Government in the U.S.	Disclosure	Reporting
4	*													Reporter
5	*					Purpose		**GOVERNOR**	**GOVERNOR**	Means				Report[start 1,end]
6	*							Purpose	Means	**GOVERNOR**	Means			
7	*								Purpose	Purpose	**GOVERNOR**			
8	*					Retailer	Seller	Seller	Provider3					
9	*					Customer	Customer	Purchaser	Consumer	Reader	Enjoyer			
10	*	Title Giver[secondary]	Name Giver[2]		Supporter	Publisher	Provider		Provider2					
11	*	Title Giver[primary]	Name Giver[1]	Writer	Author	Supporter?			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.attribtute	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	.													

Sample MSFA (for Japanese)

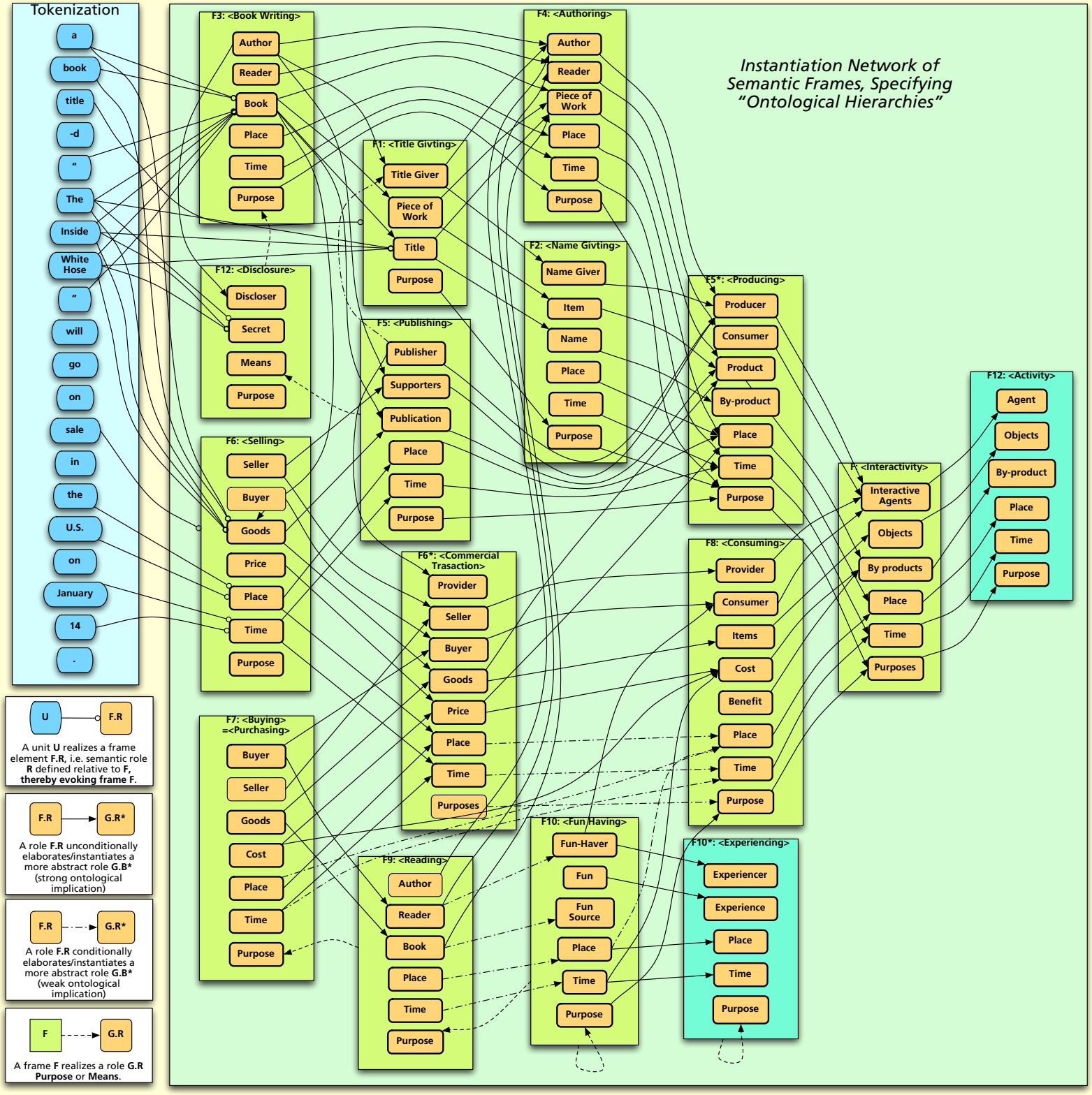
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	F-ID	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13
2	F-to-F relations	elaborates F2; constitutes F3		constitutes F5; presumes F5; elaborates F4		presupposes F3	presupposes F5; constitutes F5; presumes F7	presupposes F6; constitues F8; elaborates F9		presupposes F5	presupposes F9		constitutes F3,F5	
3	Frame idenfitier	Title Giving	Name Giving	Writing	Authoring	Publishing	Selling	Purchasing	Consuming	Reading	Having Fun	Presidential Government in the U.S.	Disclosure	Reporting
4	*													Reporter
5	*					Purpose		**GOVERNOR**	**GOVERNOR**	Means				Report[start1 ,end]
6	*							Purpose	Means	**GOVERNOR**	Means			
7	*								Purpose	Purpose	**GOVERNOR**			
8	*					Retailer	Seller	Provider	Provider[tern ary]					
9	*					Customer	Customer	Purchaser	Cosumer	Reader	Enjoyer			
10	*	Title Giver[seconda ry]	Name Giver[second ary]			Publisher			Provieder[sec ondary]					
11	*	Title Giver[primary]	Name Giver[primary]	Writer	Author	Supporter?			Provider[pri mary]				Revealer	
12	*			Purpose1									**GOVERNOR**	
13	「	MARKER[1,2]	MARKER[1,2]	Book.attribu te	Work.attrib ute	Publication.a ttribute	Goods.attribu te	Goods.attribu te	Commodity. attribute	Book.attribut e	Fun.attribute			Report[start2 ,end]
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	Publicationk	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			Purpose2		**GOVERNOR**	**GOVERNOR**						Means	
29	され					EXTENDER1	EXTENDER1							
30	る					EXTENDER2	EXTENDER2							
31	。													

Remarks

- MSFA is used:
  - to identify and specify as many frames as possible; each column, with a "Frame ID" (local variable) and a "Frame Name" (global variable) specifies a frame
  - to specify explicitly how frames are interrelated using Frame-to-Frame Relations (global) on the second row
- Conventions
  - The relative order of columns is not significant.
  - "Null instantiations" are indicated by * if they are position-neutral, and by ** if they are position-specific

What MSFA Does?

- MSFA above specifies, for example:
  - *Author*, as a role-denoting noun, designates an Agent-class role specific to <Authoring>.
  - *Writer*, as a role-denoting noun, designates an Agent-class role specific to <Writing>, a subclass of <Authoring>
  - etc
- This implicitly describes frame hierarchies and role/FE hierarchies like …


Tokenization
a
book
title
-d
"
The
Inside
White Hose
"
will
go
on
sale
in
the
U.S.
on
January
14
.
Instantiation Network of Semantic Frames, Specifying "Ontological Hierarchies"
F3: <Book Writing>
Author
Reader
Book
Place
Time
Purpose
F4: <Authoring>
Author
Reader
Piece of Work
Place
Time
Purpose
F1: <Title Givting>
Title Giver
Piece of Work
Title
Purpose
F2: <Name Givting>
Name Giver
Item
Name
Place
Time
Purpose
F12: <Disclosure>
Discloser
Secret
Means
Purpose
F5: <Publishing>
Publisher
Supporters
Publication
Place
Time
Purpose
F5*: <Producing>
Producer
Consumer
Product
By-product
Place
Time
Purpose
F12: <Activity>
Agent
Objects
By-product
Place
Time
Purpose
F6: <Selling>
Seller
Buyer
Goods
Price
Place
Time
Purpose
F6*: <Commercial Trasaction>
Provider
Seller
Buyer
Goods
Price
Place
Time
Purposes
F8: <Consuming>
Provider
Consumer
Items
Cost
Benefit
Place
Time
Purpose
F: <Interactivity>
Interactive Agents
Objects
By products
Place
Time
Purposes
F7: <Buying> =<Purchasing>
Buyer
Seller
Goods
Cost
Place
Time
Purpose
F9: <Reading>
Author
Reader
Book
Place
Time
Purpose
F10: <Fun Having>
Fun-Haver
Fun
Fun Source
Place
Time
Purpose
F10*: <Experiencing>
Experiencer
Experience
Place
Time
Purpose
U
F.R
A unit U realizes a frame element F.R, i.e. semantic role R defined relative to F, thereby evoking frame F.
F.R
G.R*
A role F.R unconditionally elaborates/instantiates a more abstract role G.B* (strong ontological implication)
F.R
G.R*
A role F.R conditionally elaborates/instantiates a more abstract role G.B* (weak ontological implication)
F
G.R
A frame F realizes a role G.R Purpose or Means.

Frame-to-Frame Relations

- The partial list of Frame-to-Frame relations we have defined so far is:
- "*F* elaborates *G*" (deals with Inheritance, "Is-A")
- "*F* constitutes *G*" (deals with "Part-Of" relation)
- "*F* presupposes *G*", "*F* negates *G*" (deals with "implications")
- "*F* motivates *G*" (can be used to specify <Reason>)
- "*F* realizes *G*" (can be used to specify <Purpose>)

Benefits of Multilayered Analysis

- Multilayered analysis has its own benefits.
  - it allows us to explore the details of *frame-to-frame relations*
    - Full text analysis tells much more about them
  - it allows us to explore and specify *multiple, simultaneous role realization* by a lexical material.
  - it allows us to avoid *frame conflation*
    - this happens all the time
  - by disentangling a complicated relationship among frames evoked in a sentence

Role Multiplex and Relativized Role

- In the following sentence, the role of 松葉づえ(で) is essentially ambiguous:

  (1) 研究室で友人と話していると，そこに太郎が松葉づえで入ってきた.

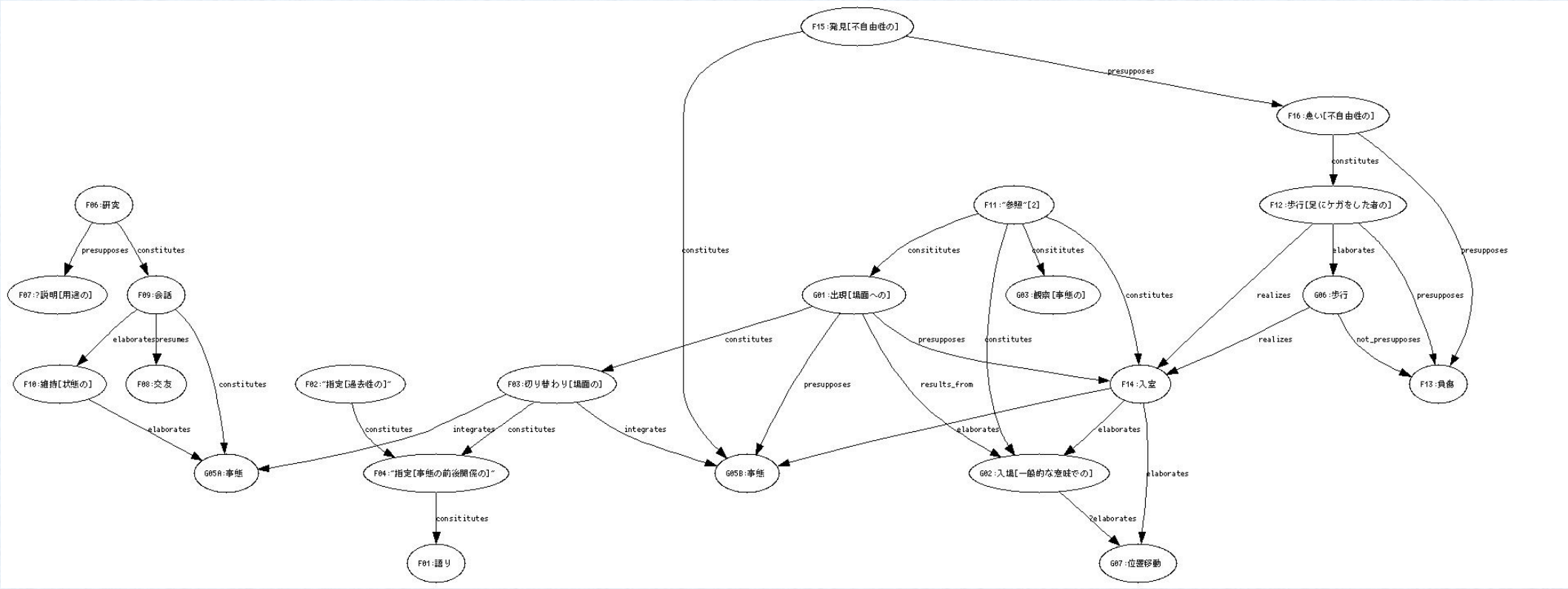
- in that it realizes at least the following two roles simultaneous and inclusively:
  - R1: <歩行者>「太郎」にとっての<歩行の道具> IS-A <Instrument>
  - R2: 観察者 (語り手とその友人にとっての<出現の際の様態> IS-A <Manner>

例文(1)のMSFA

Frame ID (Local)	F01	F02	F04	F03	G05a	G05b	F06	F07	F09	F10	F08	F11	F11	G01	G03	F14	G02	G07	F12	G06	F13	F15	F16
Frame-to-Frame Relations (Global)		constitutes F04	consititute s F01	constitutes F04; integrates G05a,G05b			presuppose s F07; constitutes F09		presumes F08; elaborates F10; constitutes G05a	elaborates G05a		constitutes F14,G02	consititute s G01,G03	constitutes F03; presuppose s F14,G05b; results_fro m G02		elaborates G02,G05b, G07	?elaborate s G07		presuppose s F13; realizes F14; elaborates G06	not_presup poses F13; realizes F14		presuppose s F16; constitutes G05b	presuppose s F13; constitutes F12
Frame Name (Global)	語り	~指定[過去性の]~	~指定[事態の前後関係の]~	切り替わり[場面の]	事態	事態	研究	?説明[用途の]	会話	維持[状態の]	交友	~参照~[1]	~参照~[2]	出現[場面への]	観察[事態の]	入室	入場[一般的な意味での]	位置移動	歩行[足にケガをした人の]	歩行	負傷	発見[不自由性の]	患い[不自由性の]
*	聞き手																						
*	語り手	指定者	指定者	場面[1]	参与者[1]		研究者 [potential] [1]	利用者[1]	参加者[1]	維持者[1]	交友者[1]	参照者	参照者	観察者[1]	観察者[1]							知覚者[1]	
研究	内容	事態	事態[1]		場所		GOVERNOR	目的[利用の]	場所	場所	場所	参照先	参照先	観察者 [1,2].Attr[1,2]	観察者 [1,2].Attr[1,2]								
室							場所	GOVERNOR															
で					MARKER				MARKER	MARKER	MARKER												
友人					参与者[2]		研究者 [potential] [2]	利用者[2]	参加者[2]	状態	相手: EVOKER			観察者[2]	観察者[2]							知覚者[1]	
と					MARKER				MARKER		MARKER			観察者 [1,2].Attr[2,2]	観察者 [1,2].Attr[2,2]								
話し					参与の仕方				GOVERNOR		?手段												
て									状態	GOVERNOR	MARKER												
いる																							
と			EVOKER	EVOKER																			
、			EXTEDNER	EXTDNER																			
そこ			事態[2]	場面[2]		場所						参照元	参照元	場面	場所	目的地点	場所	目的地点	到着地点	到着地点			
に						MARKER								MARKER		MARKER	MARKER	MARKER	MARKER	MARKER			
太郎						参与者	研究者 [potential] [3]	利用者[3]			相手 [potentila]			出現者	事態	入室者	入場者	位置移動者	歩行者	歩行者	負傷者	不自由者	?患者
が						MARKER								MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER
松葉づえ						参与の仕方								様態	様態	様態[入室の] OR 手段[入室の]	様態[入場の] OR 手段[入場の]	様態[入場の] OR 手段[入場の]	道具		EVOKER	様態: EVOKER	様態: EVOKER
で														MARKER	MARKER	MARKER	MARKER	MARKER	MARKER				
入っ														EVOKER	EVOKER	GOVERNOR	GOVERNOR	EVOKER	EVOKER	EVOKER		不自由性を伴う行動	不自由性を伴う行動
て																EXTENDER	EXTENDER	EXTENDER					
き																							
た		GOVERNOR				EXTEDNER																	
。	EXTENDER	EXTENDER	EXTENDER	EXTEDNER																			

http://www.kotonoba.net/~mutiyama/cgi-bin/hiki/hiki.cgi?c=view&p=msfa-matsubadzue-DE

Network of Frames



http://www.kotonoba.net/~mutiyama/cgi-bin/hiki/hiki.cgi?c=view&p=msfa-matsubadzue-DE

	出現[場面への]	観察[事態の]	入室	入場[一般的な意味での]	位置移動	歩行[足にケガをした人の]	歩行	負傷	発見[不自由性の]	患い[不自由性の]
*										
*	観察者[1]	観察者[1]							知覚者[1]	
研究	観察者[1,2].Attr[1,2]	観察者[1,2].Attr[1,2]								
室										
で										
友人	観察者[2]	観察者[2]							知覚者[1]	
と	観察者[1,2].Attr[2,2]	観察者[1,2].Attr[2,2]								
話し										
て										
いる										
と										
、										
そこ	場面	場所	目的地点	場所	目的地点	到着地点	到着地点			
に	MARKER		MARKER	MARKER	MARKER	MARKER	MARKER			
太郎	出現者	事態	入室者	入場者	位置移動者	歩行者	歩行者	負傷者	不自由者	?患者
が	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER
松葉づえ	様態	様態	様態[入室の] OR 手段[入室の]	様態[入場の] OR 手段[入場の]	様態[入場の] OR 手段[入場の]	道具		EVOKER	様態: EVOKER	様態: EVOKER
で	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER				
入っ	EVOKER	EVOKER	GOVERNOR	GOVERNOR	EVOKER	EVOKER	EVOKER		不自由性を伴う行動	不自由性を伴う行動
て			EXTENDER	EXTENDER	EXTENDER					
き										
た										
。										

Web Sites that Host MSFA Samples

- FOCAL Hiki Site 1
  - http://www.kotonoba.net/~mutiyama/cgi-bin/hiki/hiki.cgi?FrontPage
  - Unrestricted access to MSFA-based sample annotation of copyright-free J-E translations
- FOCAL Hiki Site 2
  - http://www.kotonoba.net/~mutiyama/cgi-bin/hiki2/hiki.cgi?FrontPage
  - Restricted access (you need an account) to MSFA-based sample annotation of J-E newspaper article alignment data.

Dealing with Multiplicity

WSD Needs to Be Done Frame-wise

- A single "entity" referred to in a text is likely to *realize different frame elements frame-wise simultaneously*.
- For instance, *book* in the previous example realizes:
  - <Information Carrier> FE in <Reading> frame
  - <Goods> FE in <Selling/Buying> frame
  - <Piece of Work> FE in <Writing> frame
  - <Publication> FE in <Publishing> frame
- This means that word sense disambiguation *needs to be done frame-wise*, explaining why WSD isn't enough for text understanding.

How to Deal with "Markers"?

- Multiple, simultaneous realization of a set of FEs poses the following problem:
  - How to treat prepositions in languages like English and postpositions in languages like Japanese?
- This boils down to the following question:
  - Are prepositions or particles really *parts* of FEs, or just *markers* of them?
- We adopted the idea that markers can be (and need to be) separated from FEs themselves to facilitate encoding of multiple role realization.

For Future

Current Status

- So far, MSFA was applied to
  - a portion of copyright-free Japanese-English alignment data (67 sentences), open data
  - a tiny portion of *Kyoto University Corpus* (KUC) texts (3 articles, 63 sentences); semi-closed data
- Characteristics
  - MSFA assigns a sentence 20-30 frames on average (depending on the length)
  - 2157 frames (=927+1227: type count), 6846 frame elements (=3031+3815: type count) are recognized
  - But no evaluation is done yet.

What Should We Do Next?

- Our own frame definitions need to be linked to BFN ones. But, at the time being, it seems to be unrealistic to establish a mapping between BFN frames and the frames we defined with MSFA
  - granularity difference matters
- But it would be useful if we try to *share an annotation scheme* to facilitate such mappings in the future, and we want to try out best to do it.
- Putting aside the granularity issue, the major problem is with the treatment of markers like prepositions and postpositions.

Thank You

Acknowledgements

My research is indebted to contribution by the following people:

Toshi-yuki KANAMRU, Kyoto University
Keiko NAKAMOTO, Bunkyo University
Jea-Ho LEE, Hajime NOZAWA, Yoshikata SHIBUYA, Masao UTIYAMA, NICT
Koichiro NAKAMOTO