A Formal Model of the Mechanism of Semantic Analysis in the Brain

BICA 2018

2018-08-23

Yuuji Ichisugi and Naoto Takahashi

National Institute of Advanced Industrial Science and Technology (AIST), Japan

Our research goals

- Long term goal : **Human-like intelligence** by imitating the architecture of the whole brain

http://www.irasutoya.com/2015/05/ai.html

- Short term goals:
  - Implement a cerebral cortex model
    - Our working hypothesis :
      The cerebral cortex is a kind of Bayesian network
  - Implement **visual areas, language areas, motor areas, prefrontal areas, etc.** using the cerebral cortex model

Models of cerebral cortex based on Bayesian networks

- Various functions, illusions, neural responses and anatomical structure of the visual cortex were reproduced by Bayesian network models.
  - [Tai Sing Lee and Mumford 2003]
  - [Dileep George and Hawkins 2005]
  - [Rao 2005]
  - [Ichisugi 2007]
  - [Litvak and Ullman 2009]
  - [Chikkerur, Serre, Tan and Poggio 2010]
  - [Hosoya 2012]
  - ...

The cerebral cortex seems to be a huge Bayesian network with layered structure like Deep Neural Networks.

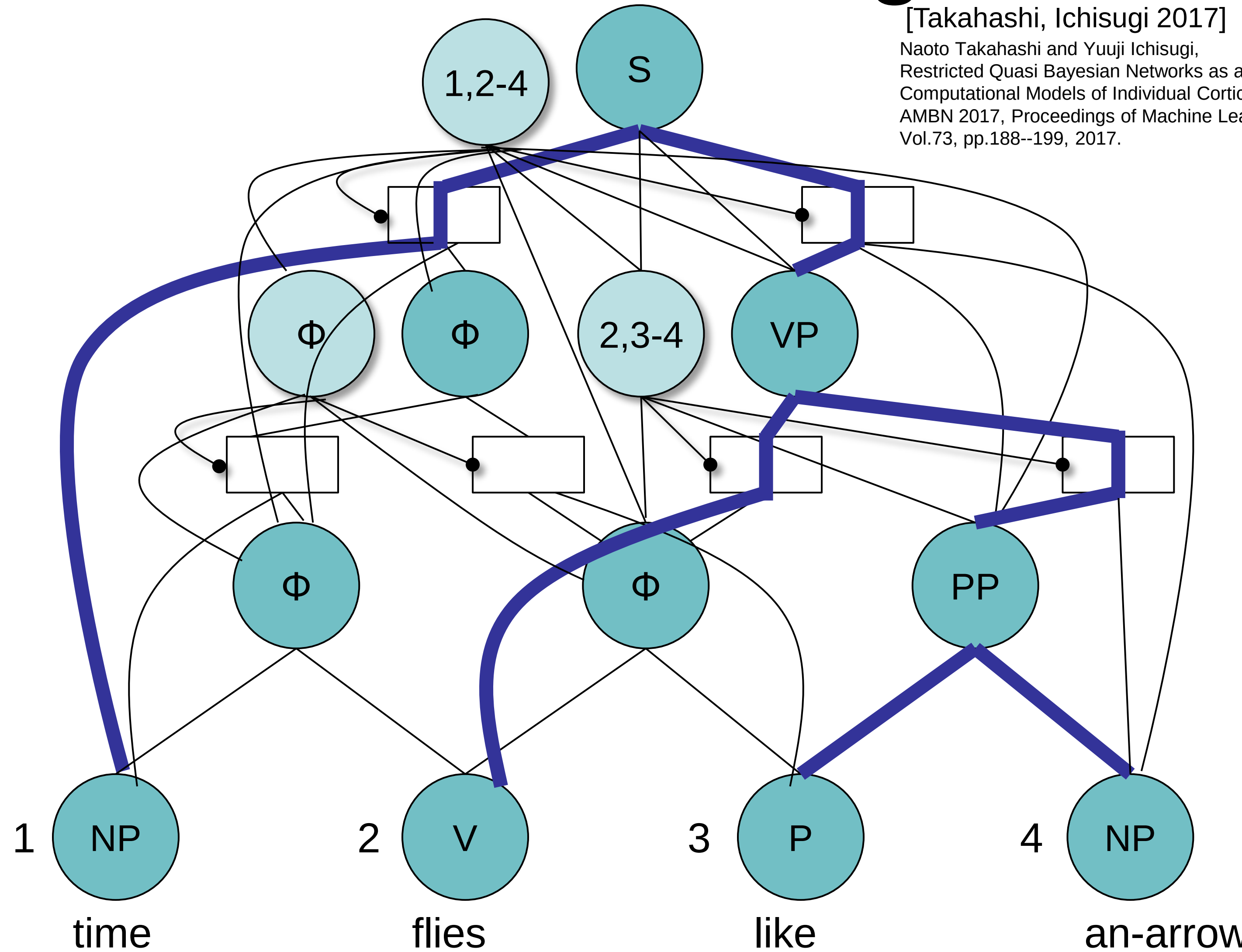
Motivation of this work

- If the cerebral cortex is a kind of Bayesian network, we can build a system that reproduces the behavior of the human language areas using a Bayesian network.
- This will become a new evidence for the Bayesian network hypothesis of the cerebral cortex.

Previous work: Bayesian network parser for context free grammar

[Takahashi, Ichisugi 2017]

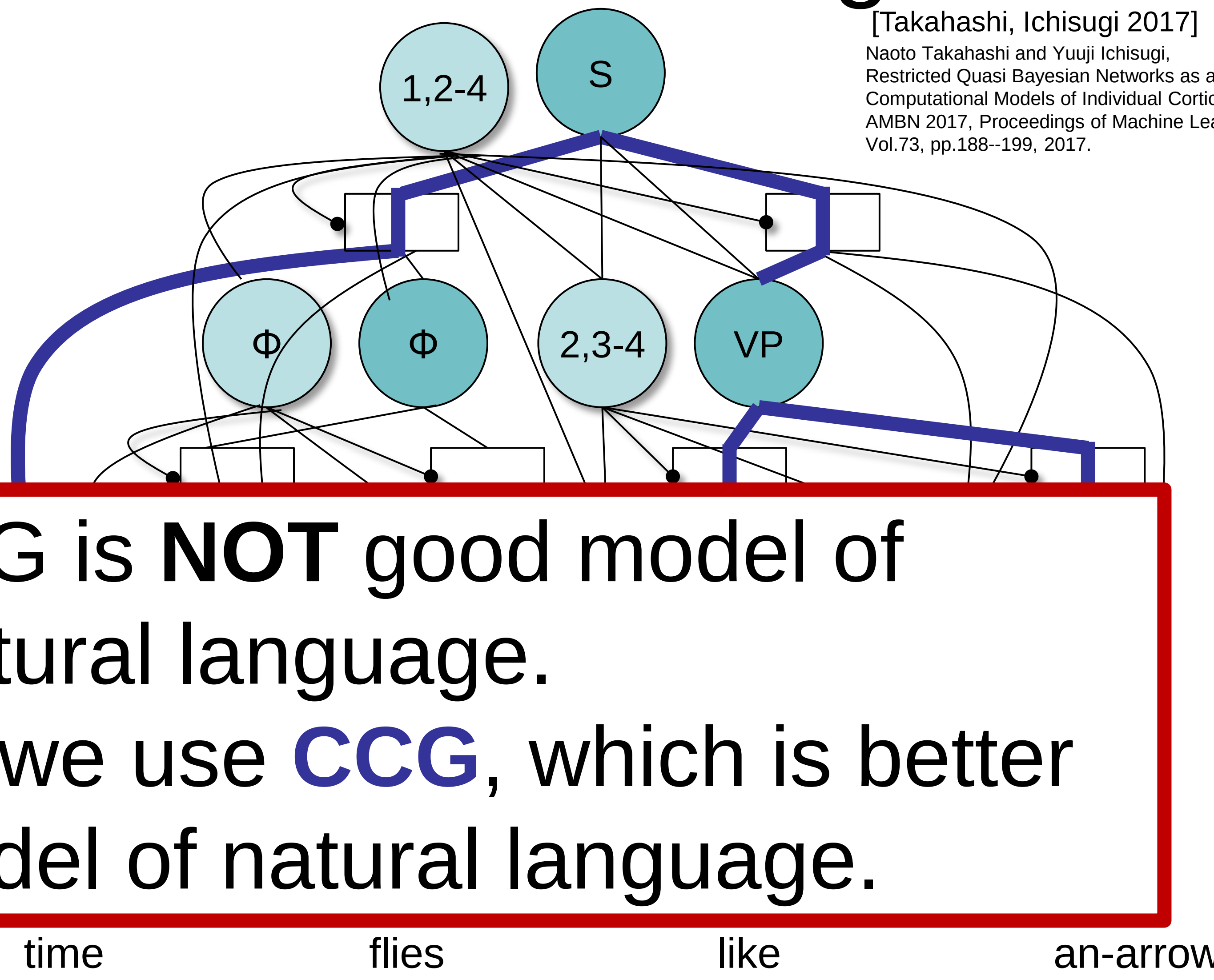
Naoto Takahashi and Yuuji Ichisugi,
Restricted Quasi Bayesian Networks as a Prototyping Tool for Computational Models of Individual Cortical Areas, In Proc. of AMBN 2017, Proceedings of Machine Learning Research, Vol.73, pp.188--199, 2017.



Previous work: Bayesian network parser for context free grammar

[Takahashi, Ichisugi 2017]

Naoto Takahashi and Yuuji Ichisugi,
Restricted Quasi Bayesian Networks as a Prototyping Tool for Computational Models of Individual Cortical Areas, In Proc. of AMBN 2017, Proceedings of Machine Learning Research, Vol.73, pp.188--199, 2017.



CFG is **NOT** good model of natural language.
So we use **CCG**, which is better model of natural language.

Combinatory Categorial Grammar (CCG) [Steedman 2000]

- CCG is one of the most successful frameworks of grammar description in theoretical linguistics.
  - CCG successfully explains many language phenomena.
- We consider that CCG is a promising theory of information processing of the brain.

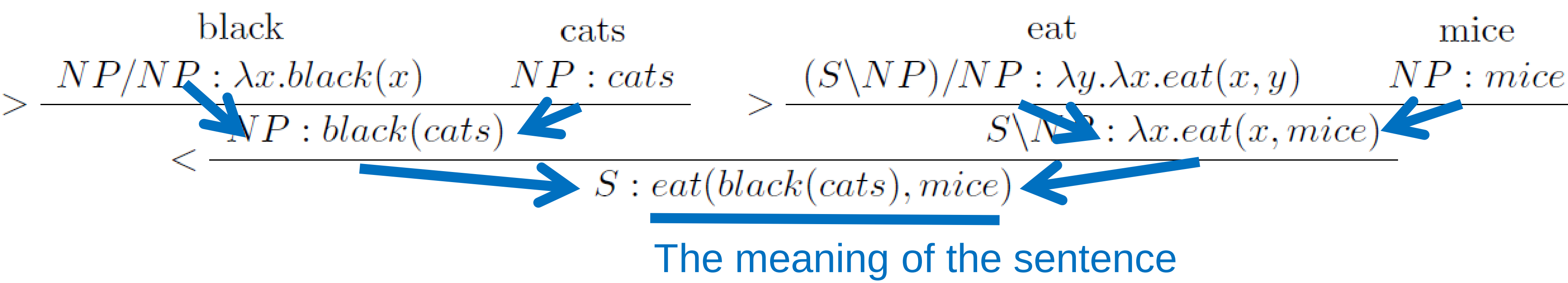
$$>\ \frac{\overset{\text{black}}{NP/NP : \lambda x.black(x)} \qquad \overset{\text{cats}}{NP : cats}}{NP : black(cats)} \qquad >\ \frac{\overset{\text{eat}}{(S\backslash NP)/NP : \lambda y.\lambda x.eat(x,y)} \qquad \overset{\text{mice}}{NP : mice}}{S\backslash NP : \lambda x.eat(x, mice)}$$

$$<\ \frac{NP : black(cats) \qquad S\backslash NP : \lambda x.eat(x, mice)}{S : eat(black(cats), mice)}$$

An example of CCG parsing

Semantic analysis in CCG

・The meaning of each word is represented as **lambda term**.
・The meaning of the whole sentence is composed by **function applications** and **function compositions**.



Problems as a model of language areas:
Because CCG uses variable-length data structure,
it is hard to be implemented by neural networks in the brain.

Hierarchical address representation

Address: (C,R,F)

Clause:
C ∈{sconj, c1, c2}
Semantic Role:
R ∈{action, agent, patient,...}
Feature:
F ∈ {entity, color, size, ...}

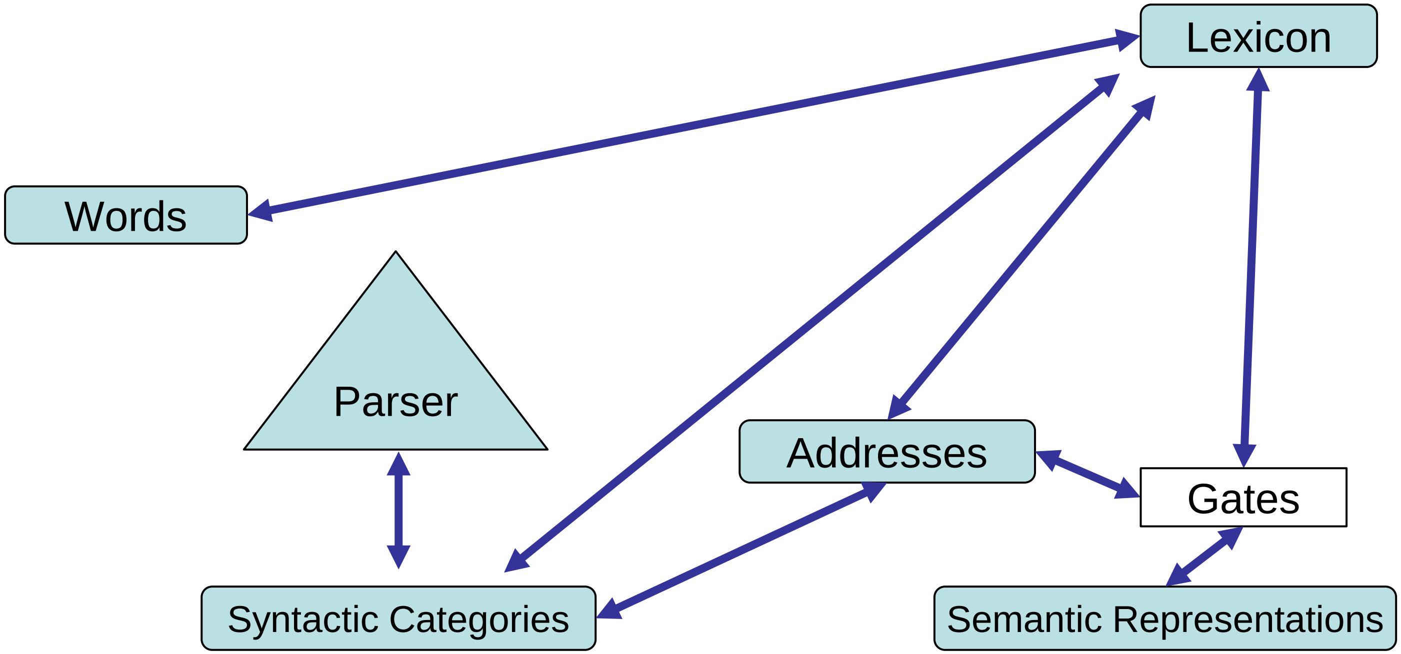
The meaning of a sentence is represented as a set of pairs of addresses and semantic representations.
-> **flat and fixed-size structure**

This model **does not use variable-length data structure** such as lambda terms.

Address	SR
$(sconj, -, -)$	if
$(c1, agent, size)$	big
$(c1, agent, color)$	*
$(c1, agent, entity)$	dogs
$(c1, modality, -)$	*
$(c1, action, -)$	chase
$(c1, patient, size)$	small
$(c1, patient, color)$	*
$(c1, patient, entity)$	mice
$(c2, agent, size)$	*
$(c2, agent, color)$	black
$(c2, agent, entity)$	cats
$(c2, modality, -)$	may
$(c2, action, -)$	eat
$(c2, patient, size)$	*
$(c2, patient, color)$	*
$(c2, patient, entity)$	mice

"if small mice areChasedBy big dogs black cats may eat mice"

Modules of the proposed model

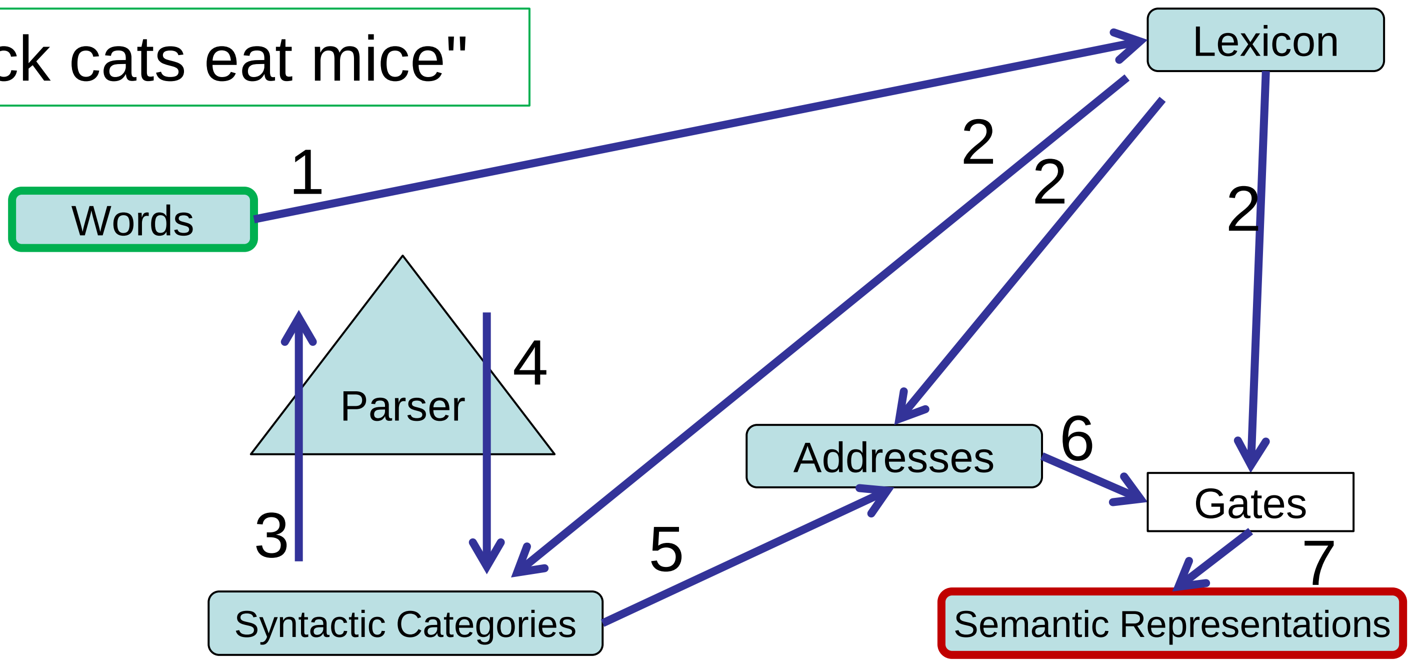


Modules are reciprocally (bidirectionally) connected like Bayesian networks, and like cerebral cortex.

Parsing

Input words:

"black cats eat mice"



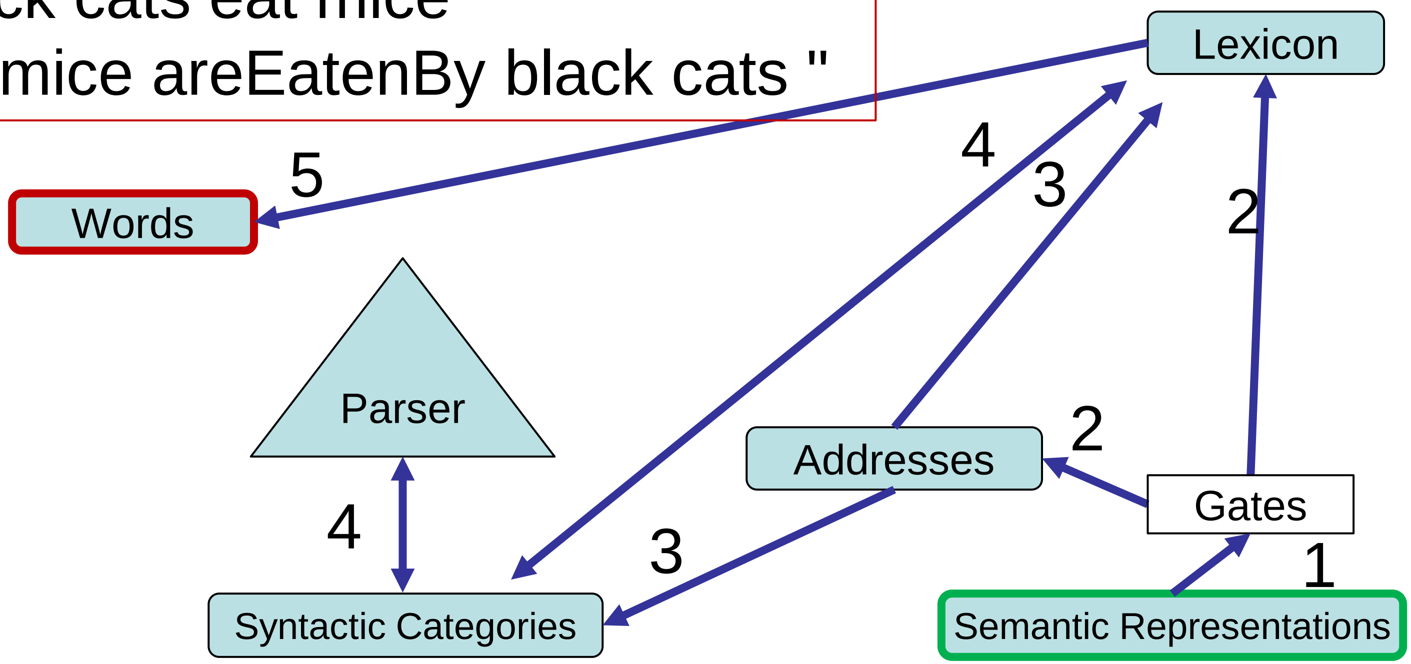
Output:

Address	SR
$(c1, agent, color)$	black
$(c1, agent, entity)$	cats
$(c1, action, -)$	eat
$(c1, patient, entity)$	mice

Utterance(Speech)

Output words:

"black cats eat mice"
or " mice areEatenBy black cats "



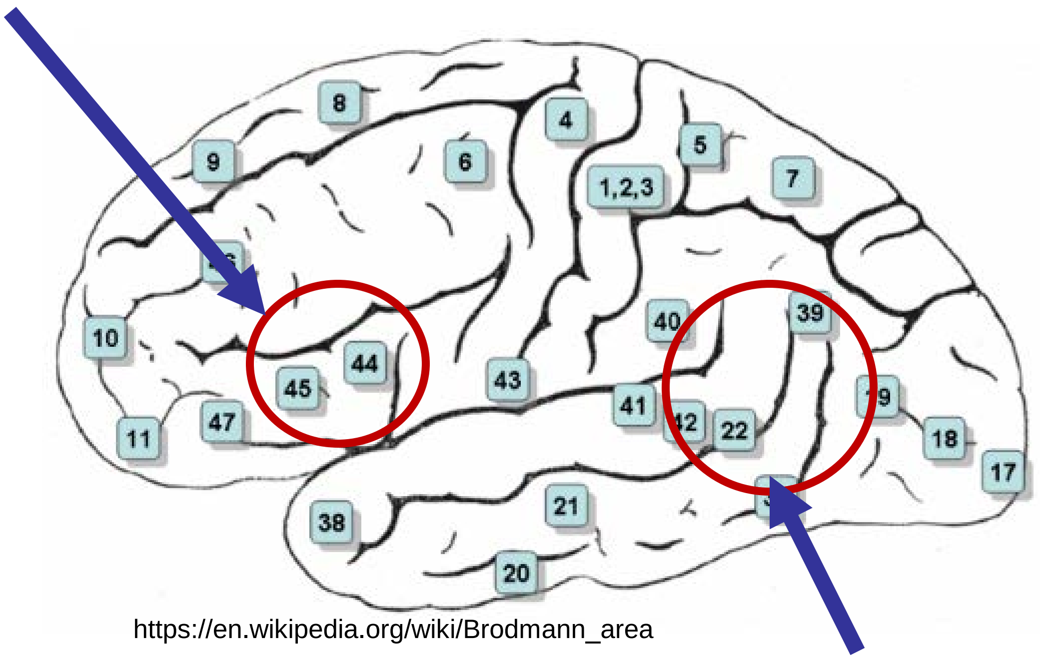
Input:

Address	SR
$(c1, agent, color)$	black
$(c1, agent, entity)$	cats
$(c1, action, -)$	eat
$(c1, patient, entity)$	mice

Correspondence to cortical areas

Language Areas

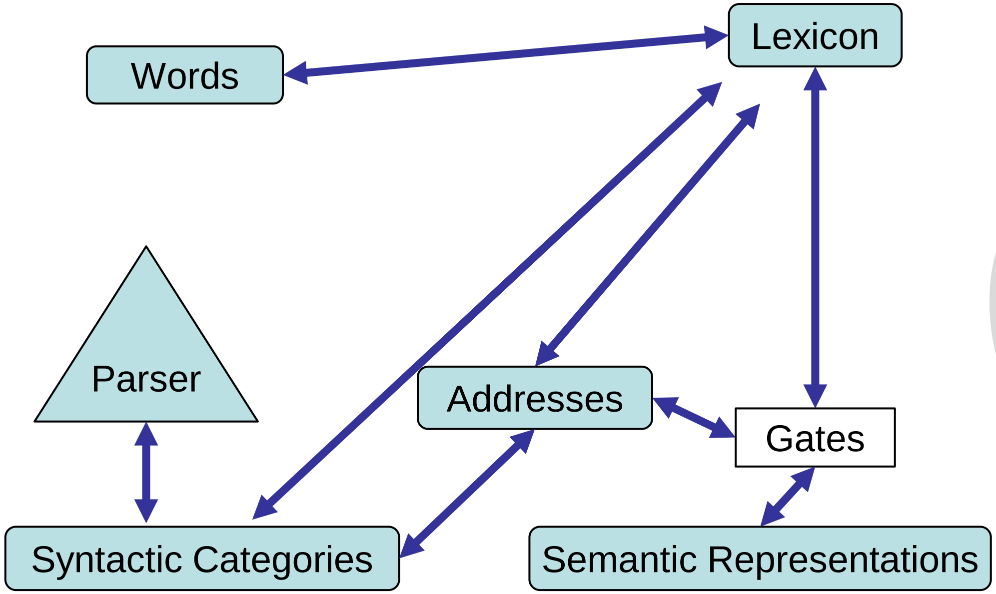
Broca's area（BA 44,45）



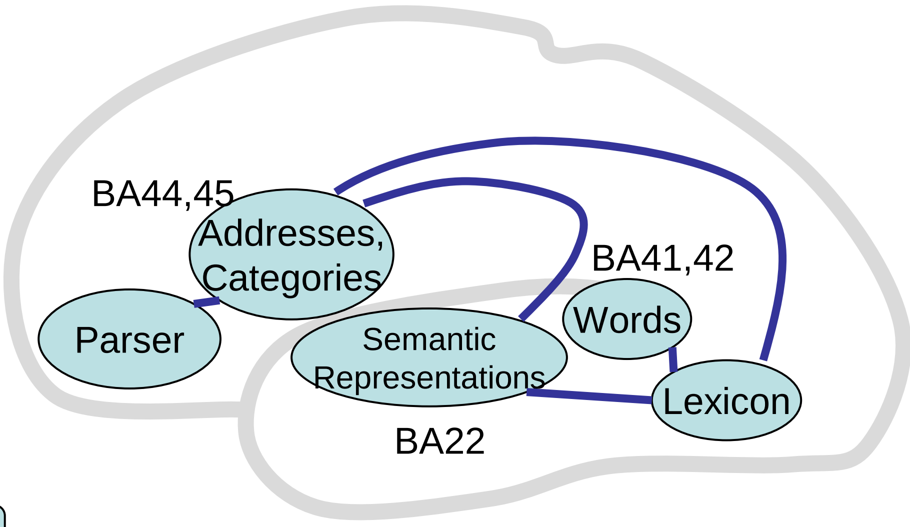
https://en.wikipedia.org/wiki/Brodmann_area

Wernicke's area (BA22)

A possible correspondence between the modules in the model and cortical areas



Proposed model



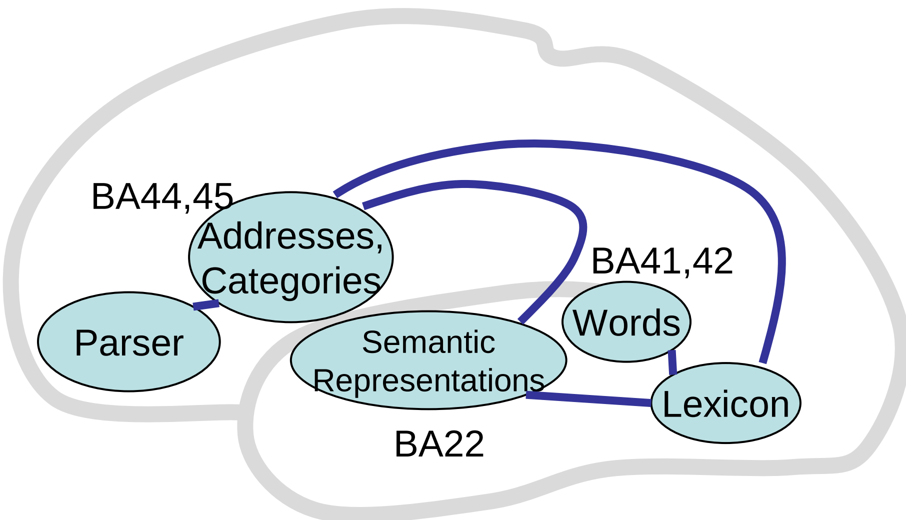
Cortical areas

We can reproduce utterance of aphasia by disabling some modules.

Broca's area and Wernicke's area

- Broca's area : Grammar processing
  - **Symptoms of Broca's aphasia :** Utterance consists of scattering words that do not constitute sentences
- Wernicke's area : Association between speech sounds and concepts
  - **Symptoms of Wernicke's aphasia :** Utterance that is fluent but does not make sense because of mistakenly selected words.

Normal utterance



Input semantic representation:

Address	SR
$(c1, agent, color)$	black
$(c1, agent, entity)$	cats
$(c1, action, -)$	eat
$(c1, patient, entity)$	mice

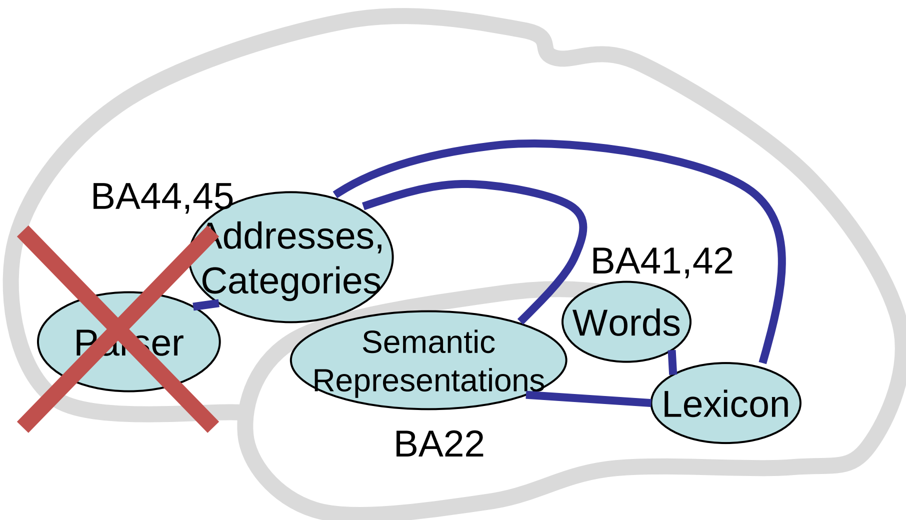
```
?- M=[[[c1,agent,color],black],
[[c1,agent,entity],cats], [[c1,action,-],eat],
[[c1,patient,entity],mice]],
Ws=[W1,W2,W3,W4], maplist(lexicalItem,
Ws, Cs, As, Ds), parse([], Cs, s(c1)),
maplist(bind(M), As, Ds), print(Ws), nl, fail.
```

Output:
[black,cats,eat,mice]
[mice,areEatenBy,black,cats]

- The proposed model has been implemented in the Prolog language.
- Simplified English with many limitations

- A semantic representation is given.
- The model infers all possible sentences that consist of four words

Reproduction of Broca's aphasia



Address	SR
$(c1, agent, color)$	black
$(c1, agent, entity)$	cats
$(c1, action, -)$	eat
$(c1, patient, entity)$	mice

Disable parser module

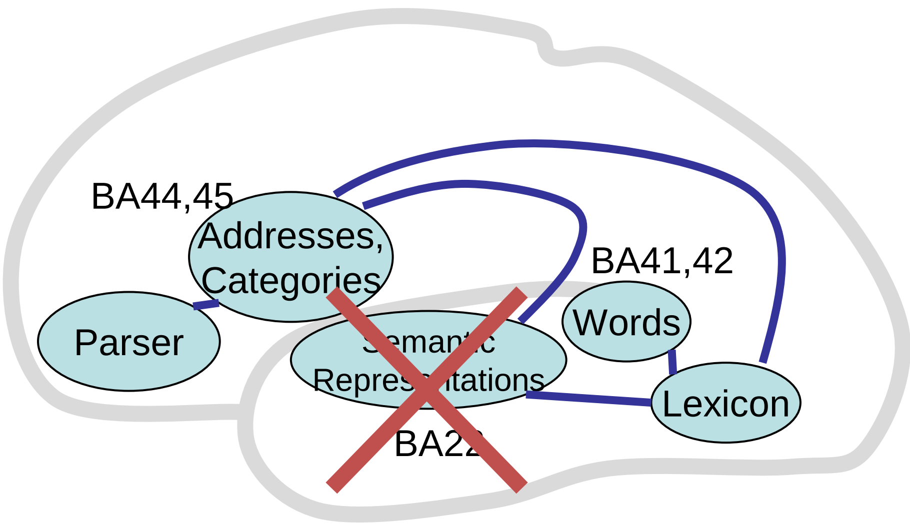
```
?- M=[[[c1,agent,color],black],
[[c1,agent,entity],cats], [[c1,action,-],eat],
[[c1,patient,entity],mice]],
Ws=[W1,W2,W3,W4], maplist(lexicalItem, Ws,
Cs, As, Ds), /* parse([], Cs, s(c1)),*/
maplist(bind(M), As, Ds), print(Ws), nl, fail.
```

Output:
[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]
...

Sentences are syntactically **incorrect.**
Selected words are semantically **correct**.

Resembles Broca's aphasia

Reproduction of Werniche's aphasia



Infer sentences without giving concrete semantic representation

?- **/* 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.

Output:
[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]
...

Sentences are syntactically **correct.**
Selected words are semantically **incorrect**.

Resembles Werniche's aphasia

Conclusion

- We proposed a model of the mechanism of the semantic analysis that **does not use variable-length data structure** such as lambda terms.
- The model **must be realized as a cortex-like Bayesian network** in the future.
- Utterance of aphasia is reproduced.
- This research will connect computational neuroscience and theoretical linguistics.