

言論マップ生成のための 事象間類似・対立関係の認識

村上浩司 松吉俊(NAIST) 増田祥子(大阪府立大/NAIST)
松本裕治(NAIST) 乾健太郎(NICT)

Webアクセスの現状

Web : 様々な情報に容易にアクセス可能

- 知りたいことに関しての多くのページを参照
- 便利は有益か？
 - 特定のサイトの情報を信頼、信用
 - 適切、不適切？
(不正確、偏りのある意見、古い情報)

クエリによるWeb検索

- 広範囲の情報収集が可能
- 有益な情報の特定が困難
- 情報検索のランキングは信憑性を反映しない

偏在する特定の言論を中心的に捉えてしまう危惧

大量の情報を俯瞰する

情報アクセスの代償

- ユーザが個々の情報の信憑性、信頼性を判断
- 限られた時間である程度大量の情報を調査

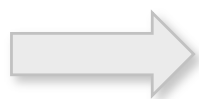
個人での大量の情報の俯瞰は極めて困難

言論の相対的な関係の把握

- Web中の数的優位な立場の言論は正論か？
- 多様な言論を抽出
- 言論間の論理的関係(類似、対立、含意など)の解析
- ユーザへ言論の信頼性判断を支援

Webアクセスの現状

- ・ 情報検索のランキングは信憑性を反映しない
(例: Googleにおいて“マイナスイオン”の上位10位)
- ・ 偏った言論を鵜呑みにしてしまう危険



情報の俯瞰が必要

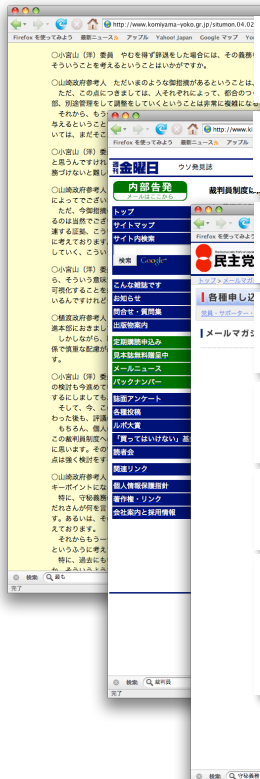
言論の相対的な関係の把握

- ・ 俯瞰するための情報をユーザに提供
- ・ 言論マップを生成

Web上のテキスト情報間の論理的関係解析

例：裁判員制度ってどうなんだろう？
どんな問題があるのかな？

裁判員制度に関するページ検索



もっとも問題だと言われているのが守秘義務のことだと思います。

裁判員制度には問題点が多々ある

最も問題なのは、守秘義務だと考えています。

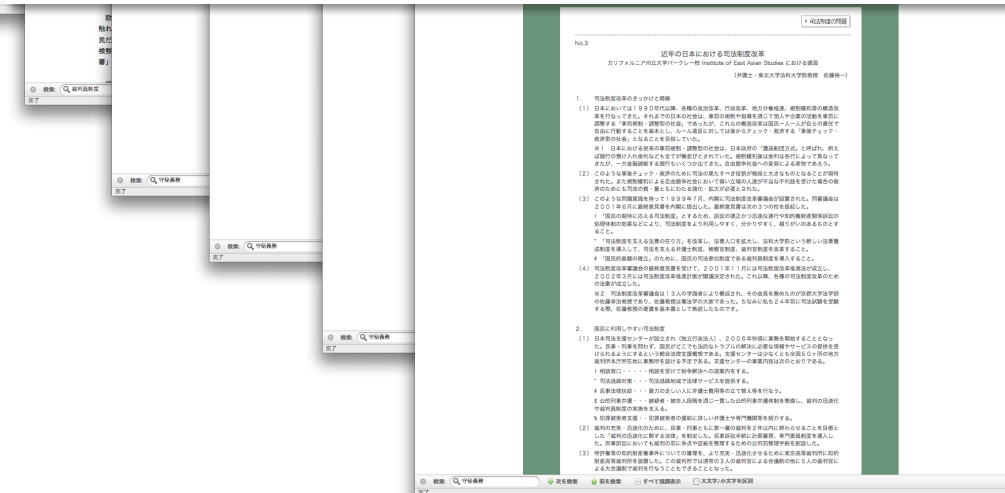
裁判員制度を導入することは問題が多すぎる。

裁判員への広範な守秘義務、裁判員への情報提供を処罰するなど問題点

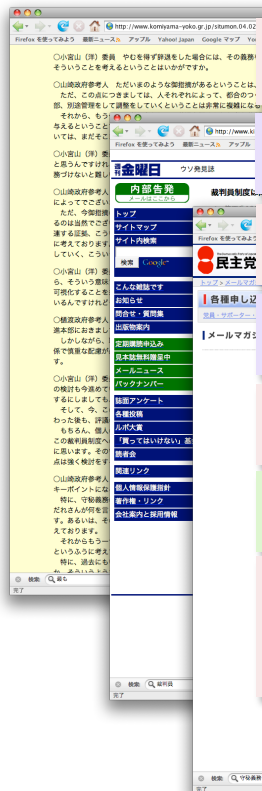
司法への国民参加の制度として画期的な意義を持つ

裁判上知りえた事実を守秘義務によって終生口外してはならない規定が裁判員の負担になる

裁判員制度は司法に国民が参加するという画期的な制度である



言論間のクラスタリングと論理的関係解析



もっとも問題だと言われているのが守秘義務のことだと思います。

裁判員制度には問題点が多々ある

最も問題なのは、守秘義務だと考えています。

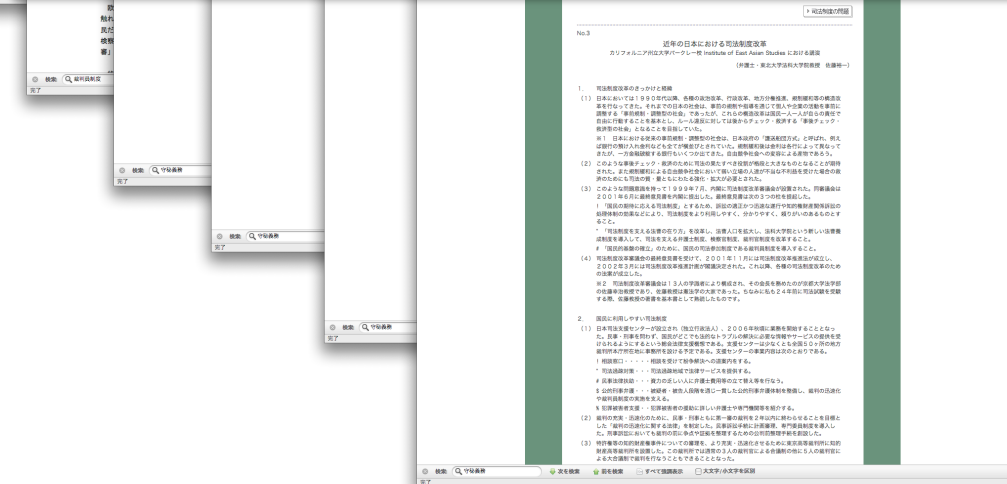
裁判員制度を導入することは問題が多すぎる。

裁判員への広範な守秘義務、裁判員への情報提供を処罰するなど問題点

司法への国民参加の制度として画期的な意義を持つ

裁判上知りえた事実を守秘義務によって終生口外してはならない規定が裁判員の負担になる

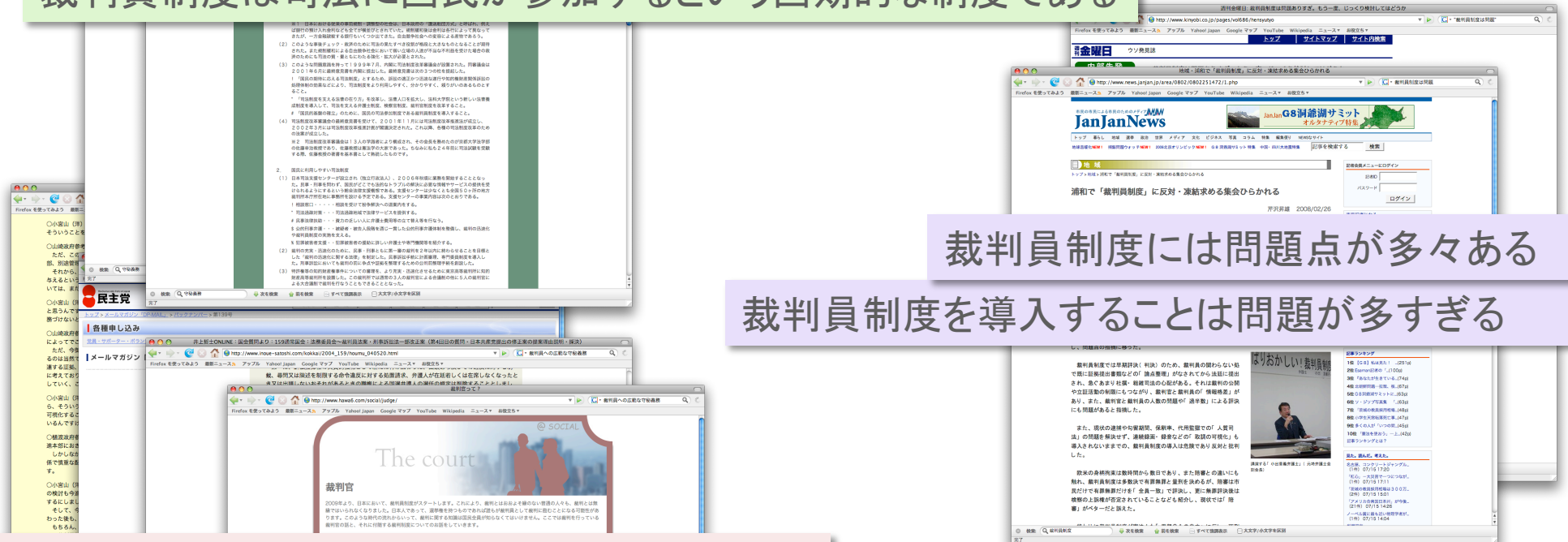
裁判員制度は司法に国民が参加するという画期的な制度である



言論のクラスタリングと論理的関係解析

司法への国民参加の制度として画期的な意義を持つ

裁判員制度は司法に国民が参加するという画期的な制度である



裁判員制度には問題点が多々ある
裁判員制度を導入することは問題が多すぎる

最も問題なのは、守秘義務だと考えています

もっとも問題だと言われているのが守秘義務のことだと思います

裁判員への広範な守秘義務、裁判員への情報提供を処罰するなど問題点
裁判上知りえた事実を守秘義務によって終生口外してはならない規定が裁判員の負担になる

言論のクラスタリングと論理的関係解析

司法への国民参加の制度として画期的な意義を持つ

裁判員制度は司法に国民が参加するという画期的な制度である

<対立>

裁判員制度には問題点が多々ある

裁判員制度を導入することは問題が多すぎる

<根拠・例示>

最も問題なのは、守秘義務だと考えています

もっとも問題だと言われているのが守秘義務のことだと思います

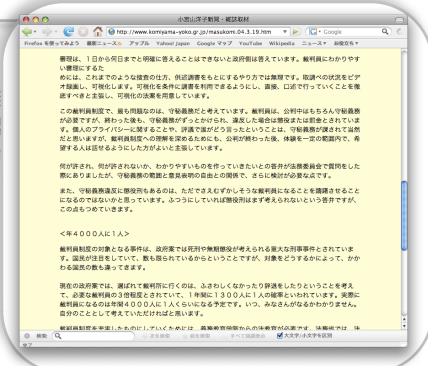
裁判員への広範な守秘義務、裁判員への情報提供を処罰するなど問題点

裁判上知りえた事実を守秘義務によって終生口外してはならない規定が裁判員の負担になる

論理的関係の解析

- ・司法への国民参加の制度として画期的な意義を持つ
- ・裁判員制度は司法に国民が参加するという画期的な制度である
- ・従来から問題になっている裁判のスピードアップ化という裁判員制度の一つの理念には賛成している
- ・裁判員制度の導入によって、現在の裁判の、解りにくいとか時間がかかりすぎるなどの問題が

- ・裁判員制度は問題が多い
- ・裁判員制度の開始にあたっては残された問題も多くあります。
- ・裁判員制度を導入することは問題が多すぎる



クラスタ: 守秘義務は問題だ

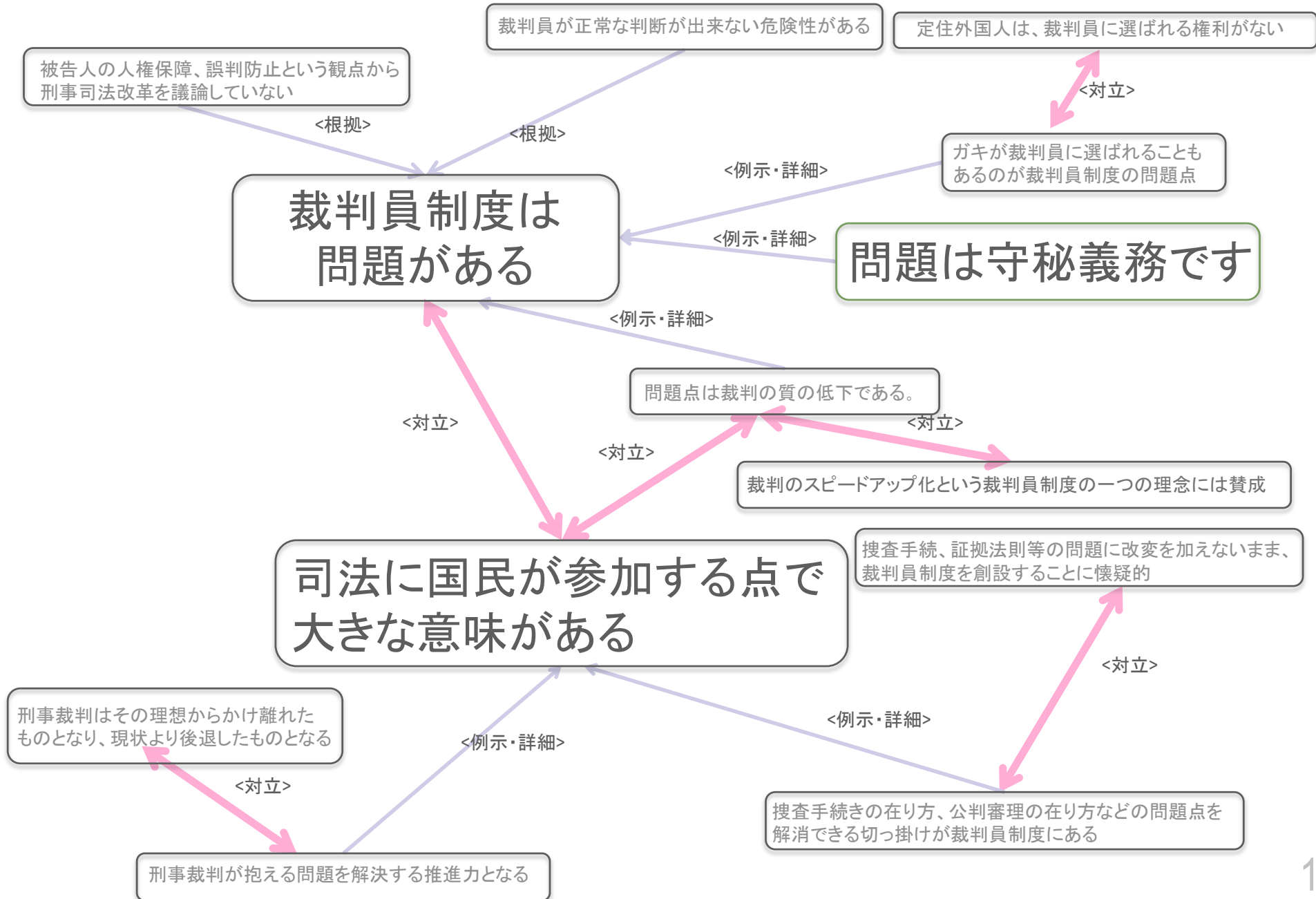
言論数 : 4

- ・この裁判員制度の中で最も**問題**だと言われているのが**守秘義務**のこと**だ**と思います。
- ・この制度で、最も**問題**なのは、**守秘義務**だと思っています。
- ・裁判員法案は、裁判員への広範な**守秘義務**、裁判員への情報提供を処罰するなど**問題点**はあります
- ・裁判上知りえた事実を**守秘義務**によって終生口外してはならない規定が裁判員の**負担**になる

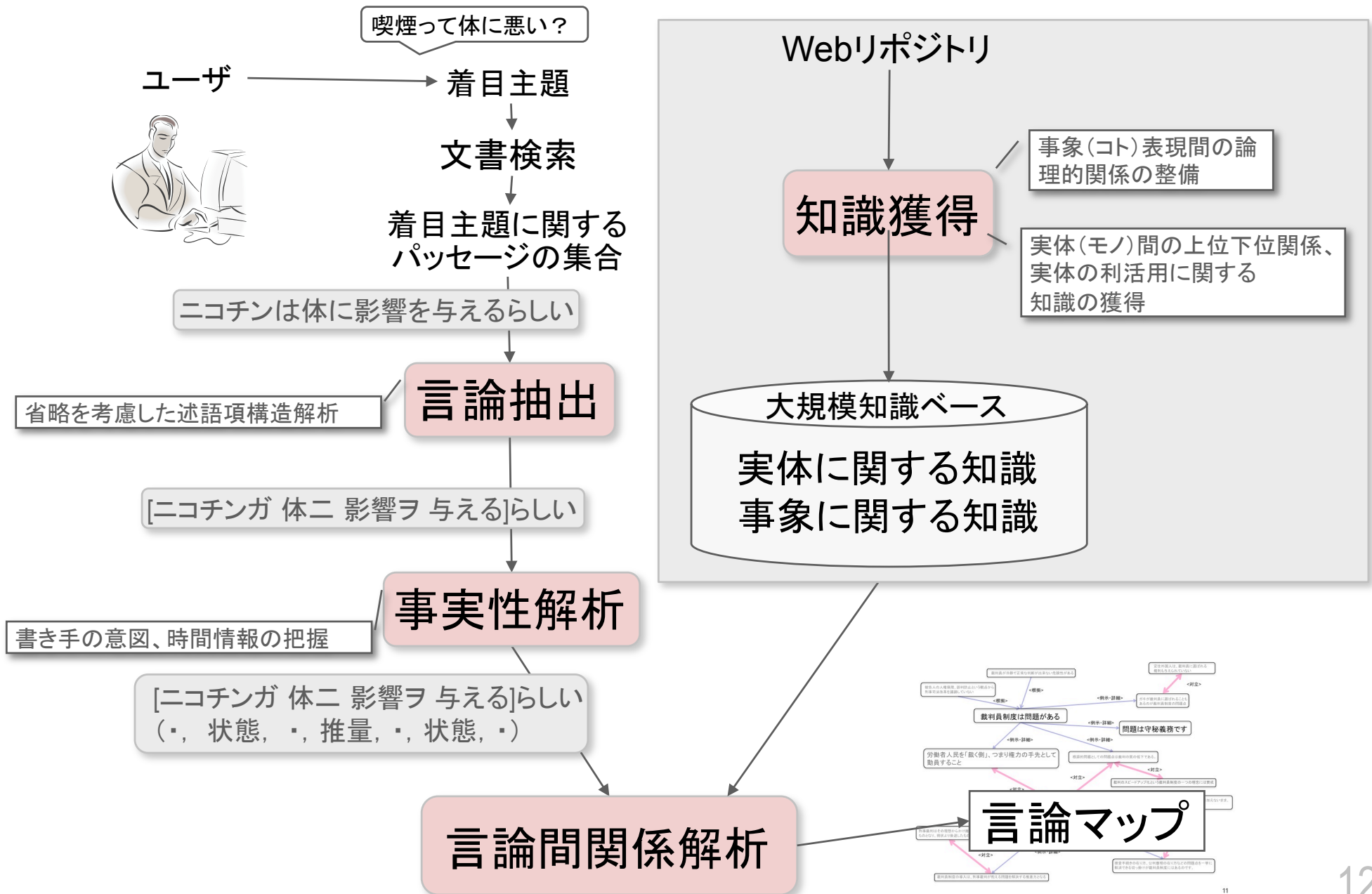
言論マップとは

言論間の論理的関係を解析し、個々の言論をノード、その間の関係をエッジとするグラフ

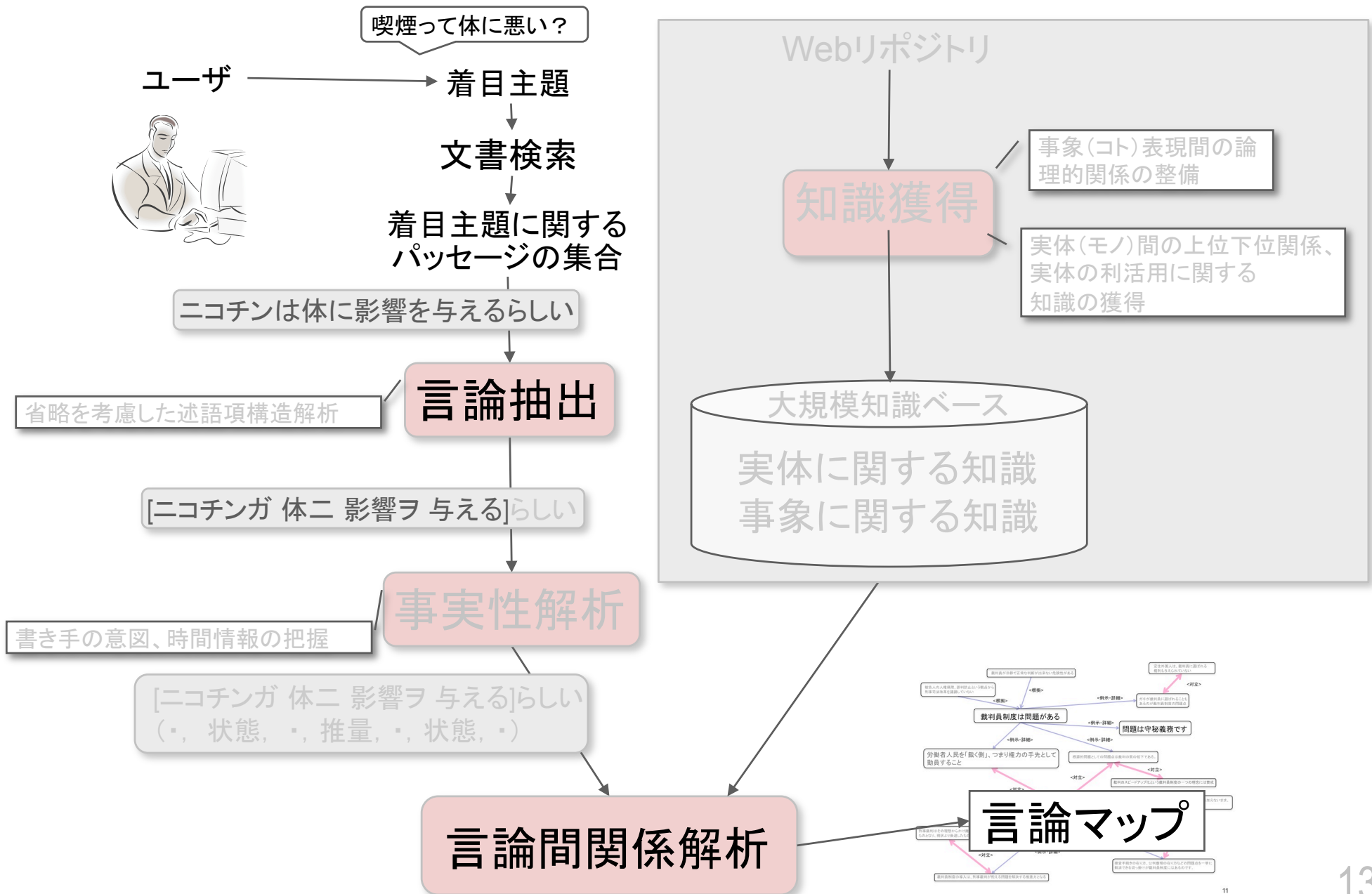
言論マップ出力例



言論マップ生成 全体の処理



言論マップ生成 全体の処理



言論の単位と言論間の論理的関係

例) 鯨の頭数が十分に回復している
⇔ 鯨が絶滅の危機にさらされている

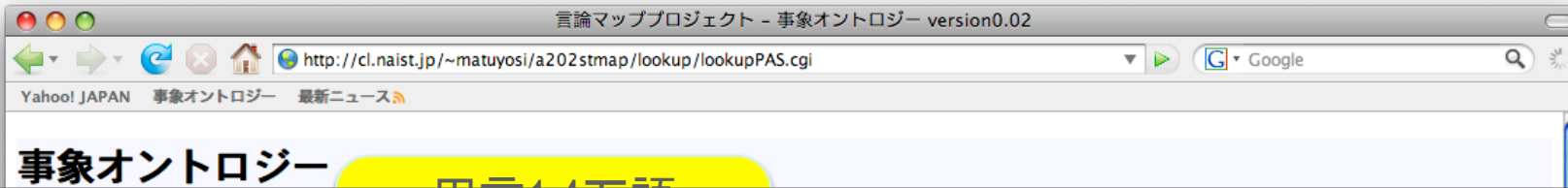
単純命題: 述語項構造(節)に概ね相当する言語表現
+ 述語の極性 + 否定
+ テンス + アスペクト + モダリティ

近い将来的には、より複雑な構造(仮定節など)も
考慮した「複合命題」に対応

取り扱う言論間の論理的関係

- ・ 同義・類義
- ・ 矛盾・対立・反論

事象間関係知識の整備 (1)



「吸収する」

事象間関係知識のエントリの例

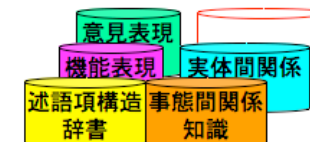
ID	04992 / 0011538-0-0-0-0-x0
表記	吸収する (きゅうしゅうする)
意味クラス	状態変化あり.位置変化.位置関係の変化 (物理) .吸入/排出.吸入
が格	<何か>.causer
を格	<何か B>.対象

1. <何か>が<何か B>を吸収するということは、<何か>が<何か B>を<何か>に吸い込むということで、逆も成り立つ
2. <何か>が<何か B>を吸収するということは、<何か>が<何か B>を吸い取るということで、逆も成り立つ
3. <何か>が<何か B>を吸収するということは、同時に、<何か>が<何か B>を取り入れることも起こる
4. <何か>が<何か B>を吸収するということは、<何か>が<何か B>を自分のものとするということであるが、逆は成り立つかどうか分からない

を格 <何か B>.対象

1. <何か>が<何か B>を吸収するということは、<何か>が<何か B>を<何か>に吸い込むということで、逆も成り立つ
2. <何か>が<何か B>を吸収するということは、<何か>が<何か B>を吸い取るということで、逆も成り立つ
3. <何か>が<何か B>を吸収するということは、同時に、<何か>が<何か B>を取り入れることも起こる
4. <何か>が<何か B>を吸収するということは、<何か>が<何か B>を自分のものとするということであるが、逆は成り立つかどうか分からない

ID	27432 / 0011538-0-0-0-0-x2
表記	吸収する (きゅうしゅうする)
意味クラス	状態変化あり.位置変化.位置関係の変化 (物理) .吸入/排出.吸入
が格	<何か>.動作主
を格	<何か B>.対象



完了

事象間関係知識の整備 (2)

事象間関係知識として述語項構造辞書の編纂

- ・ 2つの言語資源を統合、整理
- ・ 10種類の関係
 - 同義(言い換え)、同義・上位、反義語、
前提条件、結果(状態)、付帯状況、不可分、
原因・理由、目的、手段
- ・ 1.4万の用言について46,000対の関係知識を整備
- ・ 約4,000の主要な動詞に対して意味クラス情報を付与

言論マップ評価コーパス

言論マップ生成システムの出力を評価

- ・ 単純命題を対象とした評価コーパス
 - システムへの入力と共通の言論を対象
 - 同義・類似言論をクラスタ化
 - クラスタを単位として、類義以外の関係をクラスタに対して付与

◇	A	B	C	D	E	F	G	H	I	J	K	L	M
1	subtopic	id	fre	flag	arguments	presicate	cid	代表	付	不完	antonym	id2	example
4044	かゆみ	72	2	1	かゆみに	もちいる:V られる	0	x				112	ステロイド軟膏をやめた為か一時
4048	かゆみ	158	3	1	ステロイド(かゆみ止む)を土曜から	のむ:V	1	x	*			118	非ステロイド性抗炎症剤の働きで
4049	かゆみ	19	3	1	炎症・かゆみは(ステロイド剤)が	抑える:V	2				100,101,102,103	1	●小学校2年生の女の子ですが
4052	かゆみ	50	2	1	かゆみも	抑える:V	2				100,101,102,103	6	ステロイドじゃないから安心という
4056	かゆみ	116	2	1	かゆみを ステロイドで	抑える:V	2	x			100,101,102,103	15	症状が悪化しかゆみがつよくなっ
4058	かゆみ	123	2	1	かゆみなどを	抑える:V	2				100,101,102,103	17	商品特徴「デリケアクリーム」は、
4060	かゆみ	144	16	1	かゆみ・かぶれを	鎮める:V	2				100,101,102,103	28	...494円(税込送料別)ぶりがえす
4062	かゆみ	48	2	1	「かゆみ」を	取り除く:V	2				100,101,102,103	58	/しかし脱ステも、ただステロイドを
4063	かゆみ	98	2	1	かゆみが	弱いなる:V	2				100,101,102,103	59	ステロイドを塗ったり、かゆみ止め
4070	かゆみ	17	10	1	かゆみ・かぶれを	治療する:V	2				100,101,102,103	61	ステロイド剤は、かゆみを止める。
4071	かゆみ	141	2	1	かゆみが	治る:V	2				100,101,102,103	64	先レスにも何人か お書きになっ
4076	かゆみ	112	5	1	かゆみは	治まる:V	2				100,101,102,103	65	やはり完全にステロイドを辞める
4079	かゆみ	61	7	1	かゆみを	止める:V	2				100,101,102,103	72	アレルギー疾患の方はご存じかと
4082	かゆみ	93	2	1	かゆみは ステロイド剤で	軽減する:V	2				100,101,102,103	79	なんといっても、ステロイドなしで
4088	かゆみ	140	2	1	かゆみを	軽減する:V	2				100,101,102,103	83	こっちは非ステロイドの抗炎症剤
4091	かゆみ	39	2	1	かゆみを	長引く:V せるない	2				100,101,102,103	33	しかし、最も効果的なかゆみ止め
4093	かゆみ	150	2	1	かゆみが	長びく:V ない	2				100,101,102,103	36	ステロイド様作用があり、抗炎症
4096	かゆみ	119	3	1	かゆみが	なくなる:V	2				100,101,102,103	119	/ステロイド剤を使うとかゆみがな

単純命題レベルの言論間の関係解析

単純命題レベルの言論間関係解析

(述語項構造＋否定、受け身、使役の助動詞)

言論マップ生成システム

1. クエリから関連文書検索
2. 論点リストの獲得
3. 関連文書集合から単純命題の抽出
4. 否定表現の抽出
5. 取り立て助詞(は、も)を格助詞に変換
6. 同一命題をグループ化
7. 項構造を包含する単純命題をグループ化
8. 格の交代関係にある項を持つ単純命題をグループ化
9. 事象間関係知識を用いて命題のクラスタ化
10. 受け身、使役の修正
11. クラスタ間に反義関係を認定
12. 不必要な言論クラスタの除去

評価実験

コーパス中で反義、類義関係が付与された命題を
事象間関係知識をもちいて関係を再現

認識すべき論理的関係

- ・ 反義および同義・類義

対象データ

- ・ 喫煙(10論点、980命題)
[肺がん、害、ニコチン、健康、危険性、
リスク、マナー、影響、受動喫煙、禁煙]
- ・ ステロイド(5論点、1300命題)
[医師、皮膚、かゆみ、ステロイド剤、ステロイド軟膏]

実験結果

トピック	論点数	Recall (sym)	Precision (sym)	Recall (ant)	Precision (ant)
喫煙	10	0.511 (501/980)	0.793 (501/631)	0.126 (20/158)	0.451 (14/31)
ステロイド	5	0.632 (822/1300)	0.801 (822/1026)	0.259 (14/54)	0.588 (20/34)

関係認識の成功した例

受動喫煙でもリスクは高まる ⇔ 受動喫煙もリスクを上昇させる

喫煙者に禁煙を呼び掛ける ⇔ 国民に禁煙を訴える

言論マップ デモ版(随時アップデート)

言論マップ - 村上クラスタリング -

file:///Users/kmurakami/proj/NAIST/StMap/exp0715kitsuen.html

アップル Yahoo! Japan Google マップ YouTube Wikipedia ニュース (468) お役立ち

「喫煙」の言論マップ

[ニコチン][マナー][リスク][健康][分煙][危険性][原因][受動喫煙][問題][喫煙者][女性][害][影響][必要][未成年者][禁煙][肺がん][非喫煙者]

ニコチン

C1(ニコチンを 放出する:V) [3] ⇔ C18(ニコチンを 吸引する:V), C3(ニコチンが 吸収する:V れる), C7(ニコチンが 体内に入る:V), C2(ニコチンを 含む:V), C20(ニコチンを 体内に 喫煙以外で 投与する:V)

ニコチンを 放出する:V [1]
 ニコチンを尿で 出す:V [1]
 ニコチンを身体に 供給する:V [1]

C2(ニコチンを 含む:V) [17] ⇔ C1(ニコチンを 放出する:V), C9(ニコチンが 切れる:V), C15(ニコチンを 除外する:V)

からだにほどニコチンが	入る:V [1]
ニコチンが	含む:V れる [2]
ニコチンがフィルター1つには	含む:V れる [3]
ニコチンが副流煙も	入る:V [2]
ニコチンが喫煙すると母乳には	含む:V れる [2]
ニコチンが子供に	含む:V れる [2]
ニコチンを	含む:V [3]
ニコチンを体内に	入れる:V [1]
一緒に女性ニコチン中毒者必見! 禁煙+美容を手に	入れる:V [1]

C3(ニコチンが 吸収する:V れる) [18] ⇔ C1(ニコチンを 放出する:V)

ニコチンが 吸収する:V れる [6]
 ニコチンが体内に 吸収する:V れる [2]
 ニコチンが血中に 吸収する:V れる [2]
 ニコチンが間は、 吸収する:V れる [2]
 ニコチンを 摂取する:V [4]
 ニコチンを体内に 吸収する:V [1]
 ニコチンを喫煙者は 吸収する:V [1]

C4(ニコチンを ビタミンに 変える:V) [3]

ニコチンがビタミンB3に 変化する:V [1]
 ニコチンがビタミンB3に喫煙すると 変化する:V [1]
 ニコチンをビタミンに 変える:V [1]

C5(ニコチンを 喫煙者は 銘柄でも 摂取する:V) [10]

ガムからニコチンを 摂取する:V [1]
 ニコチンを 取り入れる:V [1]
 ニコチンを体内に 取り入れる:V [2]
 ニコチンを体内に形で 取り入れる:V [1]

考察

命題間の関係をより多く、適切に捉えるために

1. 名詞に関する上位下位関係、類義関係の導入

ニコチン依存症に喫煙者が陥る ⇔ ニコチン依存に陥る^(類義)
ニコチン依存になる ⇔ ニコチン中毒になる^(類義)

2. さらに難しい問題(これからの課題)

アトピー性皮膚炎では症状が治まる

アトピー性皮膚炎は軽減する

アトピー性皮膚炎がよくなる

アトピー性皮膚炎はステロイドに救われる

アトピー性皮膚炎患者を完治に導く^(類義)

⇔ アトピー性皮膚炎を永遠に抑える

まとめと今後の課題

言論マップ生成課題

言論間の論理的関係を抽出、クラスタ化と可視化

- ・ 言論の定義
- ・ 言論間の論理的関係
- ・ 事象間関係知識
- ・ 述語項構造辞書を用いた類義、反義の再現実験
 - 高い精度で関係を再現

今後の課題

- ・ 事象間関係知識(項目の拡張)
- ・ 言論マップ生成システム
 - 関係認識精度の向上
 - 事実性解析によるモダリティ情報の導入