

脳における 文の意味解析機構のモデル

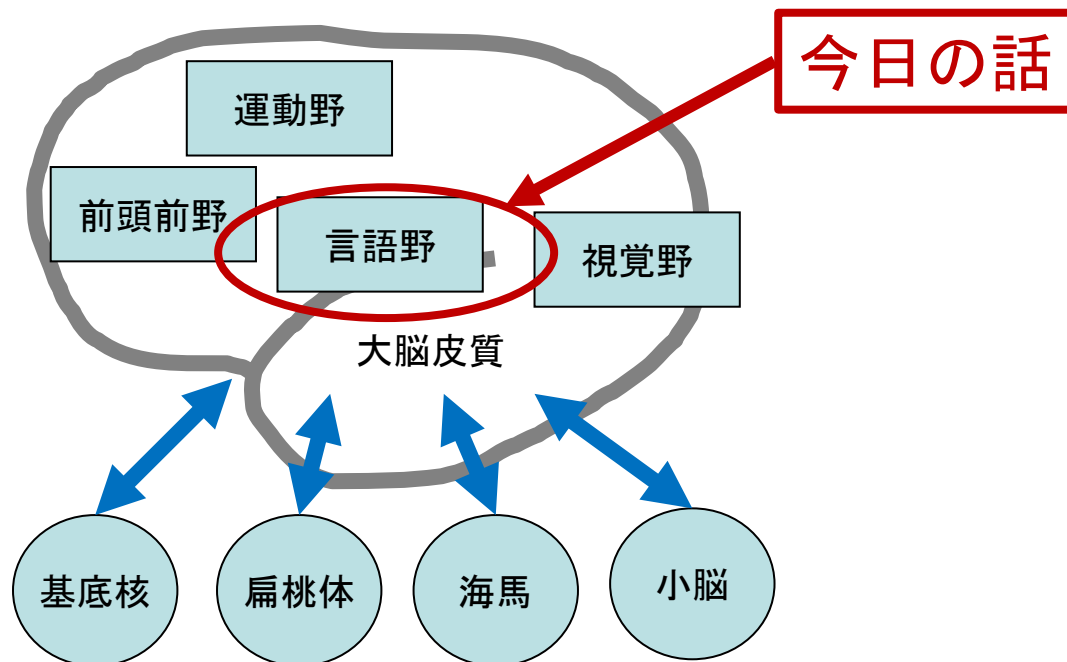
汎用人工知能研究会(SIG-AGI)

2018-03-20

産業技術総合研究所
人工知能研究センター
一杉裕志

私の研究の長期的目標

- 脳を模倣して「人間のような知能を持つ機械」を作る。



脳のリバーエンジニアリング

ここでいう「意味」とは

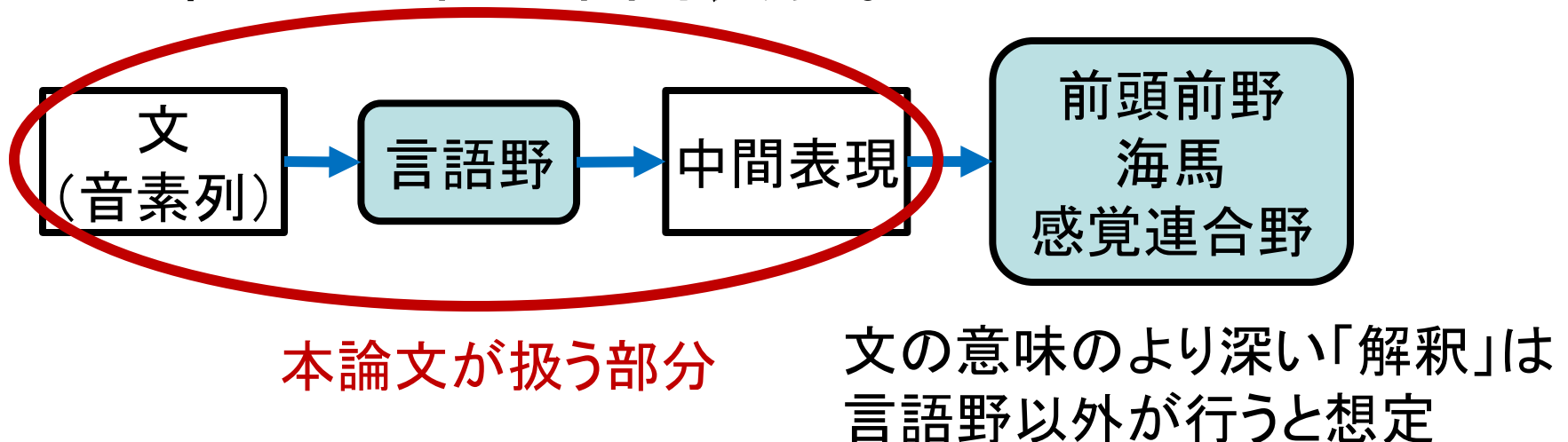
- 文が文字通りに表現する表面的意味。

例: Black cats eat mice.



$\exists e. \exists x, y. \text{black}(e, x) \wedge \text{cats}(e, x) \wedge \text{eat}(e, x, y) \wedge \text{mice}(e, y)$

- いわば意味の中間表現。

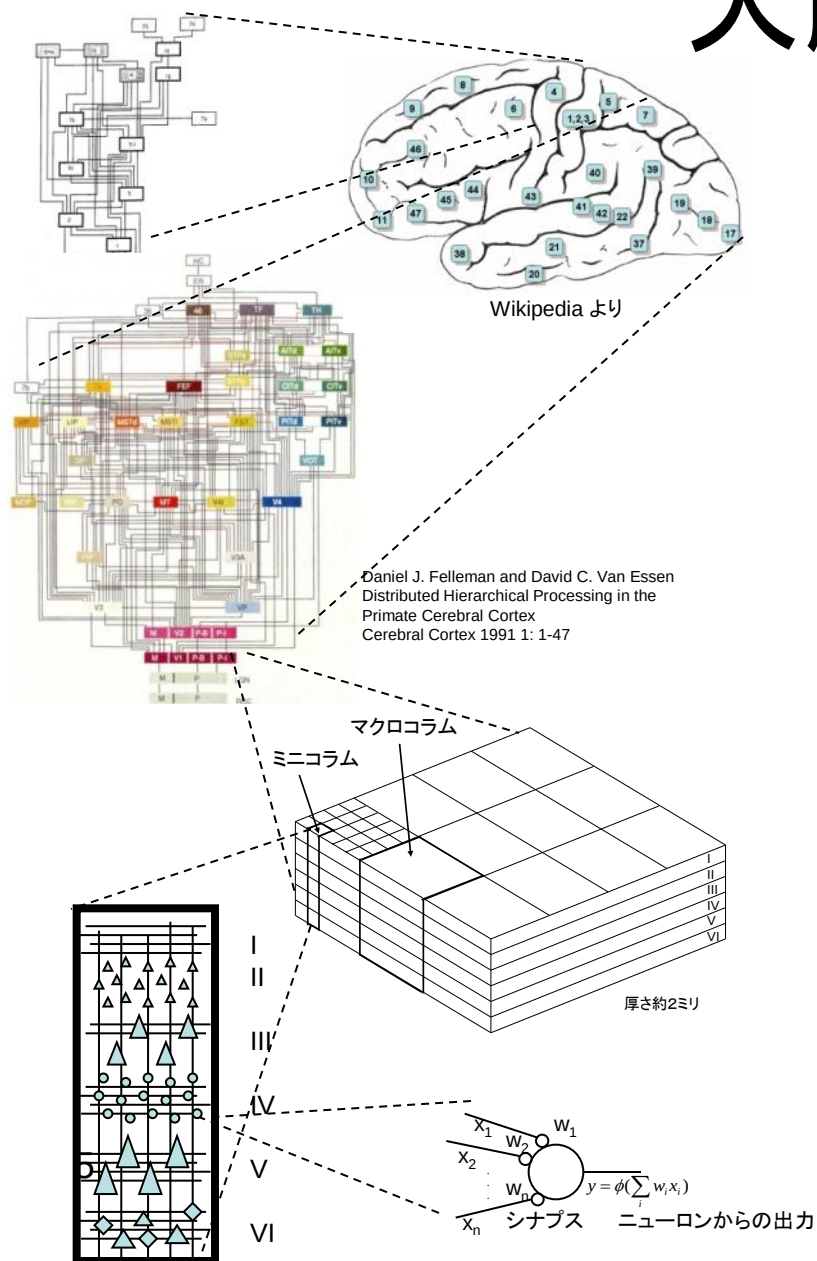


今日の話に必要な背景知識

- 大脳皮質
- ベイジアンネット
- 単一化(Unification)
- 組み合わせ範疇文法
(Combinatory Categorical Grammar, CCG)

大脳皮質

- 脳の様々な高次機能（認識、意思決定、運動制御、思考、推論、言語理解など）が、**たった50個程度**の領野のネットワークで実現されている。



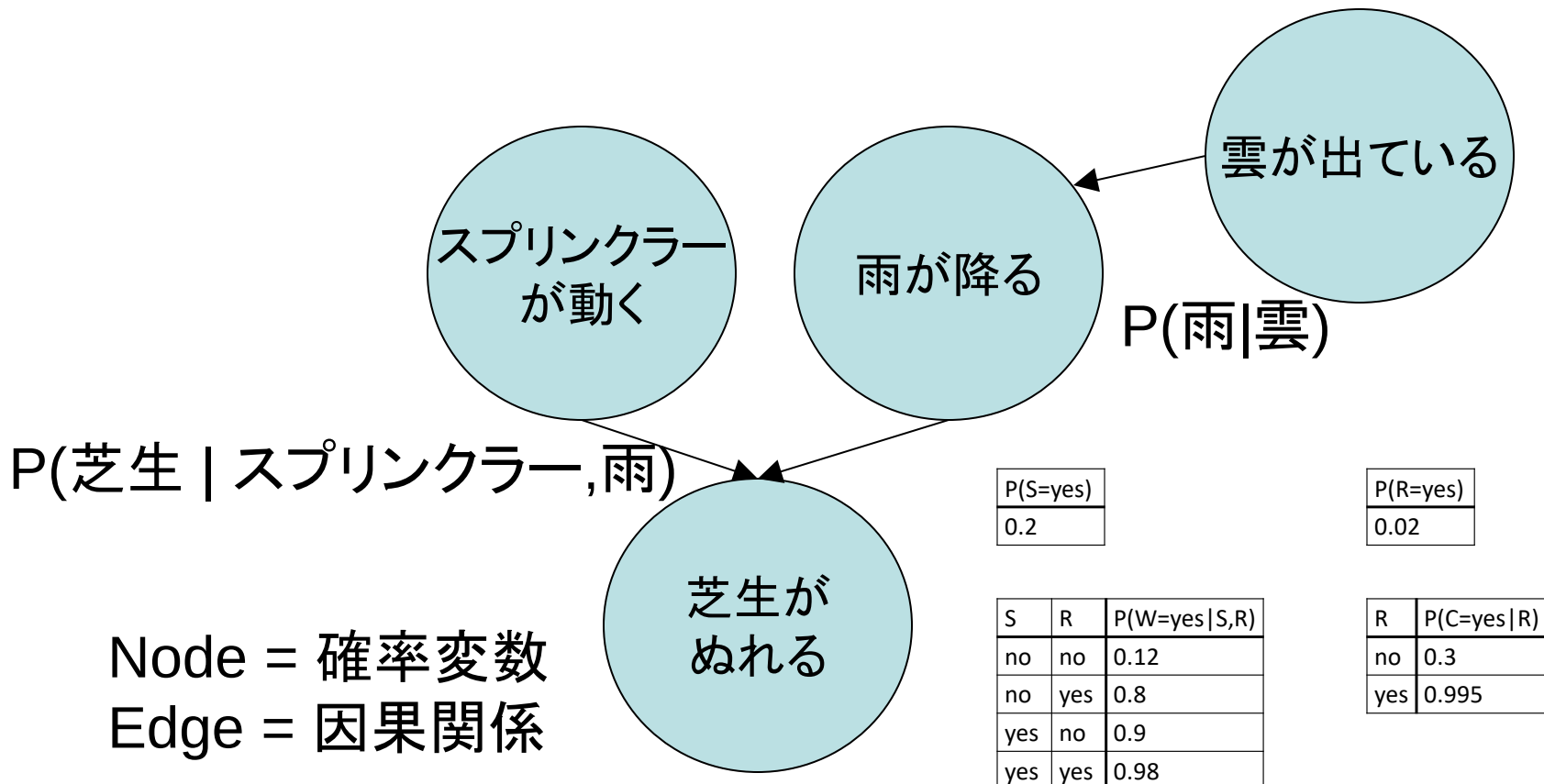
大脳皮質の動作原理解明が
最大の課題

ベイジアンネットを使った 大脳皮質モデル

- 視覚野の機能、運動野の機能、解剖学的構造、電気生理学的現象などを説明
 - [Lee and Mumford 2003]
 - [George and Hawkins 2005]
 - [Rao 2005]
 - [Ichisugi 2007] [Ichisugi 2010] [Ichisugi 2011] [Ichisugi 2012]
 - [Rohrbein, Eggert and Korner 2008]
 - [Hosoya 2009] [Hosoya 2010] [Hosoya 2012]
 - [Litvak and Ullman 2009]
 - [Chikkerur, Serre, Tan and Poggio 2010]
 - [Hasegawa and Hagiwara 2010]
 - [Dura-Bernal, Wennekers, Denham 2012]

ベイジアンネット [Pearl 1988] とは

- 確率変数の間の因果関係をグラフで**効率的**に表す**知識表現**の技術。
- すべてのノードが**入力にも出力にもなり得る**。



単一化文法

- 「**単一化**」という操作を使って構文解析を進めていく文法枠組みの総称
- **理論言語学における文法記述の枠組みの多くが単一化文法**
 - LFG, HPSG, **CCG**, TAG, ...

自然言語の性質を表現する道具として洗練され到達した1つの結果が単一化文法であるという事実は、脳の言語野で何らかの形で単一化が行われていることを強く示唆している。

単一化(unification)とは

- 計算機科学でもよくつかわれる。
 - prolog 言語, 型推論など
- 変数を含む2つの項が等しくなるように、変数の値を割り当てる操作。

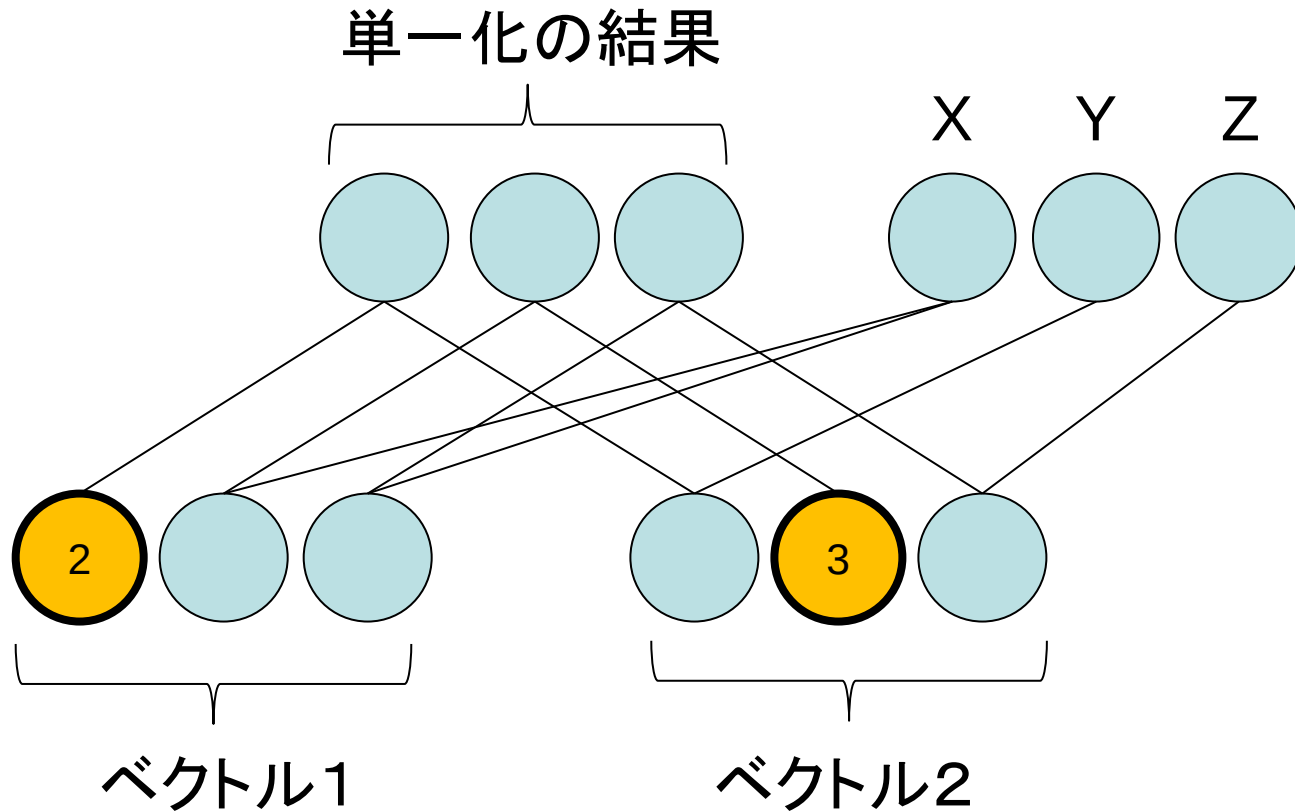
例:

ベクトル 1	ベクトル 2	単一化した結果	変数の割り当て
$(2, X)$	$(Y, 3)$	$(2, 3)$	$X = 3, Y = 2$
$(2, X)$	$(1, 3)$	単一化失敗	-
(X, Y)	(Z, Z)	(X, X)	$X = Y = Z$
$(2, X, X)$	$(Y, 3, Z)$	$(2, 3, 3)$	$X = Z = 3, Y = 2$

単一化を行うベイジアンネット

ベクトル $(2, X, X)$ と $(Y, 3, Z)$ を単一化する回路

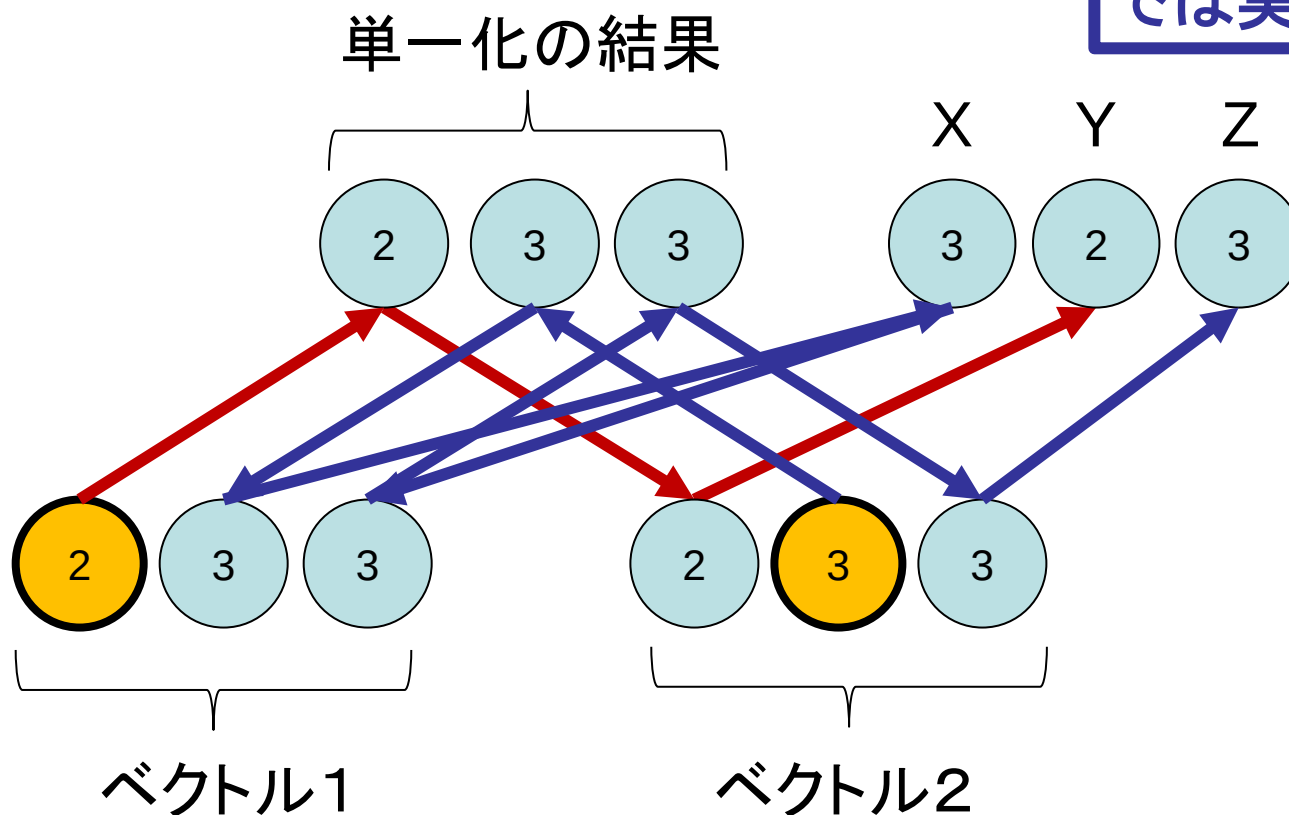
結合された変数は必ず同じ値になるように結合の重みを設定



非観測変数の値の推定結果

ベクトル $(2, X, X)$ と $(Y, 3, Z)$ を単一化する回路
情報が上の層と下の層を何度も往復し伝搬

feedforward NN
では実現しにくい



ここまでのまとめ

- 理論言語学における単一化文法の成功は、大脳皮質の言語野が単一化を行っていることを示唆。
- 単一化はベイジアンネットで簡潔に実現可能。
- 大脳皮質がある種のベイジアンネットである強い状況証拠。

組み合わせ範疇文法(CCG)

組み合わせ範疇文法^[Steedman 2000] (Combinatory Categorical Grammar, CCG)

- 単一化文法の1つ。
- 単純な枠組みで多くの言語現象を説明できる。
- 日本語の文法も記述されている。



チョムスキー階層 (Chomsky 1956)

句構造文法階層	文法	言語	オートマトン	生成規則
タイプ-0	--	帰納的可算	チューリングマシン	制限なし
タイプ-1	文脈依存	文脈依存	線形拘束オートマトン	$\alpha A \beta \rightarrow \alpha \gamma \beta$
タイプ-2	文脈自由	文脈自由	プッシュダウン・オートマトン	$A \rightarrow \gamma$
タイプ-3	正規	正規	有限オートマトン	$A \rightarrow a$ および $A \rightarrow aB$ または $A \rightarrow Ba$



複雑な枠組み

単純な枠組み

<https://ja.wikipedia.org/wiki/チョムスキー階層>

- ・ タイプ1と2の中間(タイプ 1.5)にある文法フレームワークが複数提案されている。

CCG, TAG, HG, LIG

弱文脈依存言語 [Joshi 1985]

文法的な単語列と 非文法的な単語列の例

a man who I think that Keats likes

* a man who I think that likes Keats

Give a teacher an apple and policeman a flower.

* Policeman a flower and give a teacher an apple.

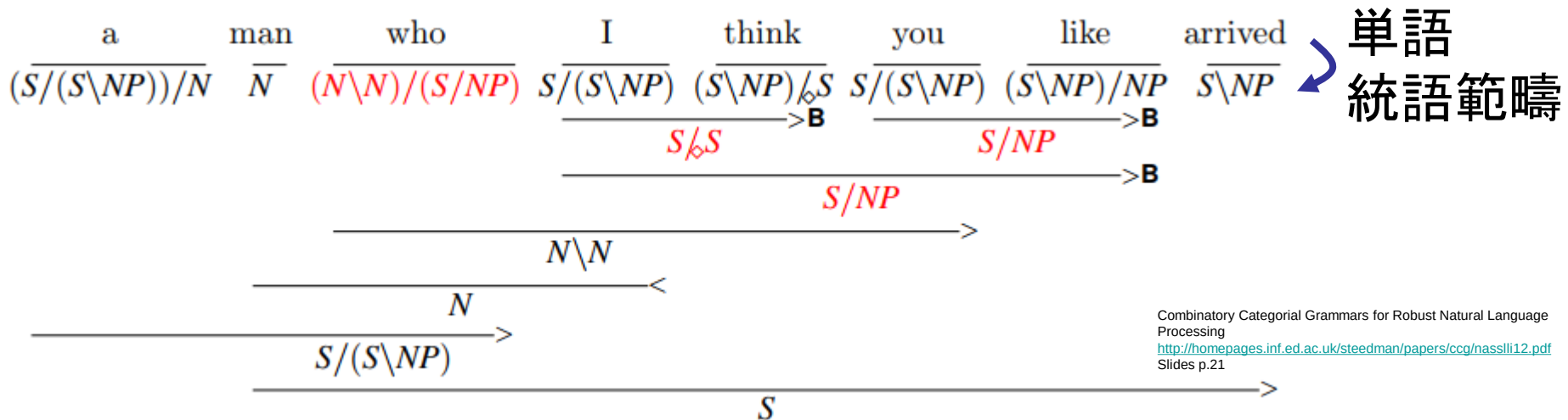
* は非文法的であることを示す印。

CCG における 構文解析の流れ

1. まず辞書を使って各単語を統語範疇(≡非終端記号)に変換
2. 2つの統語範疇に推論規則を適用していき統語範疇 S を導出

$$\begin{array}{ll}
 > \frac{X/Y \quad Y}{X} & < \frac{Y \quad X/Y}{X} \\
 > B \frac{X/Y \quad Y/Z}{X/Z} & < B \frac{Y/Z \quad X/Y}{X/Z} \\
 > T \frac{X}{T/(T/X)} & < T \frac{X}{T/(T/X)}
 \end{array}$$

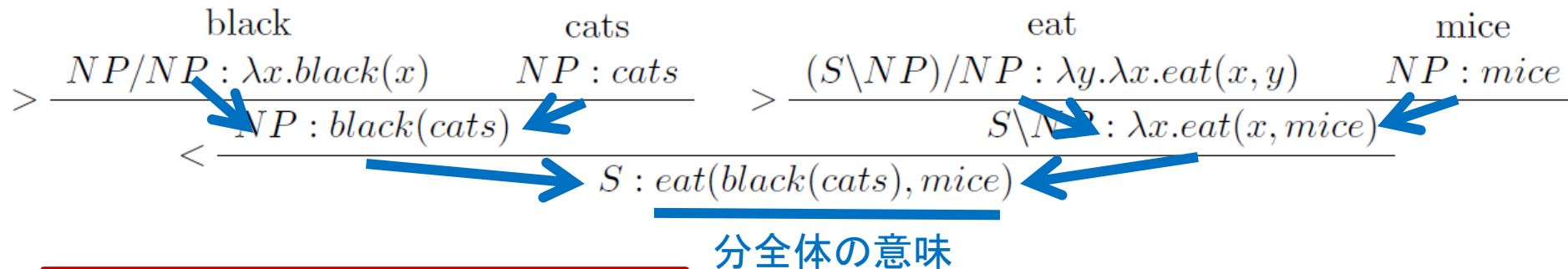
少数の推論規則(≡生成規則)



CCGでの意味解析

$$\begin{array}{ll}
 > \frac{X/Y : \underline{f} \quad Y : \underline{a}}{X : \underline{fa}} & < \frac{Y : \underline{a} \quad X \backslash Y : \underline{f}}{X : \underline{fa}} \\
 > B \frac{X/Y : \underline{f} \quad Y/Z : \underline{g}}{X/Z : \underline{\lambda x.f(gx)}} & < B \frac{Y \backslash Z : \underline{g} \quad X \backslash Y : \underline{f}}{X \backslash Z : \underline{\lambda x.f(gx)}}
 \end{array}$$

構文木に沿って関数適用・関数合成を行っていく。



脳のモデルとしての問題点

- ・ **ラムダ計算は強力すぎる**: 言語野の情報処理の特徴が浮き彫りにならない。
- ・ **可変長データを扱っている**: 神経回路で扱いにくい。

提案モデル

今回の提案モデルの制限

- 単純化した英語の文法に従う文を扱う。
 - 従位接続詞(if など)を使った複文または関係代名詞 which を高々1度だけ使った文を扱う。
- 文の長さ・統語範疇の長さなど、すべて有限とする。
- 文を構成する単語を一度に与える。
 - インクリメンタルな処理は今回は扱わない。
- ある程度複雑な文は言語野以外の領域も使って処理すると考え、今回は扱わない。

提案モデルの特徴： 階層アドレス表現

- 文の意味を、木構造をした論理式ではなく、アドレスとそこに書き込む値の組の集合というフラットな構造で表現。
- アドレスごとに書き込まれた値が何を表現するかがあらかじめ固定されている。
- 可変長データを不要にし、神経回路で扱いやすくすることを意図している。

階層アドレス表現

アドレス: (C,R,F)

節 Clause:

$C \in \{\text{sconj}, c1, c2\}$

意味役割 Semantic Role:

$R \in \{\text{action}, \text{agent}, \text{patient}, \dots\}$

特徴 Feature:

$F \in \{\text{entity}, \text{color}, \text{size}, \dots\}$

論理式に相当する情報を
アドレスと意味表現からなる組の集合で
表現

アドレス	意味表現
$(\text{sconj}, -, -)$	if
$(c1, \text{agent}, \text{size})$	big
$(c1, \text{agent}, \text{color})$	*
$(c1, \text{agent}, \text{entity})$	dogs
$(c1, \text{modality}, -)$	*
$(c1, \text{action}, -)$	chase
$(c1, \text{patient}, \text{size})$	small
$(c1, \text{patient}, \text{color})$	*
$(c1, \text{patient}, \text{entity})$	mice
$(c2, \text{agent}, \text{size})$	*
$(c2, \text{agent}, \text{color})$	black
$(c2, \text{agent}, \text{entity})$	cats
$(c2, \text{modality}, -)$	may
$(c2, \text{action}, -)$	eat
$(c2, \text{patient}, \text{size})$	*
$(c2, \text{patient}, \text{color})$	*
$(c2, \text{patient}, \text{entity})$	mice

if small mice areChasedBy big dogs
black cats may eat mice

解析の手順

1. 辞書を使って単語列を 統語範疇の列に変換

単語	統語範疇	アドレス	意味表現
'if'	$(S_{c1}/S_{c2})/S_{c1}$	$(sconj, -, -)$	if
'black'	$NP_{C,R}/NP_{C,R}$	$(C, R, color)$	black
'big'	$NP_{C,R}/NP_{C,R}$	$(C, R, size)$	big
'cats'	$NP_{C,R}$	$(C, R, entity)$	cats
'eat'	$(SC \setminus NP_{C,agent})/NP_{C,patient}$	$(C, action, -)$	eat
'areEatenBy'	$(SC \setminus NP_{C,patient})/NP_{C,agent}$	$(C, action, -)$	eat
'may'	$(SC \setminus NP_{C,R})/(SC \setminus NP_{C,R})$	$(C, modality, -)$	may
'which'	$(NP_{c1,R1} \setminus NP_{c1,R1})/(S_{c2} \setminus NP_{c2,R2})$	$(c2, R2, entity)$	$(c1, R1, entity)$
'which'	$(NP_{c1,R1} \setminus NP_{c1,R1})/(S_{c2} \setminus NP_{c2,R2})$	$(c2, R2, entity)$	$(c1, R1, entity)$

$$\begin{array}{c}
 \text{black} \qquad \qquad \text{cats} \qquad \qquad \text{eat} \qquad \qquad \text{mice} \\
 > \frac{NP_{C1,R1}/NP_{C1,R1} \quad NP_{C2,R2}}{NP_{C1,R1}} > \frac{(SC_3 \setminus NP_{C3,agent})/NP_{C3,patient} \quad NP_{C4,R4}}{SC_3 \setminus NP_{C3,agent}} \\
 < \frac{\quad}{SC_3}
 \end{array}$$

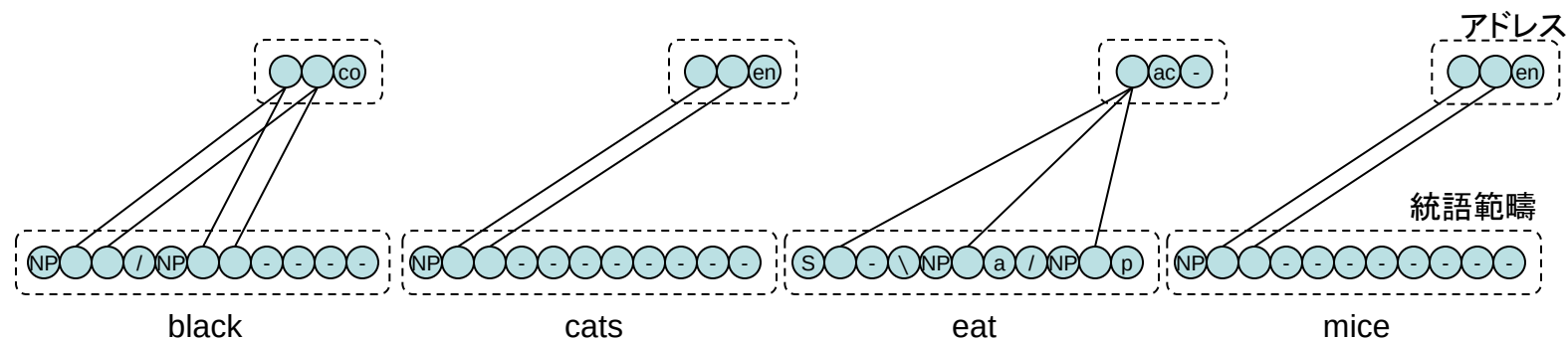
$> \frac{X/Y \quad Y}{X}$	$< \frac{Y \quad X \setminus Y}{X}$
$> B \frac{X/Y \quad Y/Z}{X/Z}$	$< B \frac{Y \setminus Z \quad X \setminus Y}{X \setminus Z}$
$> T \frac{X}{T/(T \setminus X)}$	$< T \frac{X}{T \setminus (T/X)}$

2. 統語範疇の列に 推論規則を適用して 文 S を導出

アドレス	意味表現
$(c1, agent, color)$	black
$(c1, agent, entity)$	cats
$(c1, action, -)$	eat
$(c1, patient, entity)$	mice

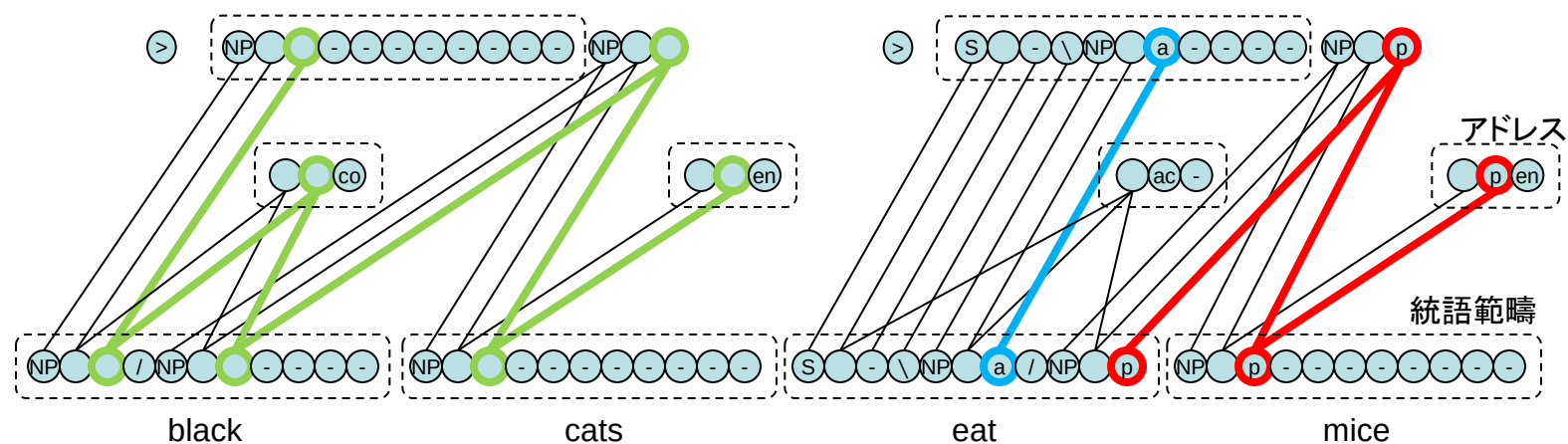
3. 個々の単語の意味を書き込む アドレスが確定

black	cats	eat	mice	単語
$NP_{C_1, R_1} / NP_{C_1, R_1}$	NP_{C_2, R_2}	$(S_{C_3} \setminus NP_{C_3, agent}) / NP_{C_3, patient}$	NP_{C_4, R_4}	統語範疇
$(C_1, R_1, color)$	$(C_2, R_2, entity)$	$(C_3, action, -)$	$(C_4, R_4, entity)$	アドレス



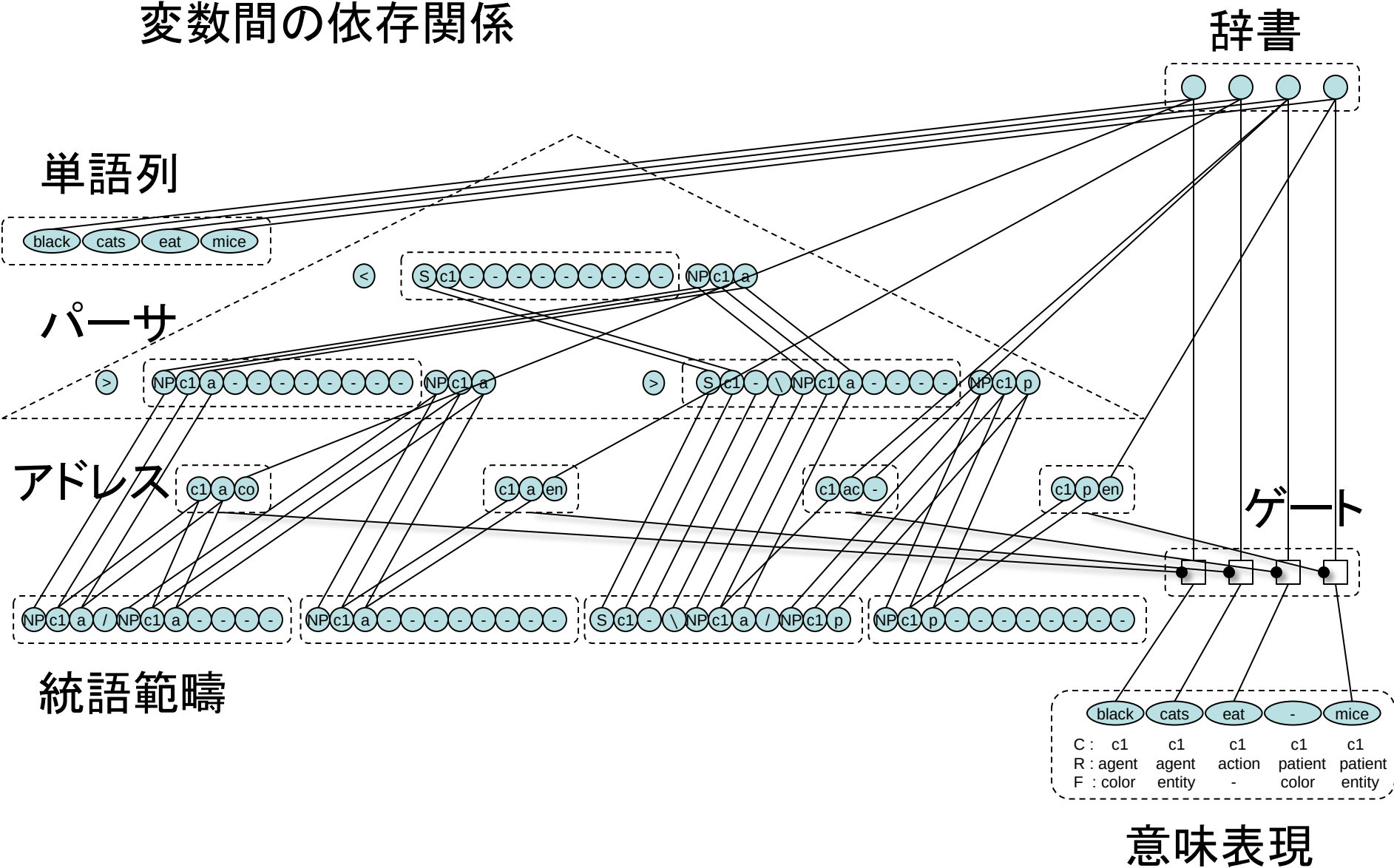
1. 辞書を検索し、単語ごとに、このような共有構造をもったアドレスと統語範疇を得る。

$$\begin{array}{c}
 \text{black} \qquad \qquad \text{cats} \\
 \frac{NP_{C_1,R_1}/NP_{C_1,R_1} \quad NP_{C_2,R_2}}{NP_{C_1,R_1}} > \frac{\text{eat} \qquad \qquad \text{mice}}{(S_{C_3} \setminus NP_{C_3,agent})/NP_{C_3,patient} \quad NP_{C_4,R_4}} > \frac{\quad}{S_{C_3} \setminus NP_{C_3,agent}}
 \end{array}$$



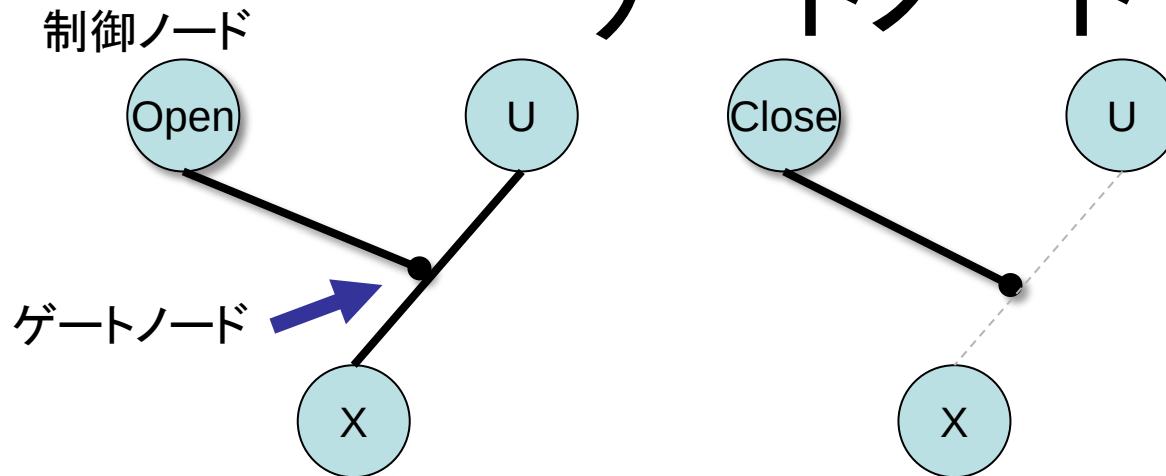
2. 推論規則を適用して統語範疇を統合していく。
 その際、統語素性が単一化されていく。

black cats eat mice を解析し終えた状態における 変数間の依存関係



大脳皮質モデル BESOM Ver.4 での実現方針

ゲートノード



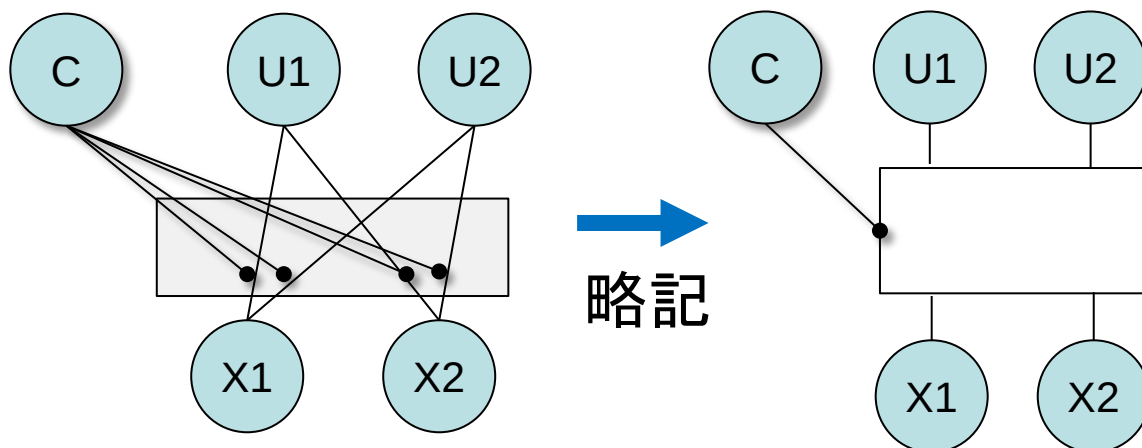
制御ノードの値が
Open なら U と X が結合、
Close なら切断

ゲートノードの条件付確率表

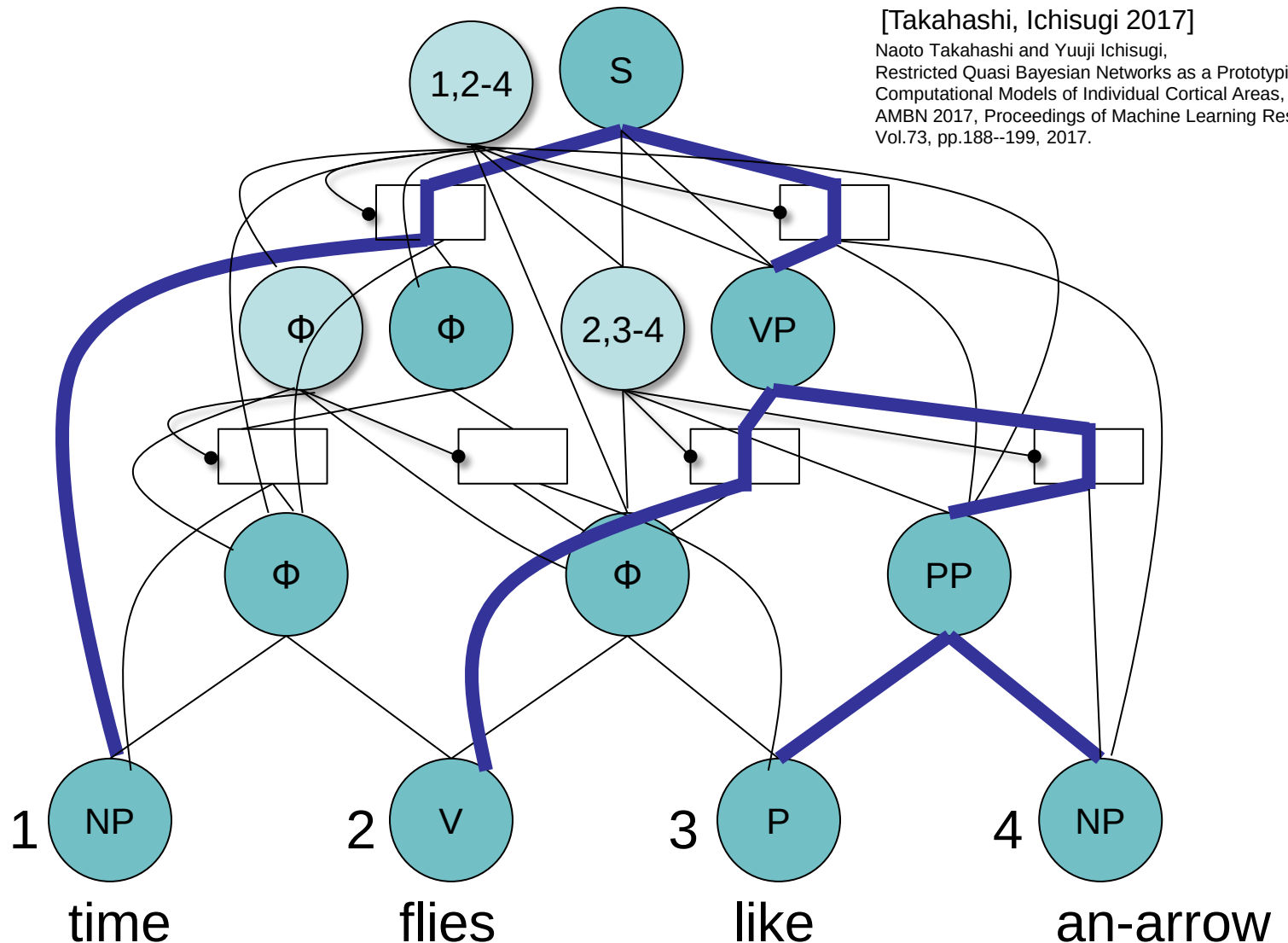
$$\begin{aligned} P(G = 0|c, u) &= 1 - (1 - w_u)^{-u} (1 - w_c)^c \\ P(G = 1|c, u) &= (1 - w_u)^{-u} (1 - w_c)^c \end{aligned}$$

- ・ 論理回路のような感覚で生成モデルを設計できる。
- ・ あくまでベイジアンネットなので、出力(下流の値)から入力(上流の値)の推定もできる。

ゲートノードの行列



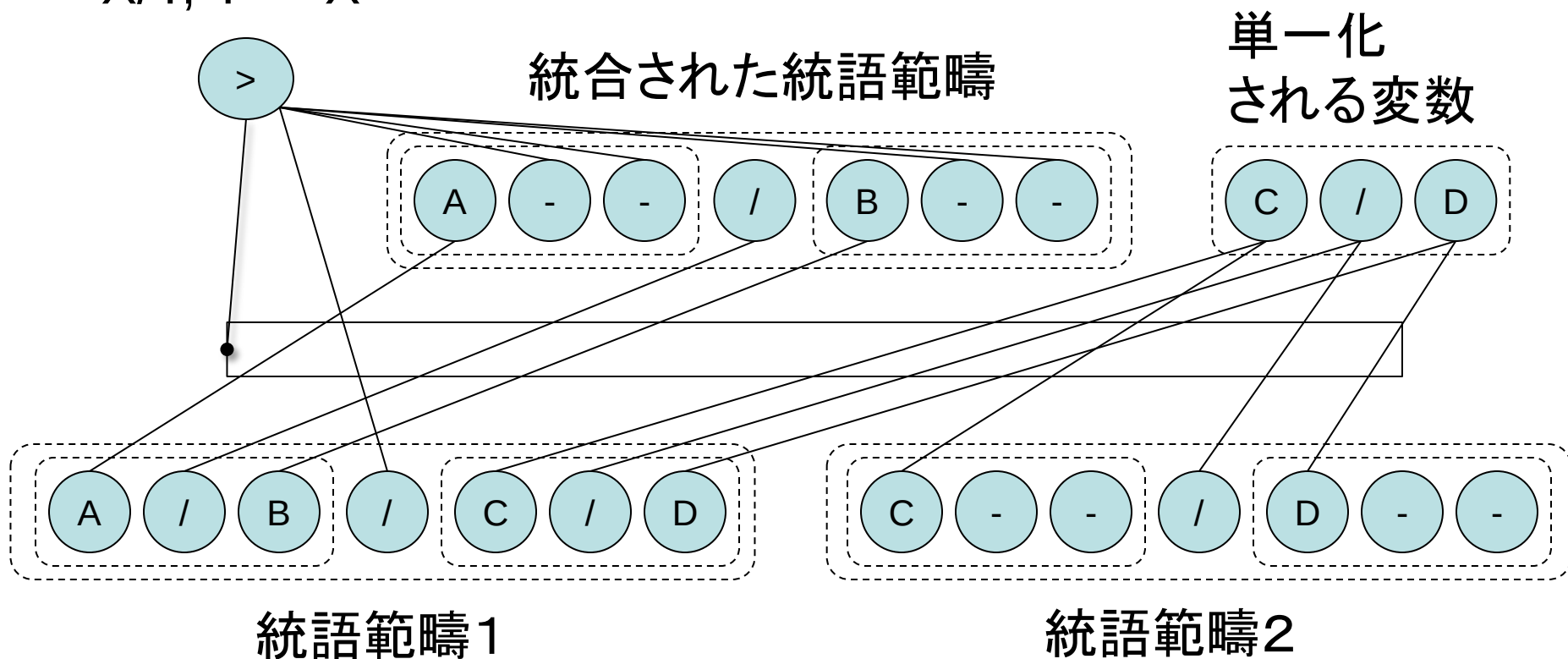
文脈自由文法のチャートパーサ



CCG の推論規則を適用する回路

推論規則

"X/Y, Y $\rightarrow$ X"



(A/B)/(C/D), C/D $\rightarrow$ A/B

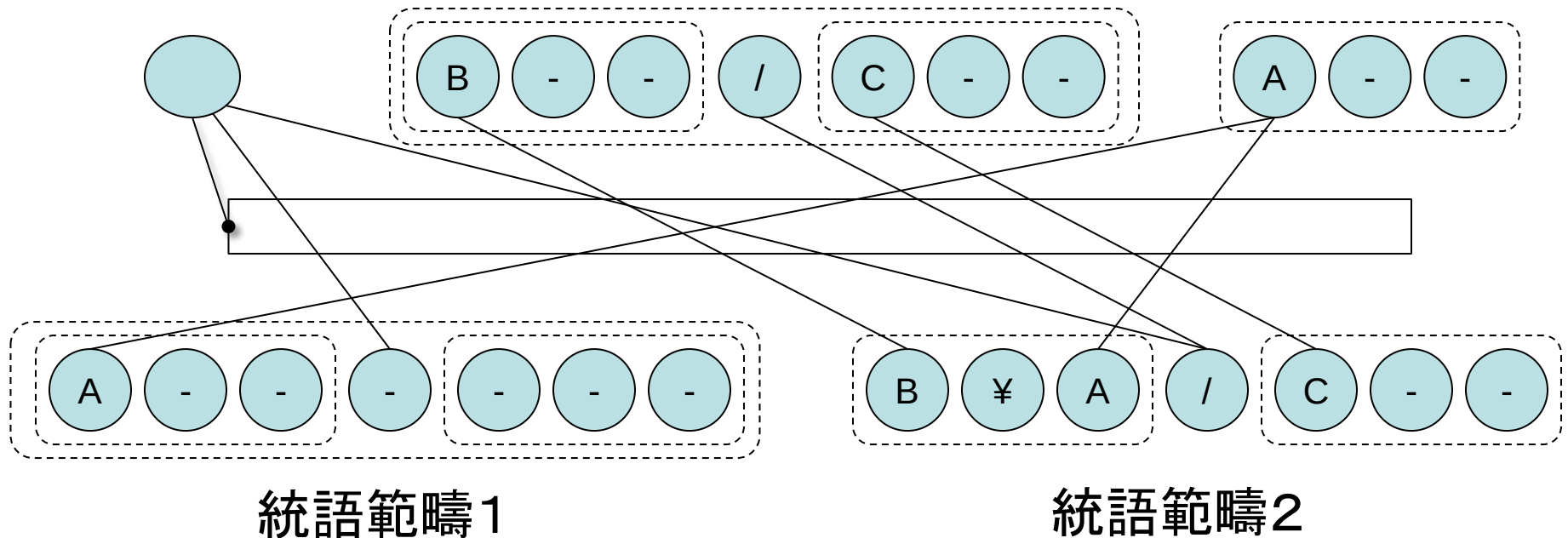
型繰り上げ規則は、他の推論規則を組み合わせて表現

$(>T) \quad X \rightarrow T / (T \neq X)$
 $(>B) \quad T / (T \neq X), (Y \neq X) / Z \rightarrow Y / Z$
 $\downarrow$
 $(>T, >B) \quad X, (Y \neq X) / Z \rightarrow Y / Z$

$T/X, T \neq X$ は
 おそらく有限個
 [Steedman 2000]
 p.44

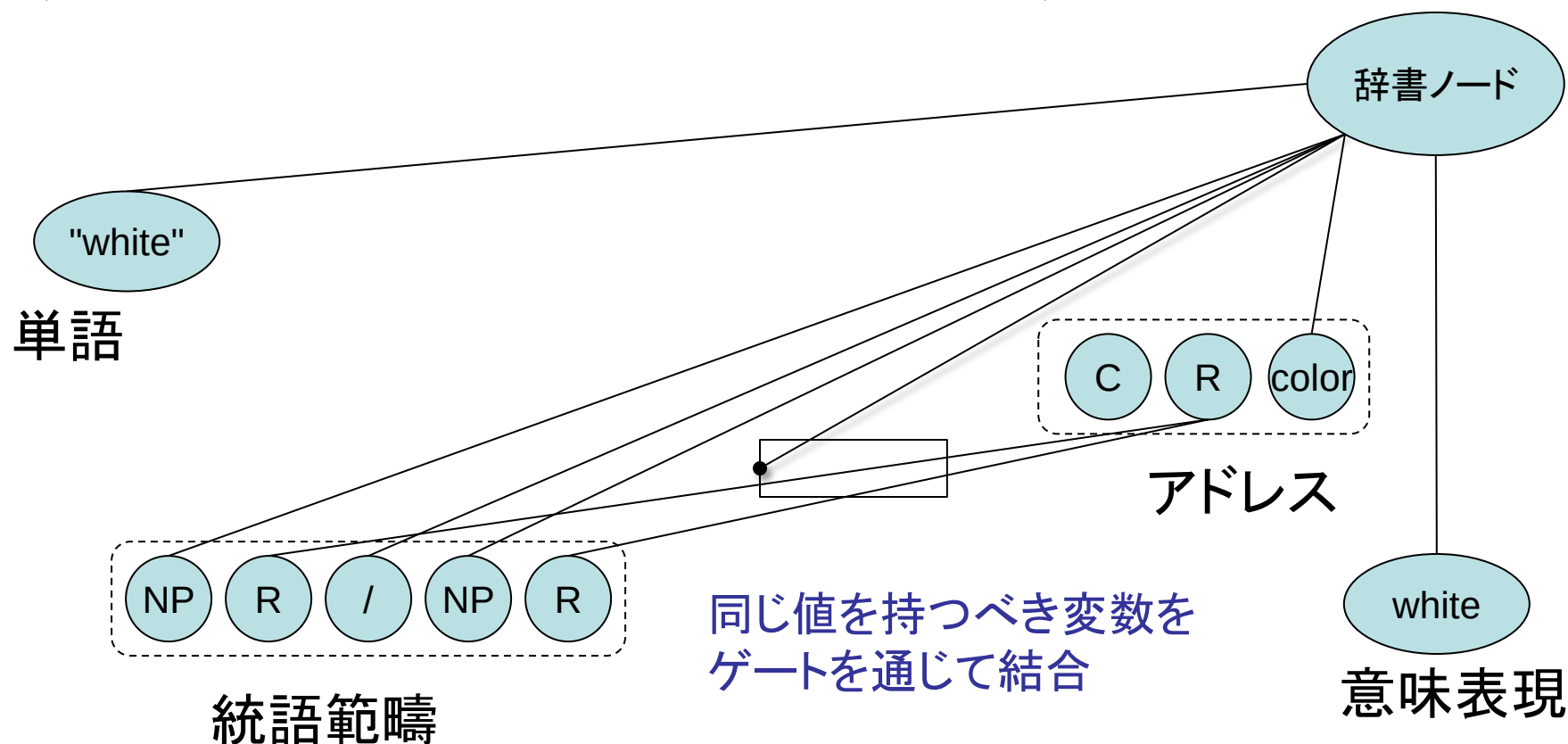
推論規則

" $X, (Y \neq X) / Z \rightarrow Y / Z$ "



語彙項目の表現方法

例: (単語, 統語範疇, アドレス, 意味表現)
= ("white", NP_R/NP_R , $(C, R, color)$, white)



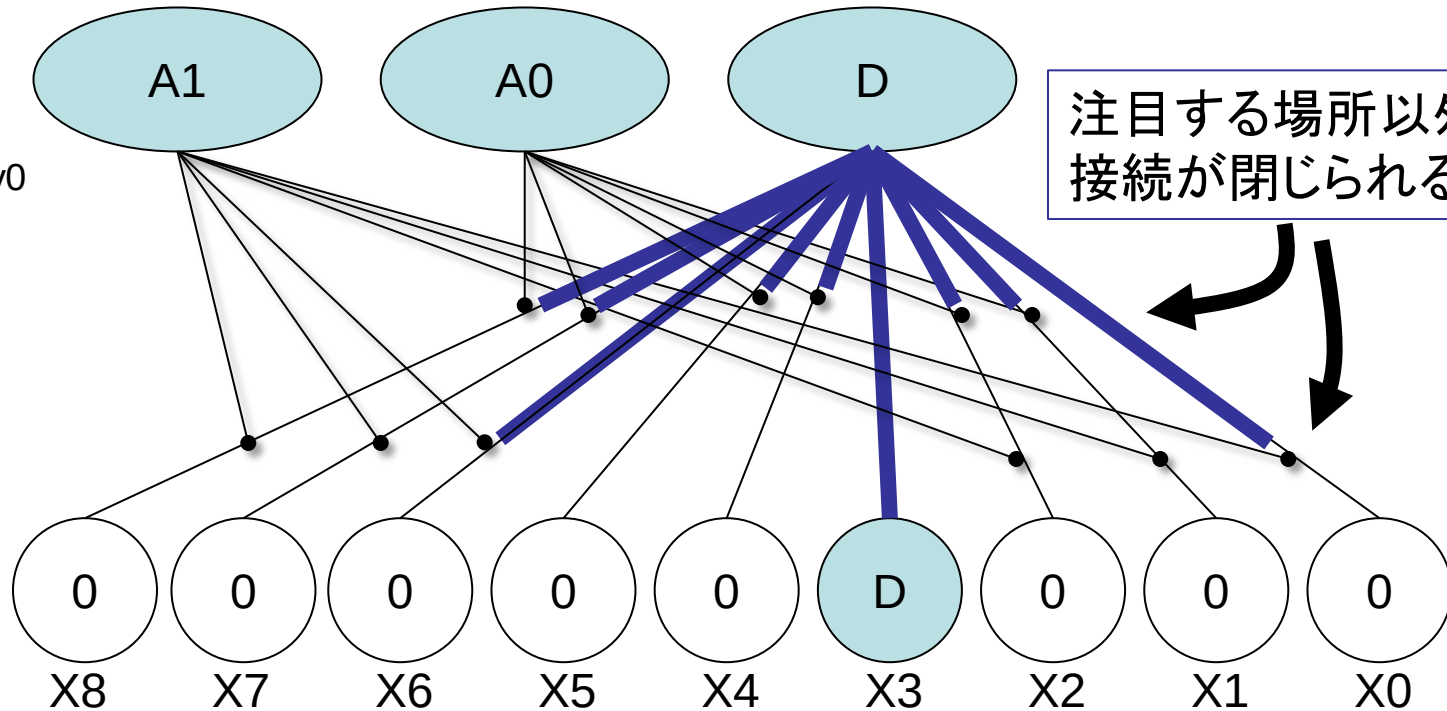
階層的アドレスとデータ

アドレス
3進数2ケタめ

アドレス
3進数1ケタめ

データ

A1=v1, A0=v0



注目する場所以外への
接続が閉じられる。

A1	G8	G7	G6	G5	G4	G3	G2	G1	G0
v0	I	I	I	I	I	I	O	O	O
v1	I	I	I	O	O	O	I	I	I
v2	O	O	O	I	I	I	I	I	I

A0	G8	G7	G6	G5	G4	G3	G2	G1	G0
v0	I	I	O	I	I	O	I	I	O
v1	I	O	I	I	O	I	I	O	I
v2	O	I	I	O	I	I	O	I	I

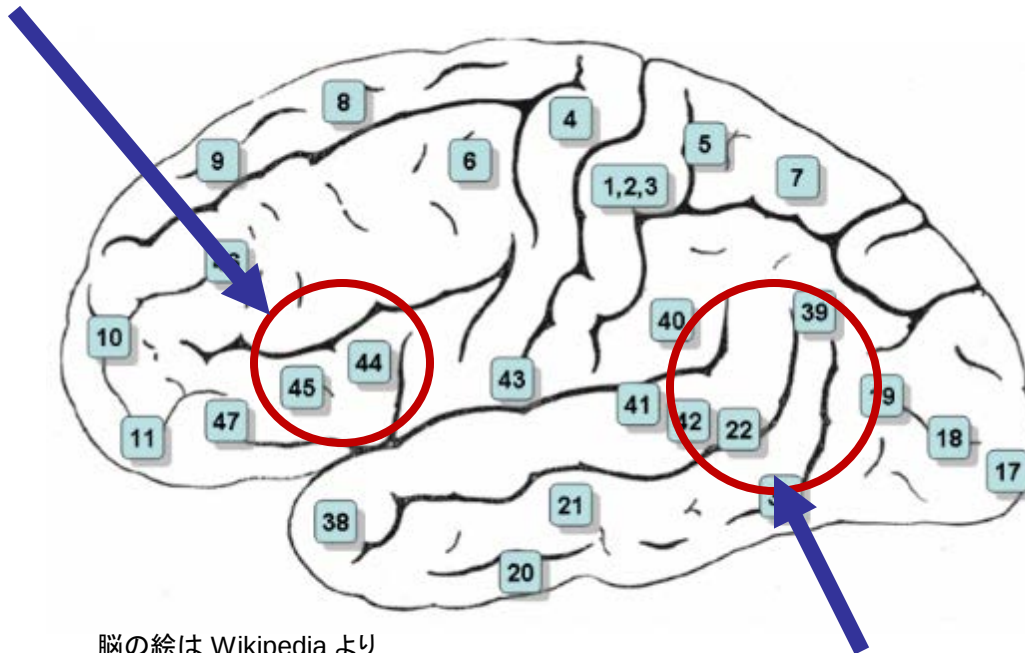
大脳皮質モデルに関して 新たに得られた仮説

- 仮説1：
2層間の結合は行列(2階テンソル)では不十分だが3階テンソルなら十分である。
- 仮説2：
学習時に結合の重みが0か1に近づくような prior が設定されている。

提案モデルと脳との対応、 失語症の発話の再現

言語野

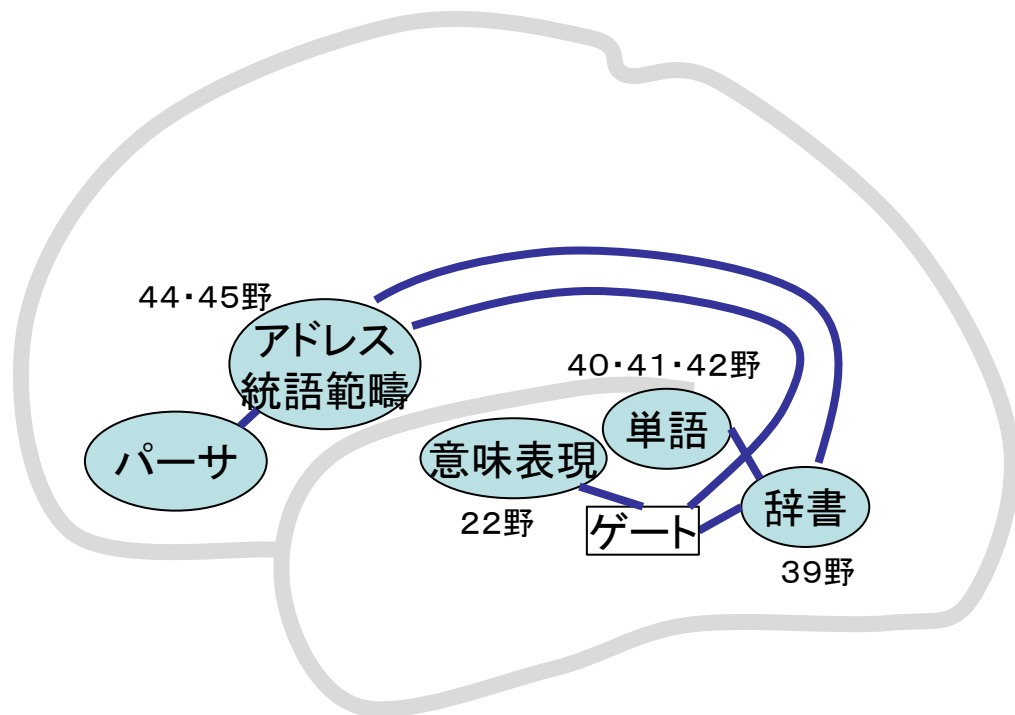
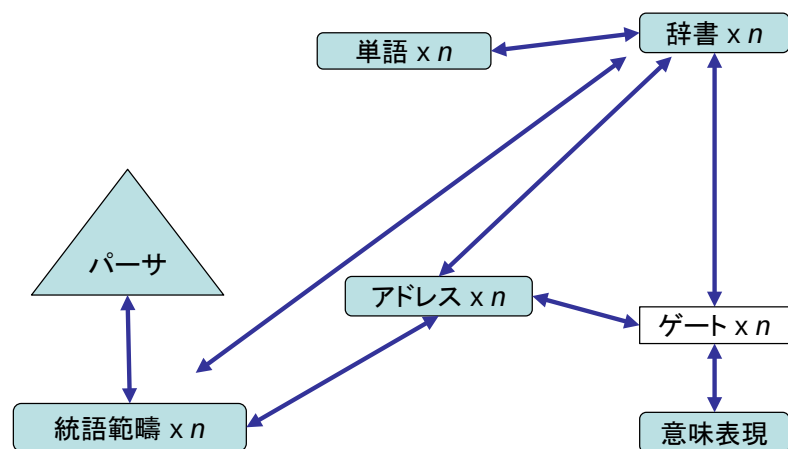
ブローカー野(44・45野)



脳の絵は Wikipedia より

ウェルニッケ野
(一次聴覚野と角回の近く)

脳との対応の1つの可能性



注:

- ・ パーサと意味表現以外のモジュールの位置については要再検討。
- ・ 運動野はこの図には入っていない。BA39 と接続する？
- ・ 縁上回 BA40 は単語列のバッファだと思うが他の部位との関係はまだよくわからない。
- ・ 言語野にも腹側経路があると考えられつつあるが、それはこの図には入っていない。
- ・ ゲートは特定の領野に対応するのではなく樹状突起上の抑制性結合であると考えている。

ブローカー野とウェルニッケ野

• ブローカー野：文法処理に関与

– ブローカー失語：

- 発話：電文体（機能語が省略される）
- 理解：文法に強く依存した文の理解が困難

• ウェルニッケ野：言語音と概念の連合に関与

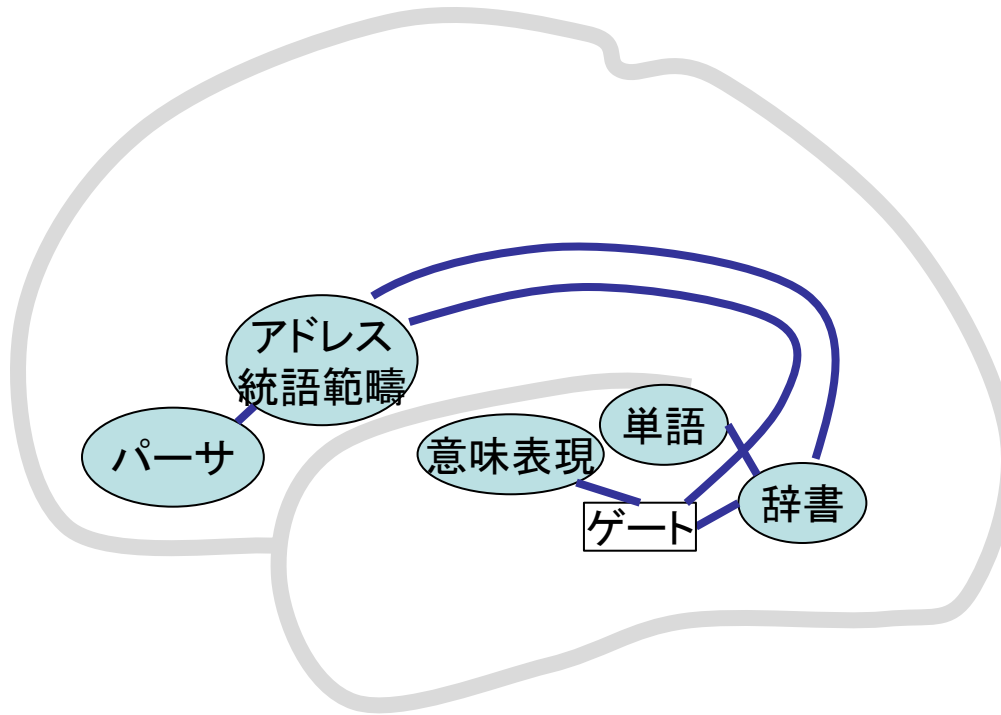
– ウェルニッケ失語：

- 発話：流暢に話すが意味のない内容
- 理解：障害されている

参考：カンデル神経科学 p.1336-1339、言語処理学事典 p.772-787

脳においては「発話」と「理解」というモジュール分割ではなく
「文法」と「意味」というモジュール分割がなされている

正常な発話



アドレス	意味表現
(c1, agent, color)	black
(c1, agent, entity)	cats
(c1, action, -)	eat
(c1, patient, entity)	mice

正常な文の生成

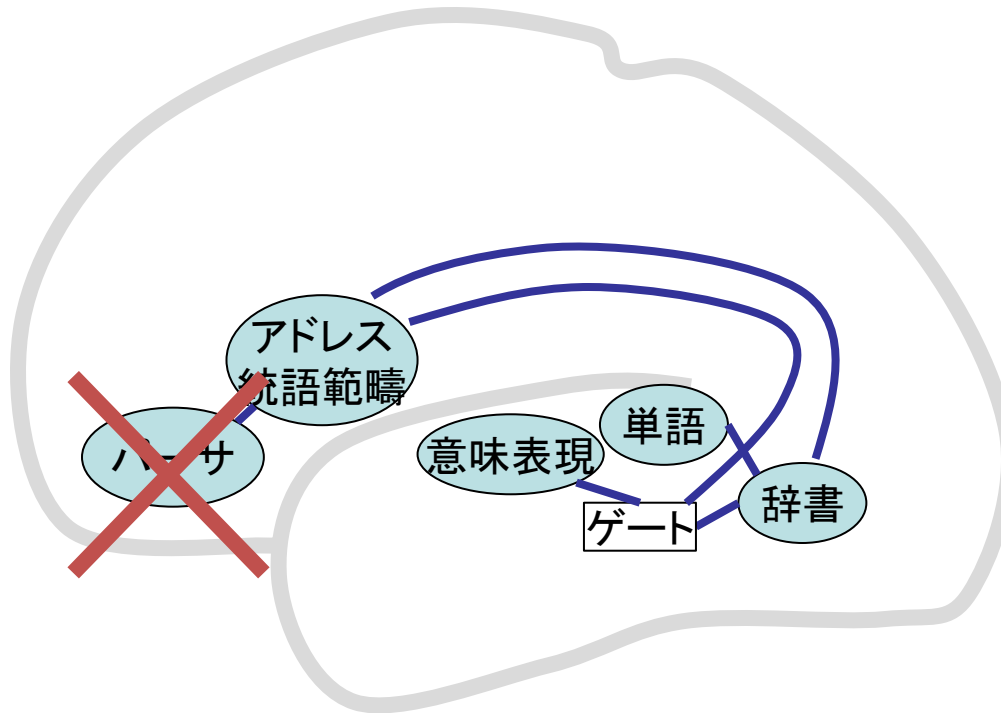
```
?- M=[[[c1,agent,color],black],  
[[c1,agent,entity],cats], [[c1,action,-],eat],  
[[c1,patient,entity],mice]],  
Ws=[W1,W2,W3,W4], maplist(lexicalItem,  
Ws, Cs, As, Ds), parse([], Cs, s(c1)),  
maplist(bind(M), As, Ds), print(Ws), nl, fail.
```

出力(重複を除く):

[black,cats,eat,mice]

[mice,areEatenBy,black,cats]

ブローカー失語と似た発話



アドレス	意味表現
(c1, agent, color)	black
(c1, agent, entity)	cats
(c1, action, -)	eat
(c1, patient, entity)	mice

パーサ部分を取り除いた場合

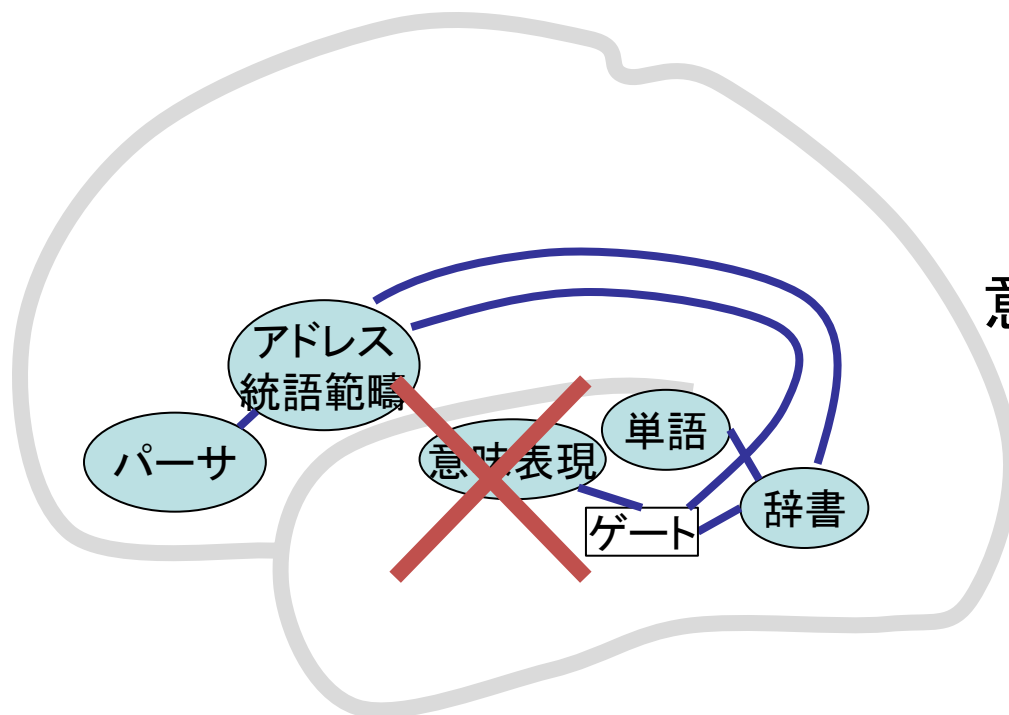
```
?- M=[[c1,agent,color],black],  
[[c1,agent,entity],cats], [[c1,action,-],eat],  
[[c1,patient,entity],mice]],  
Ws=[W1,W2,W3,W4], maplist(lexicalItem, Ws,  
Cs, As, Ds), /* parse([], Cs, s(c1)),*/  
maplist(bind(M), As, Ds), print(Ws), nl, fail.
```

出力:

```
[black,black,black,black]  
[black,black,black,cats]  
[black,black,black,mice]  
[black,black,black,eat]  
[black,black,black,areEatenBy]  
[black,black,cats,black]  
[black,black,cats,cats]  
[black,black,cats,mice]  
...
```

文法的に正しくないが、意味的に正しい単語が選択される。

ウェルニツケ失語と似た発話



意味表現を与えずに文を生成

```
?- /* M=[[c1,agent,color],black],  
    [[c1,agent,entity],cats], [[c1,action,-],eat],  
    [[c1,patient,entity],mice]],*/  
Ws=[W1,W2,W3,W4], maplist(lexicalItem,  
Ws, Cs, As, Ds), parse([], Cs, s(c1)),  
maplist(bind(M), As, Ds), print(Ws), nl, fail.
```

出力(重複を除く)

```
[white,dogs,eat,dogs]  
[white,dogs,eat,cats]  
[white,dogs,eat,mice]  
[white,dogs,chase,dogs]  
[white,dogs,chase,cats]  
[white,dogs,chase,mice]  
[white,dogs,areEatenBy,dogs]  
[white,dogs,areEatenBy,cats]  
...
```

文法的に正しいあらゆる文が生成される。

従来のCCGではこれらの症状は 再現できない

従来のCCGでは文法と意味の情報が混在している。

$$\begin{array}{c}
 \text{black} \qquad \qquad \text{cats} \qquad \qquad \text{eat} \qquad \qquad \text{mice} \\
 > \frac{NP/NP : \lambda x. \underline{\text{black}}(x) \quad NP : \underline{\text{cats}}}{NP : \underline{\text{black}}(\underline{\text{cats}})} \quad > \frac{(S \backslash NP)/NP : \lambda y. \lambda x. \underline{\text{eat}}(x, y) \quad NP : \underline{\text{mice}}}{S \backslash NP : \lambda x. \underline{\text{eat}}(x, \underline{\text{mice}})} \\
 < \frac{NP : \underline{\text{black}}(\underline{\text{cats}})}{S : \underline{\text{eat}}(\underline{\text{black}}(\underline{\text{cats}}), \underline{\text{mice}})}
 \end{array}$$

提案モデルでは、パーサはアドレスだけ进行处理しており、意味の内容そのものは分離されている。

そのため、**文法だけ障害**、**意味だけ障害**、ということが起きうる。

$$\begin{array}{c}
 \text{black} \qquad \qquad \text{cats} \qquad \qquad \text{eat} \qquad \qquad \text{mice} \\
 > \frac{NP_{C_1, R_1}/NP_{C_1, R_1} \quad NP_{C_2, R_2}}{NP_{C_1, R_1}} \quad > \frac{(S_{C_3} \backslash NP_{C_3, agent})/NP_{C_3, patient} \quad NP_{C_4, R_4}}{S_{C_3} \backslash NP_{C_3, agent}} \\
 < \frac{NP_{C_1, R_1}}{S_{C_3}}
 \end{array}$$

まとめ

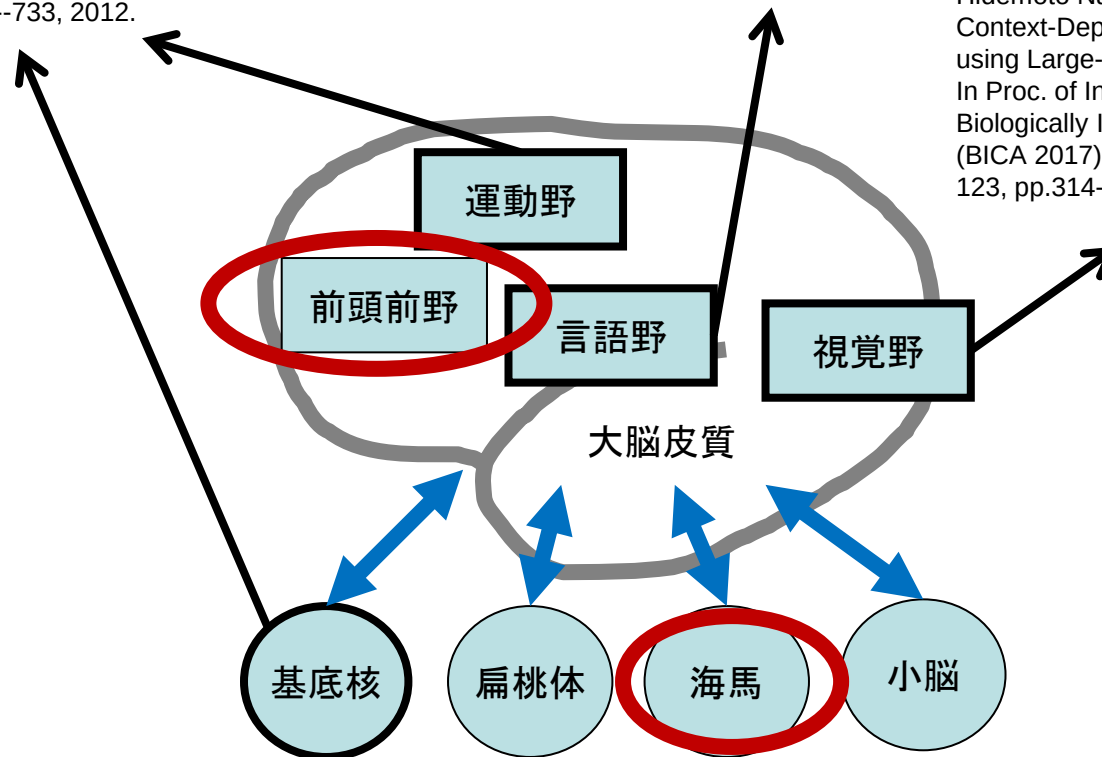
- CCG を修正し、ラムダ計算や可変長データを使わない意味解析機構を提案。
- 簡単な文法を prolog で実装。
 - 大脳皮質モデル BESOM Ver.4 で実現できる見込みが立った。
 - 大脳皮質モデルが持つべき特徴のヒントが得られた。
- 失語症の発話を再現。
- 計算論的神経科学と理論言語学を融合させ、双方を大きく進展させる可能性がある。

今後

Yuuji Ichisugi, A Computational Model of Motor Areas Based on Bayesian Networks and Most Probable Explanations, In Proc. of The International Conference on Artificial Neural Networks (ICANN 2012), Part I, LNCS 7552, pp.726--733, 2012.

今回の論文

Hidemoto Nakada and Yuuji Ichisugi, Context-Dependent Robust Text Recognition using Large-scale Restricted Bayesian Network, In Proc. of International Conference on Biologically Inspired Cognitive Architectures (BICA 2017), Procedia Computer Science, Vol. 123, pp.314--320, 2018.



脳型AGIアーキテクチャの本丸：
前頭前野と海馬を中心とする思考のモデルに取り組みたい