複層意味フレーム分析 (MSFA) に意味役割の典型的実現値の情報を付加してシソーラス化する試み

黒田 航　　井佐原 均

独立行政法人 情報通信研究機構けいはんな情報通信融合研究センター

問題意識と目的

何らかの意味型 T (e.g., [ヒト]) をもつ実体 (entity) X (ID で区別される) が意味役割 R (e.g., <犯罪者>, <犯人>) という意味クラスの実現値 (インスタンス) (e.g., 「(R_1: 強盗殺人(のR_2: 容疑)で) X を逮捕した」)になる．この実体と意味役割と意味型との関係を，なるべく網羅的に明示することは，シソーラスに期待されていながら実装されていない機能の一つ．

提案

細かい粒度の意味解析のための手法として提唱された複層意味フレーム分析 (MSFA) に，問題となる機能を融合する

背景となる理論

文 S の MSFA は一般に S の内容理解に使用される背景知識を意味フレーム $\{F_1, F_2, ..., F_n\}$ という形で網羅的に明示する．

この際，おのおのの F_i を構成する意味役割 $\{F_i.R_1, ..., F_i.R_k\}$ について，文 S という文脈で与えられた実現値以外に可能な実現値を明記することで，必要となる情報をエンコードできる．

フレーム名	意味役割名とその異名	典型表現	実現値1	実現値2	実現値3	実現値4	実現値5	実現値6	実現値7	実現値8	実現値9	実現値10	実現値11	実現値12	実現値13
~過去性の指定	*事態	<*事態>た.	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****
~過去性の指定	*指定者[過去性の]		語り手												
喚問	*(行)喚問者=裁判所	<*(行)喚問者>が, <*応喚問者>を, <*喚問の理由>で, (裁判所に) {呼び出す,喚問する}.	裁判所[+exclusive]	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****
喚問	*応喚問者,*喚問に応じる者	<*応喚問者>が, <*喚問の理由>で, <*(行)喚問者=裁判所>の{呼び出し,喚問}に, {応じる, 従う}.	被告人(~)	(~)被告	証人(~)	[人名]									
喚問	*喚問(の)内容		出廷												
喚問	*喚問の理由		証言	弁明	事情説明	反論	陳情	意見陳述	口頭弁論	証拠提示	弁論				
起訴	*起訴者	<*起訴者>が<*起訴の理由>で<*被起訴者>を起訴する	検察	告訴者	原告(側)										
起訴	*起訴の理由		起訴事実	~法違反	~の罪	~の疑い	殺人(罪)	強盗	暴行	窃盗	過失致死	過失致傷	放火	横領	~罪
起訴	*被起訴者=被告(人)	<*起訴者>が, <*被告>を, <*容疑>で, 訴える; <*起訴者>が, <*被告>に, <*容疑>の訴えを起こす; <*裁判所>が, <*起訴事実>を, 審理する	(~)容疑者[+exclusive?]	[人名]	[組織名]										
刑事裁判	*(行)裁判者=裁判官		(~)裁判官[+exclusive]	(~)裁判長											
刑事裁判	*受裁判者=被告(人)		(~)被告[+exclusive]	被告人(~)[+exclusive	[人名]	[組織名]									
刑事裁判	*起訴の理由=起訴事実=容疑		[犯罪名]の容疑												
刑事裁判	*起訴者=検察		検察[+exclusive]												
刑事裁判	*裁判機関=裁判所		裁判所[+exclusive]	地裁	高裁	最高裁									
刑罰	*行刑罰者,*加刑罰者	<*行処罰者>が, <*受刑罰者>を, <*刑罰の内容>の刑に, 処す.	権力?	権力者?	当局?										
刑罰	*刑罰の内容		[犯罪名]の刑	禁固刑	死刑	磔,はりつけ(獄門)	([n]年の)懲役	無期懲役	([n]円の)罰金	罰金([n]円)	執行猶予?				
刑罰	*受刑罰者		[判決.*有罪者]	犯人	容疑者										
逮捕[犯罪者の]	*行逮捕者	<*行逮捕者>が, <*被逮捕者>を, {逮捕,捕まえる}する>	警察(当局)	検察((当局)	当局										
逮捕[犯罪者の] 逮捕[犯罪者の]	*(被)逮捕者=逮捕者 *逮捕の理由=容疑		容疑者	違反者	犯罪者	(~の)犯人	[人名]								
検挙[犯罪者の]	*行検挙者	<*行検挙者>が, <*被検挙者>を, {検挙する,アゲる}	容疑	~の疑い											
検挙[犯罪者の]	*(被)検挙者=検挙者		警察(当局)	検察((当局)	当局										
事件報道	*受報道者		検挙者	違反者	犯罪者	(~の)犯人	[人名]								
事件報道	*報道者	<*報道者>が<*受報者>に<*報道内容>を報道して知らせる	視聴者	読者	市民	大衆	リスナー	[人名]							
事件報道	*報道内容		[報道機関名]	[報道メディア名]	(~)新聞	(~)テレビ	(~)ラジオ	(~)放送	[番組名]	(~)アナウンサー	(~)ニュースキャスター	(~)記者	[人名]		
支払い[罰金の]	*(支払い)額=罰金	<*支払い手>が<*支払先,*支払い相手>に<*(支払い)額>を罰金として支払い,<*見返り>を受ける	(~)事件	ニュース	記事	ルポ	特ダネ	ネタ	[事実]						
			罰金[n]円	罰金[n][通貨の単位]											

MSFAの実例 (部分)

アルゼンチンの元サッカー選手，ディエゴ・マラドーナ氏が六日，両国の検察当局に身柄を拘束された．[京大コーパス S-ID 950107210-002]

他の MSFA データも公開中 (要アカウント)

MSFA に現われた意味役割の実現値の付加情報 (断片)

ただし，実現パターンはそれなりに複雑で，その記述には適当なパターン記述言語が必要

一般的注意

この際，従来のシソーラスのように，語の意味分類を，語彙分類としてトップダウンに行なうのではなく，実際に文章で使われている意味クラスを，なるべく細かい粒度で認定し，その語彙的実現値のデータベース化という形で，語の意味分類をボトムアップに行なう

個別的注意

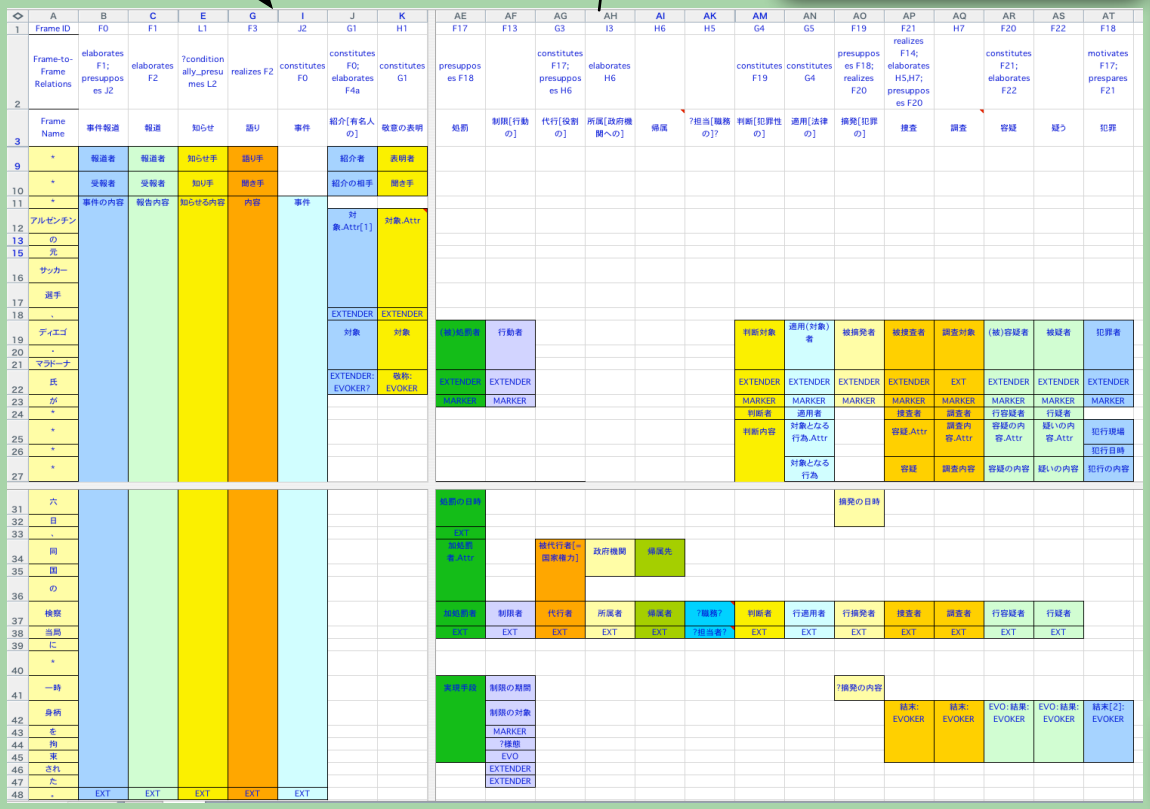
意味型に関する情報は MSFA にはハッキリ現われないので，実現値の補完表にも現われない．意味型が直接に言及されるのは実は意外に稀．Gentner 2005 など，relational nouns と object/entity nouns の区別を論じた論文にも同様の指摘がある．

意味型は実体による意味役割の実現の前提条件となることが多い．例えば，[個体 X がある災害 Y の被害者 Z であるためには X は人(か少なくとも動物) でないといけない] (この指定は Z にある) など，意味型はどっちかというと潜伏性の意味特徴＝素性のようにふるまう．

固有名は正確には意味型名でも意味役割名でもないが，強いて言えば，インスタンスを一つしかもたない意味型か意味役割？ただし，

彼は平成の円空だ

のような文では，固有名がインスタンスを複数もつように拡張されている．これが意味型名ベースの拡張なのか，意味役割名ベースの拡張なのかは不明．

謝辞

内山 将夫，中本敬子，野澤 元，横森大輔