

BESOMを使った言語野モデル の構想

産業技術総合研究所

一杉裕志

2012-12-28

2013-01-08 一部修正

BESOMとは

- 大脳皮質の機能を再現することを目指して設計・実装中の計算論的モデル。
 - 特殊なベイジアンネット。
 - ノード数 n に対して、1入力の認識・学習に必要な計算量がほぼ $O(n)$ 。
 - ネットワークの構造は非線形BSSにより自己組織化。
- 注：BESOM は未完成。
 - ただし、ゆっくりと着実に進展中。

目標:このような言語野の計算論 的モデルを作りたい。

- 学習
 - データのみから構文規則と意味規則を自己組織化(教師なし学習)できる。
- 認識
 - 並列動作により単語長 n の文を $O(n)$ の時間で理解できる。
 - 係り受けの関係を正しく認識できる。
- 神経科学的妥当性
 - BESOM の基本動作を前提とする。
 - マクロな構造や機能が言語野の知見と一致する。

対象としないこと

- 脳の中の汎用情報処理装置がからむ言語処理は対象としない。
 - おそらく前頭前野は汎用の状態遷移機械。
 - おそらく海馬は汎用の連想記憶装置。
 - この2つがあれば限りなく複雑な言語処理ができてしまう。
- ここでは、言語野が担当すると思われる、人間が直感的に瞬時に行える言語処理のみをモデル化の対象とする。

このあとのスライドのアウトライン

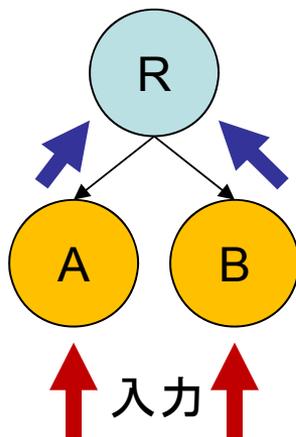
- BESOMの基本動作
- explaining away
- バインディングとオブジェクトファイル
- BESOMを使った言語理解のモデル
 - 構文解析のモデル
 - 係り受け構造に対応した意味を想起するモデル
- 将来の課題

BESOMの基本動作

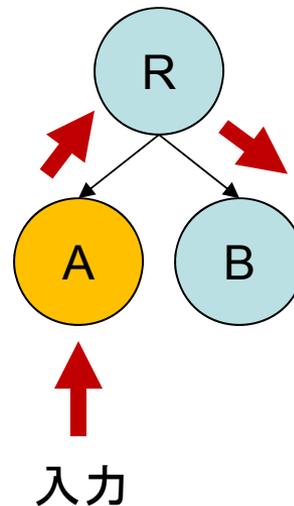
1ノードの場合

BESOMのノードは「データベース」

学習



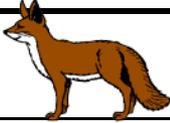
想起

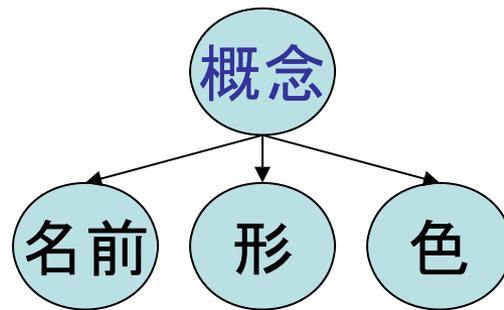


ノードRには、同時入力されるAとBの値の組がデータベースとして記憶される。

Aの値が与えられると、データベースを検索し、対応するBの値が想起される。

データベースの学習の例

ユニットID	名前	形	色
U01	きつね		茶色
U02	うさぎ		白
U03	かめ		緑
U04	からす		黒
U05	はと		灰色
U06	やぎ		白
...			

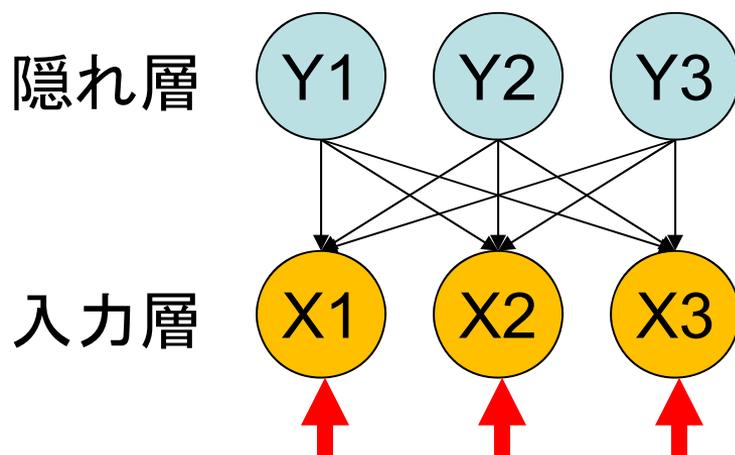


注: 競合学習を使って圧縮・抽象化して記憶する。

BESOMの基本動作 2層の場合

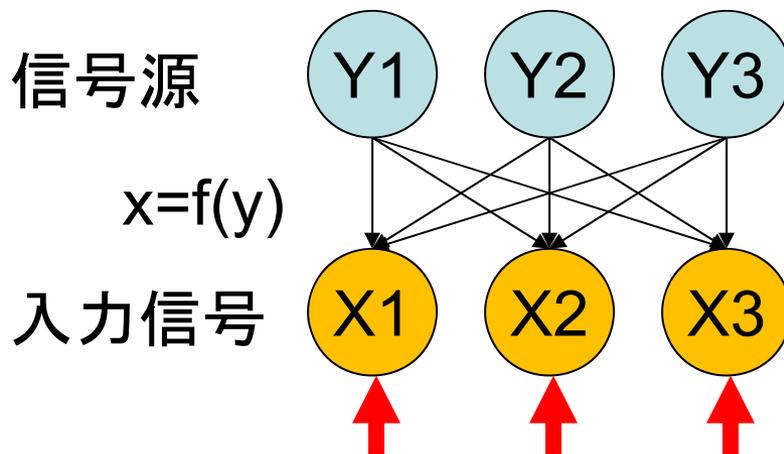
2層BESOM

- 隠れ層の各ノードは同じ入力を受け取るが、同じ学習結果になっても意味がない。
- ノードごとの役割分担を自己組織化する機構が必要。



隠れノードの意味の推定=BSS

- ベイジアンネットにおいて親ノードは子ノードの信号源。
- 入力信号だけから信号源を教師なしで推論すればよい。
 - 暗号信号源分離(BSS, blind source separation)と呼ばれる問題。

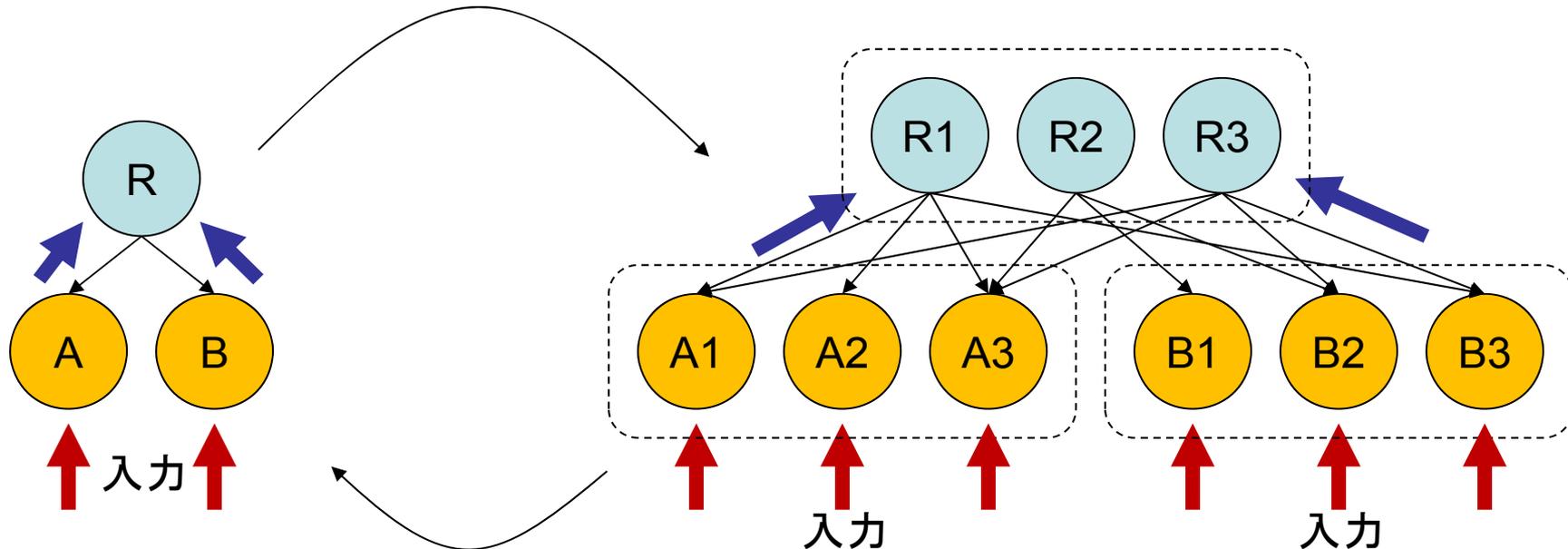


信号源の推定方法

- 解くべき問題: **非線形BSS** → **決して簡単ではない**
- 信号源に関する**事前知識**を駆使して推定
 - 信号源どうしはほぼ独立。
 - 信号源はスパースにしか活性化しない。
 - 信号源の状態が変化すると生成される信号も滑らかに変化する。
 - 信号は複数の信号源が生成する値のORで決まることが多い。
 - その他、扱うデータに固有の知識。
 - 視覚刺激はゆっくりとしか変化しない。
 - 視覚刺激は物体の形と位置という独立成分に分解できる。
 - …

2層BESOMもデータベースとして 動作

各変数を複数の独立成分に分解



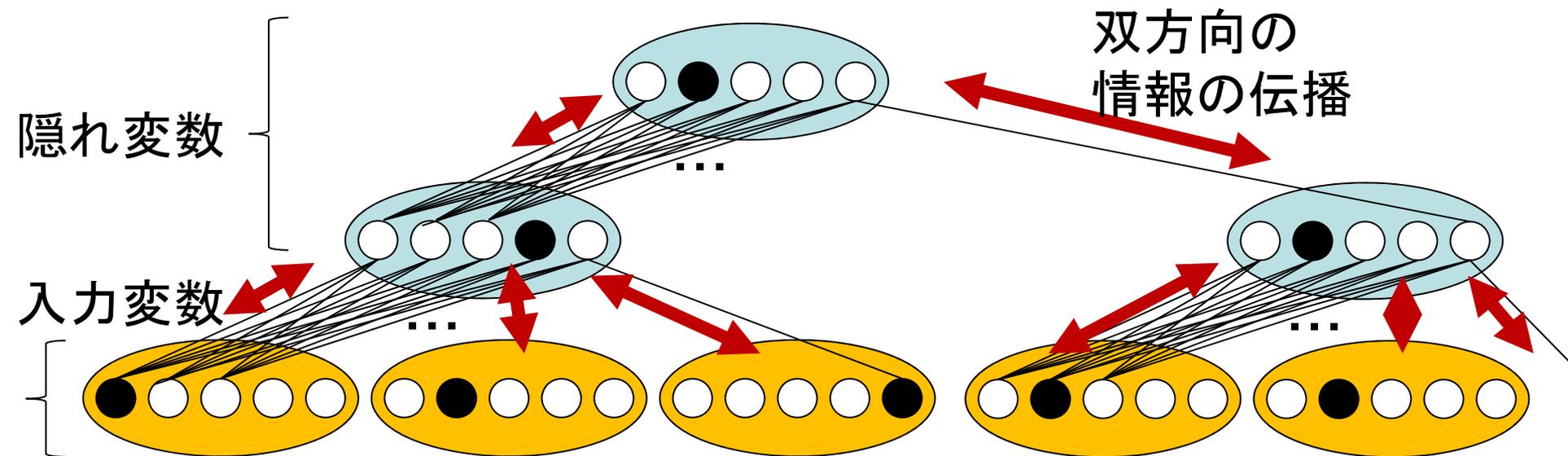
複数の変数を1つに統合

2つのベイジアンネットの機能は等価

BESOMの基本動作 多層の場合

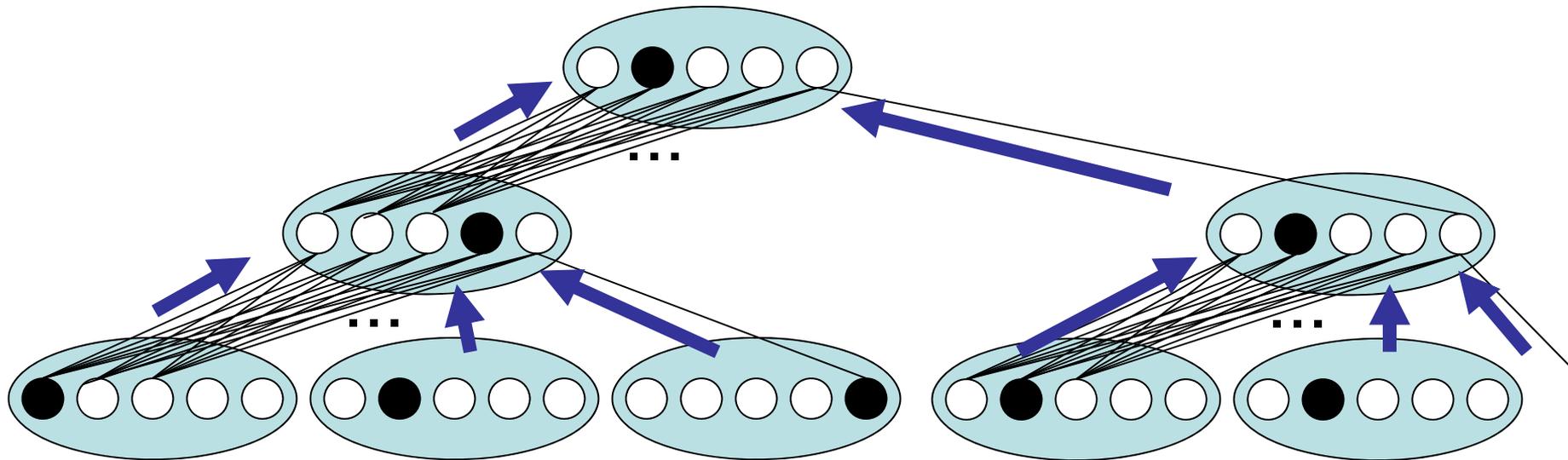
BESOMネットワークの動作：認識

- 全体がベイジアンネットとして動作
- 入力変数の観測値をもっともよく説明する隠れ変数の状態を想起



BESOMネットワークの動作：学習

- 各ノードがSOMとして動作
- 子ノードの値を圧縮・抽象化して記憶



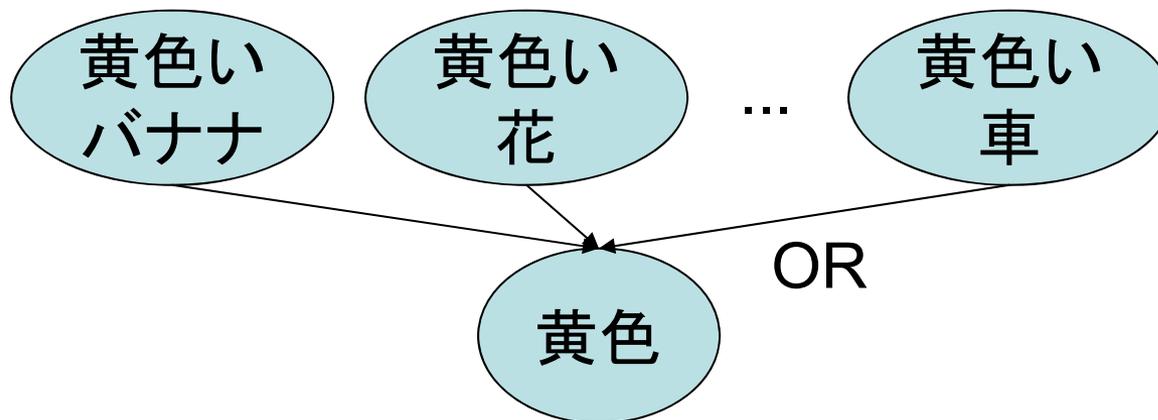
explaining away

explaining away とは

- 信号からその原因となる信号源を推定するとき、複数考え得る原因が結果の解釈を取り合うことによって起きる現象。
- 入力から出力へ一方向に情報が流れるだけのニューラルネットでは起きない。
- 情報が双方向に繰り返し流れて全体を最適化するベイジアンネットで起きる。

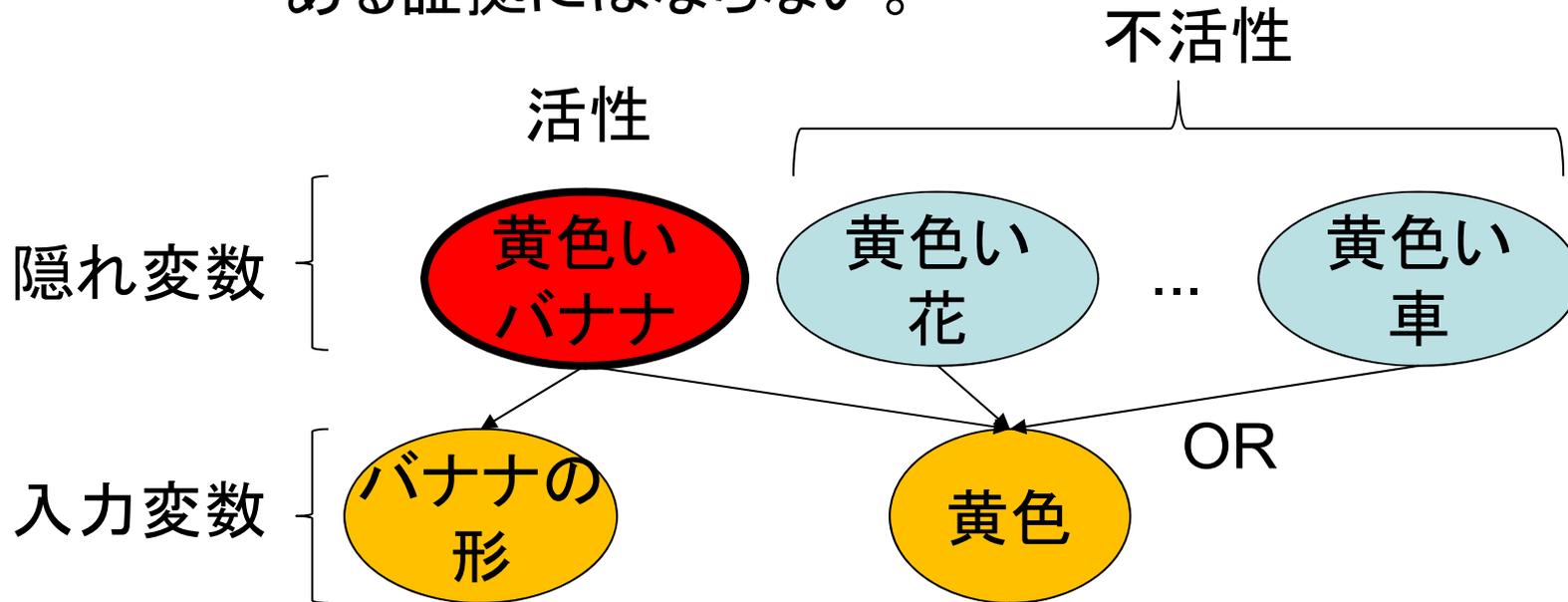
自然界の因果関係はORで表現できることが多い

- Cf. noisy-OR model [Pearl 1988]
- 特徴：
 - 活性化した信号源が信号を生成する。
 - 不活性化信号源は信号に影響を与えない。
- 例：何か1つ黄色いものが目の前にあれば、視野の一部に黄色が見える。



explaining away 現象

- 信号源のどれかが活性状態であると確定すれば、他の信号源は不活性状態であろうと推論される。
 - 黄色が見えていても、もはや黄色い花や黄色い車がある証拠にはならない。



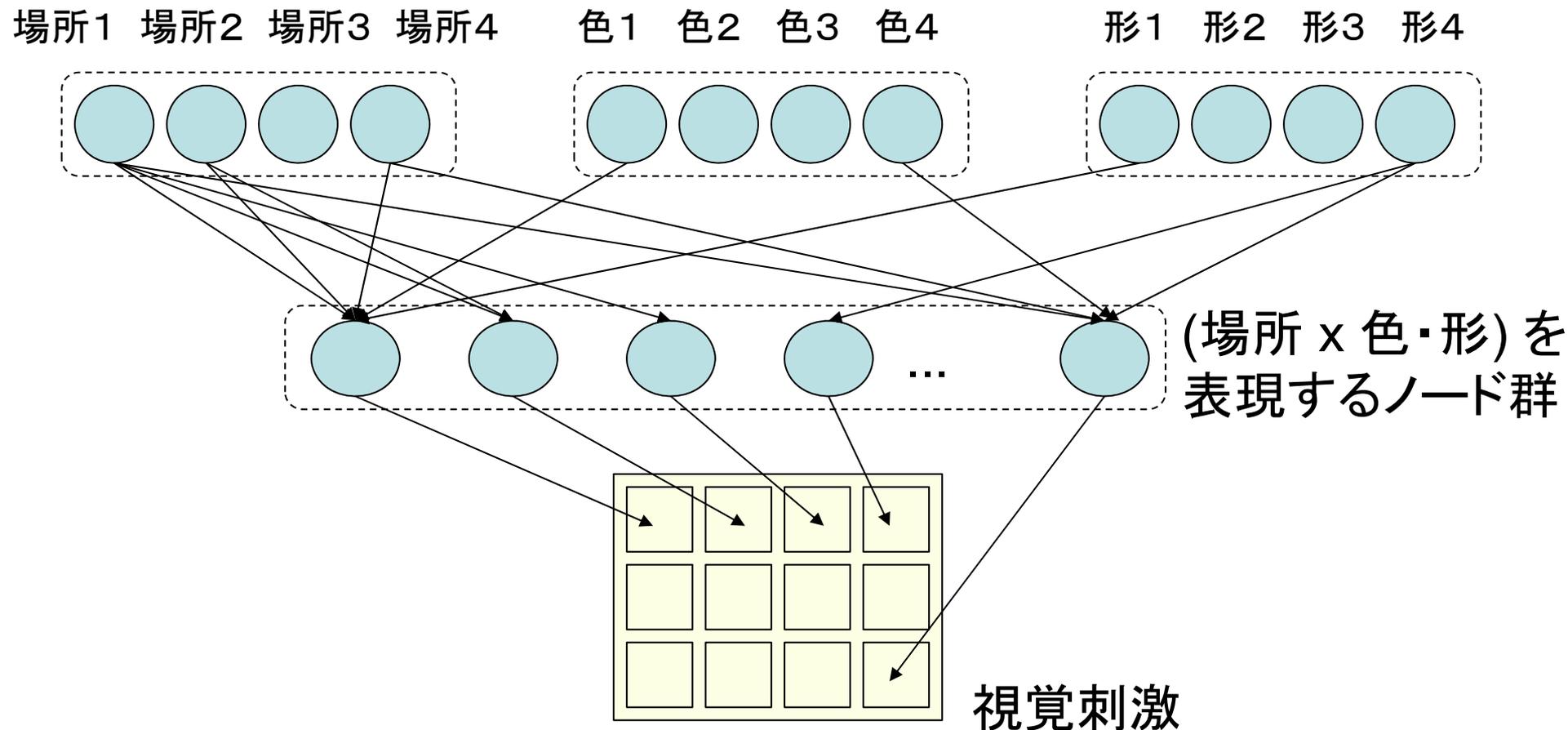
バインディングと オブジェクトファイル

オブジェクトファイル理論 [Kahneman 1992]

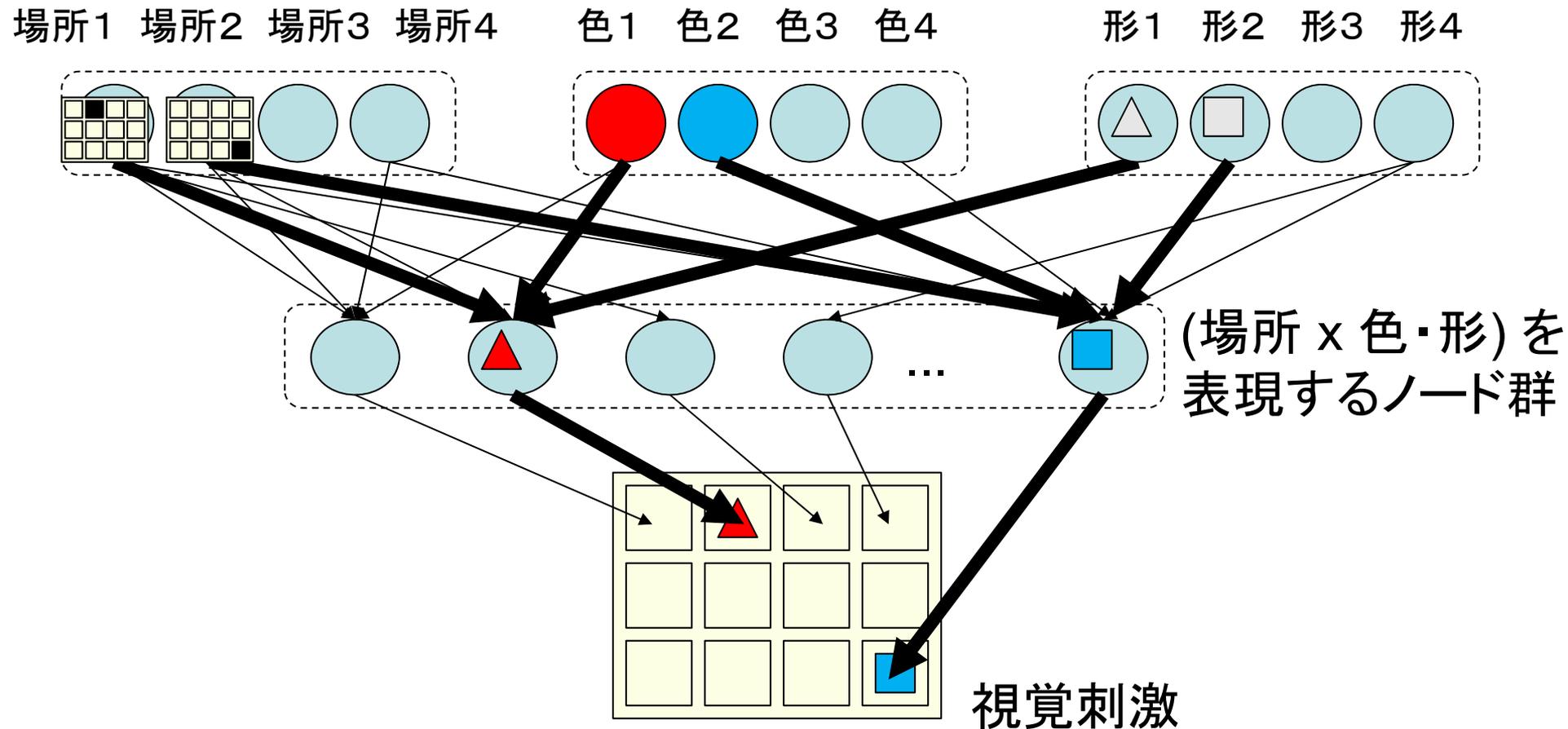
- 脳は視覚刺激の中の物体の位置、動き、形、色などの属性を「オブジェクトファイル」と呼ばれるデータ構造でむすびつけて一時的に記憶する、という仮説。
- 参考：
日本神経回路学会誌 Vol. 16(2009) No. 1
「視覚性短期記憶における結び付け問題」
齋木 潤
https://www.jstage.jst.go.jp/browse/jnns/16/1/_contents/-char/ja/
- オブジェクトファイルの数は多くても4個程度。

オブジェクトファイルの BESOM ネットによる表現(案)

オブジェクトファイルの各属性は
視覚刺激を生成する**信号源**。



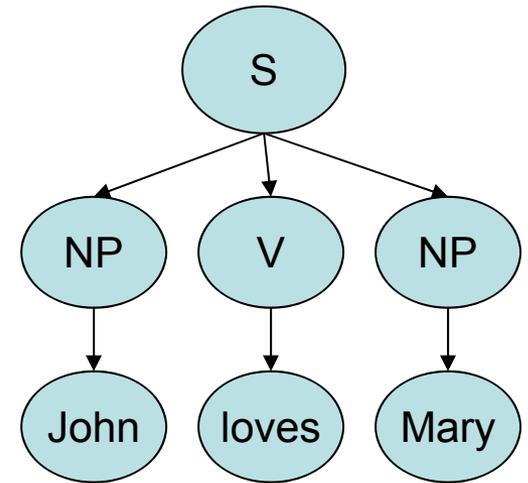
オブジェクトファイルに表現される 画像認識結果



構文解析のモデル

「文法」＝「文の生成モデル」

- 文法とは文の生成モデルである。
- 文の生成規則は、無数の文を抽象化したものである。



構文解析木

チャートパーサ

- 参考:

「チャートパーサ - Wikipedia」

<http://ja.wikipedia.org/wiki/%E3%83%81%E3%83%A3%E3%83%BC%E3%83%88%E3%83%91%E3%83%BC%E3%82%B5>

「チャートパーサ(英: Chart parser)は、自然言語などの曖昧な文法に向いた構文解析器の一種である。動的計画法を用い、中間的かつ仮説的な結果をチャート(chart)と呼ばれるデータ構造に格納しておき、再利用する。これによりバックトラッキングを省き、同時に組合せ爆発を防ぐ。」

構文解析するBESOMネットの設計方針

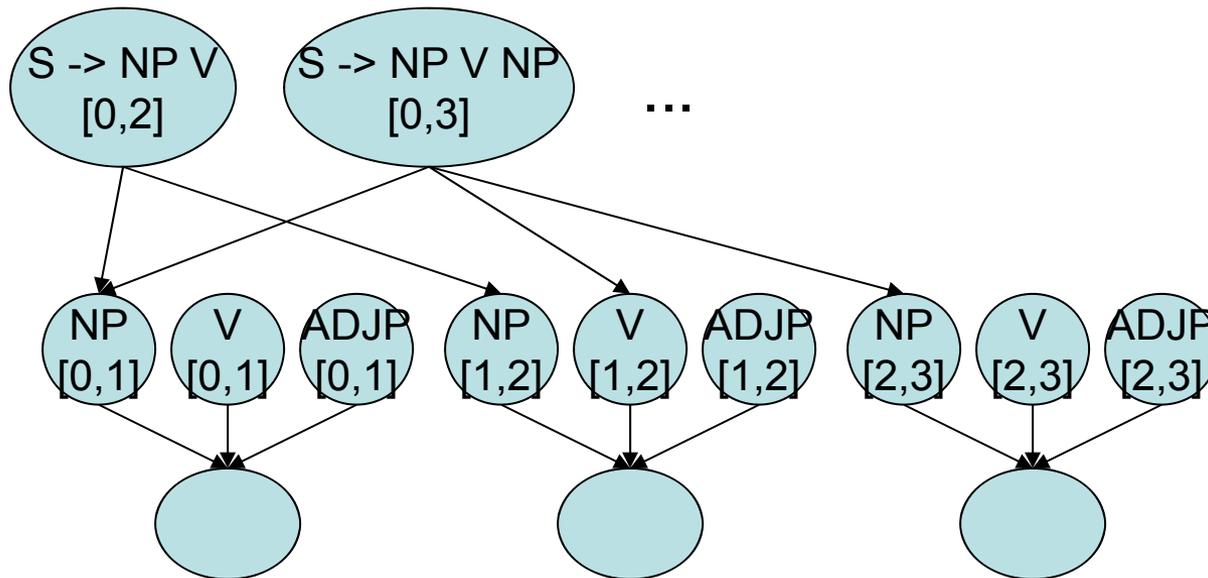
- チャートパーサに似たものを並列実行するBESOMネットを作る。
- 考えられるすべての中間的解析結果 ($O(n^2)$ 個) を、あらかじめノードで表現。
- $O(n^2)$ のノードが並列動作すれば $O(n)$ 程度の時間で構文解析できるはず。
- 再帰的な文法は、再帰的なBESOMネットで表現。

構文解析するBESOMネットの設計方針(つづき)

- 実際の言語野は単語が入力されるたびにパイプライン的に構文解析されると思われるが、ここでは第一歩として、文全体をバッファに蓄えて構文解析結果を静的に表現するようなモデルを考えてみる。

再帰的でない単純な文法を構文解析するBESOMネット(案)

- $[i, j]$ の i は単語列における開始位置、 j は終了位置を表す。
- 可能な i, j の組み合わせの数だけノードを用意する。



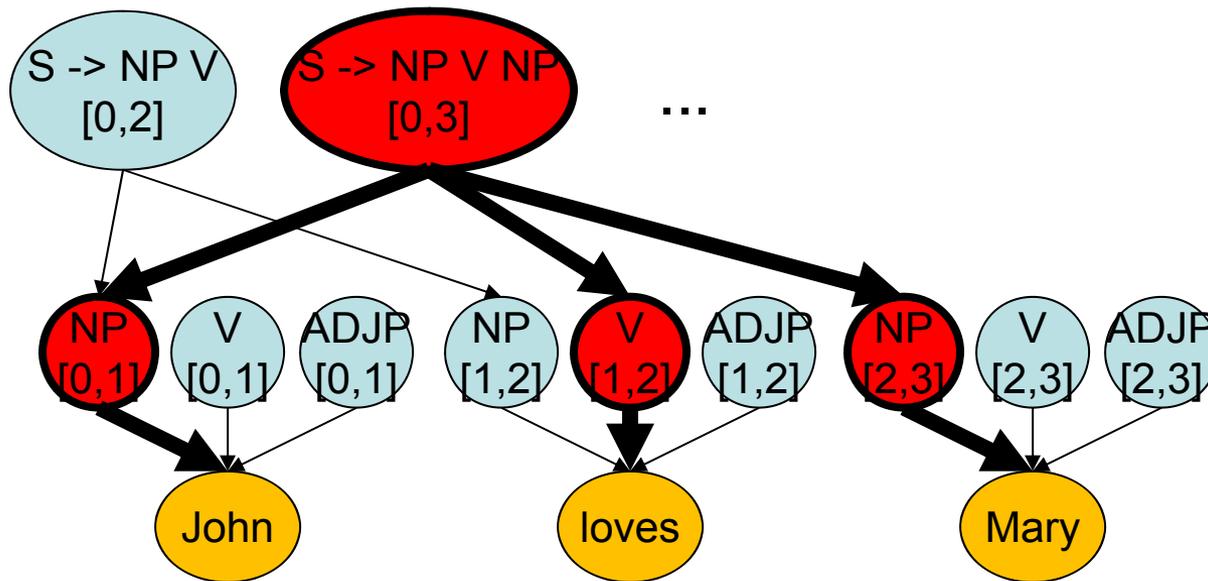
文法ノード
生成規則の数 $\times O(n^2)$

品詞ノード
品詞の数 $\times n$

単語ノード n

認識動作の例

隠れ変数の値の推論結果 = 構文解析木



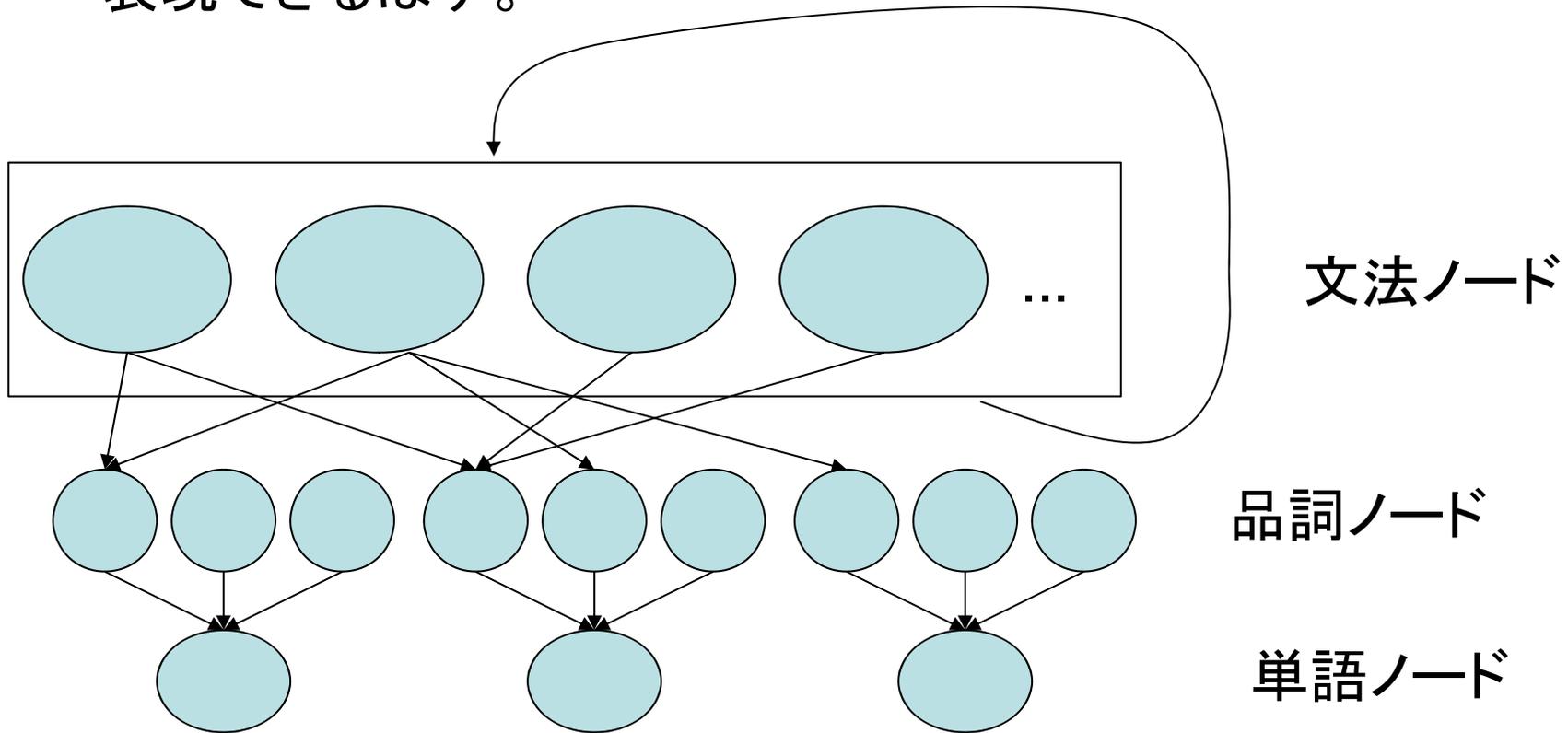
文法ノード

品詞ノード

単語ノード
(観測値)

再帰的文法を構文解析する BESOMネット(案)

リカレントなBESOMネットであれば、生成規則の無限の階層を表現できるはず。



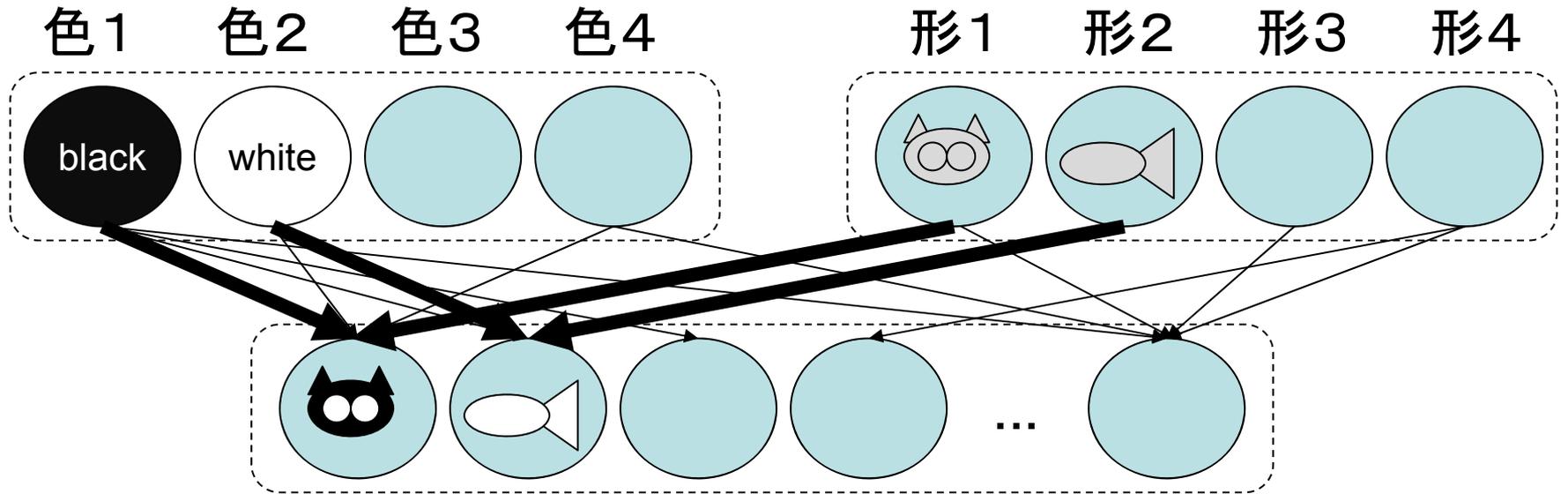
係り受け構造に対応した意味を
想起するモデル

自然言語の「意味」とは？

- ここでは、「言語野以外の領野の発火パターンの変化の仕方」が「意味」として考える。
- "A black cat eats a white fish." という文を聞くと、それに相当するイメージが視覚野などの発火パターンとして表現される。
- このとき、黒いのはネコ、白いのは魚、という係り受け構造が適切に処理されなければならない。
- 主語はネコであって魚ではない、ということも発火パターンで表現できなければならない。

係り受けの正しい理解

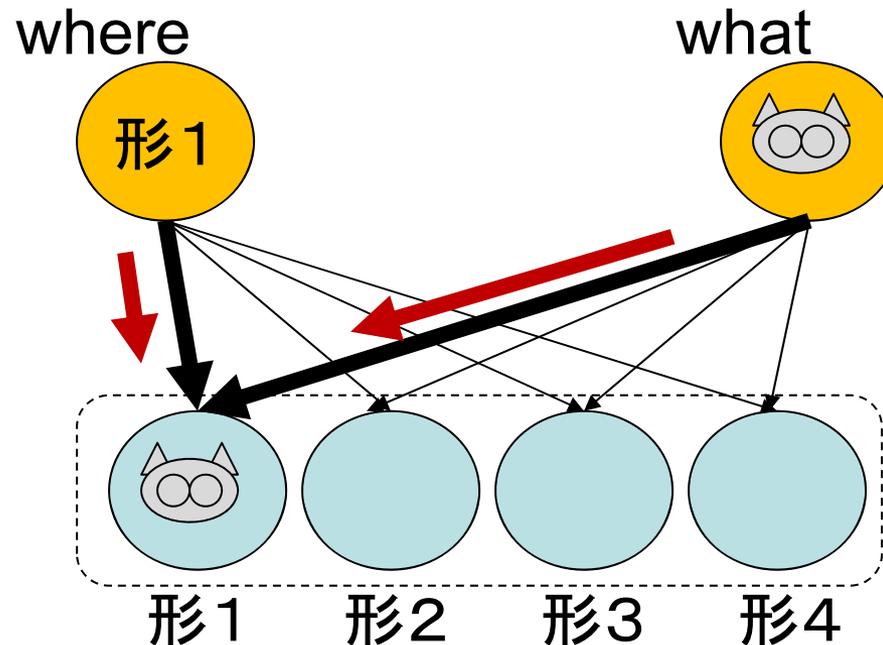
"A black cat eats a white fish." という文で、視覚野に下記のような発火パターンを想起させたい。



(場所 x 色・形) を
表現するノード群

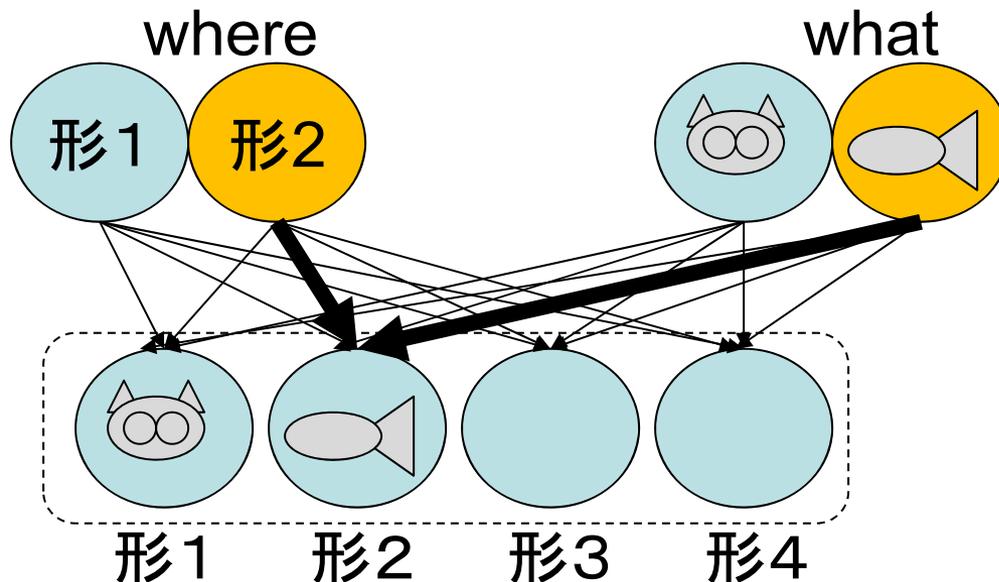
オブジェクトファイルへの値の書き込み

- オブジェクトファイルの番号を表すノードと、書き込む値を表すノードが親ノードにあれば、そこに値が書き込まれると仮定。



オブジェクトファイルへの追加の書き込み

- 2つの値が混ざらずに書き込めるかどうかは自明ではないが、とりあえず書き込めると仮定する。

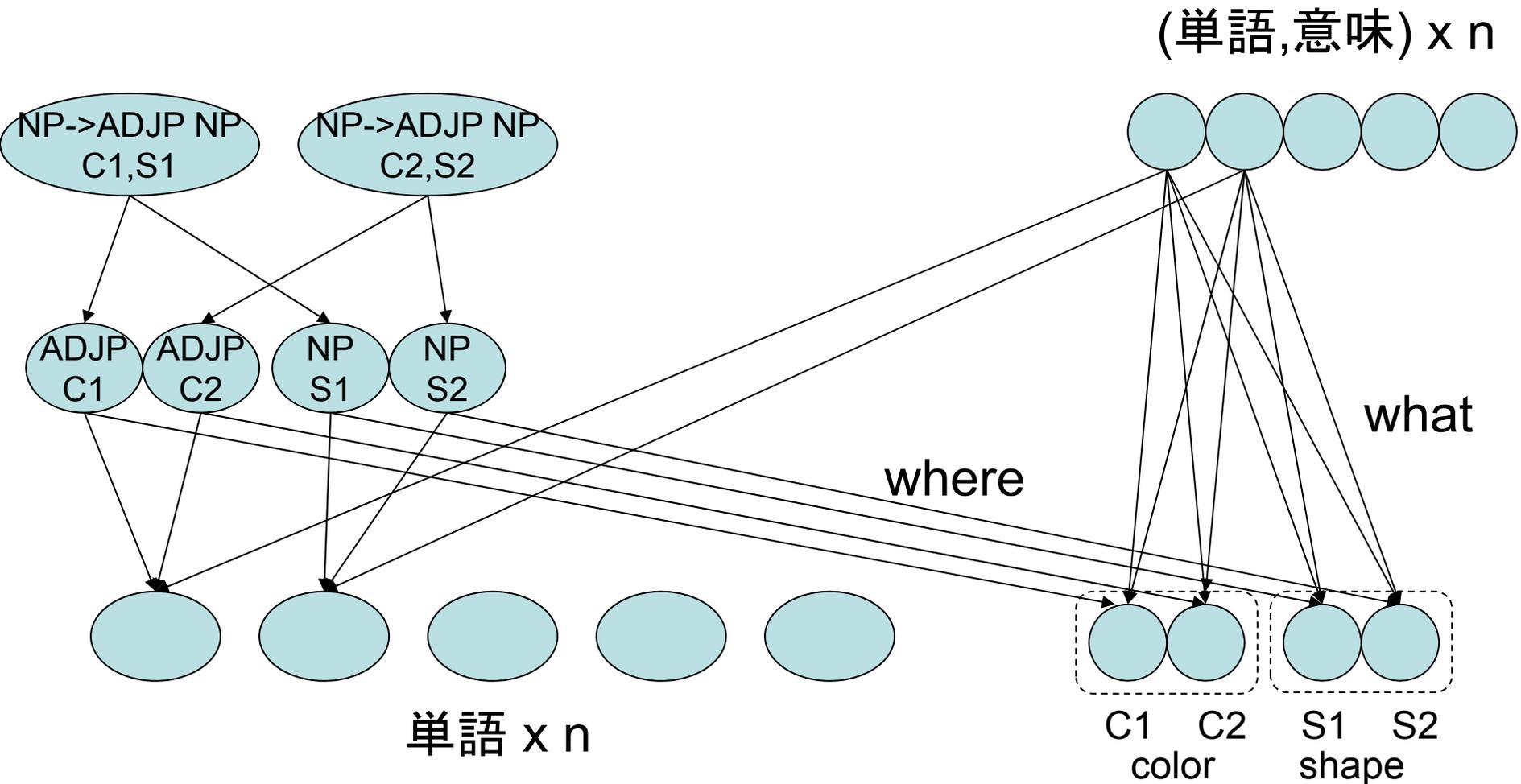


係り受けを正しく理解するBESOM ネットの設計方針

- オブジェクトファイルの数を m とする。($m=4$ 程度。)
 - 品詞ノードの数を m 倍に増やす。
 - k 個のノードを統合する文法ノードは、最大 ${}_m P_k$ 倍に増やす。 (順列 ${}_m P_k = m(m-1)\dots(m-k+1)$)
 - (最悪のケースを想定。オブジェクトファイルの使い方にバイアスがあればノード数は減らせる。例: 主語は必ず1番目のオブジェクトファイルを使う、など。)
- 単語 \times 意味を学習するノードを n 倍用意する。

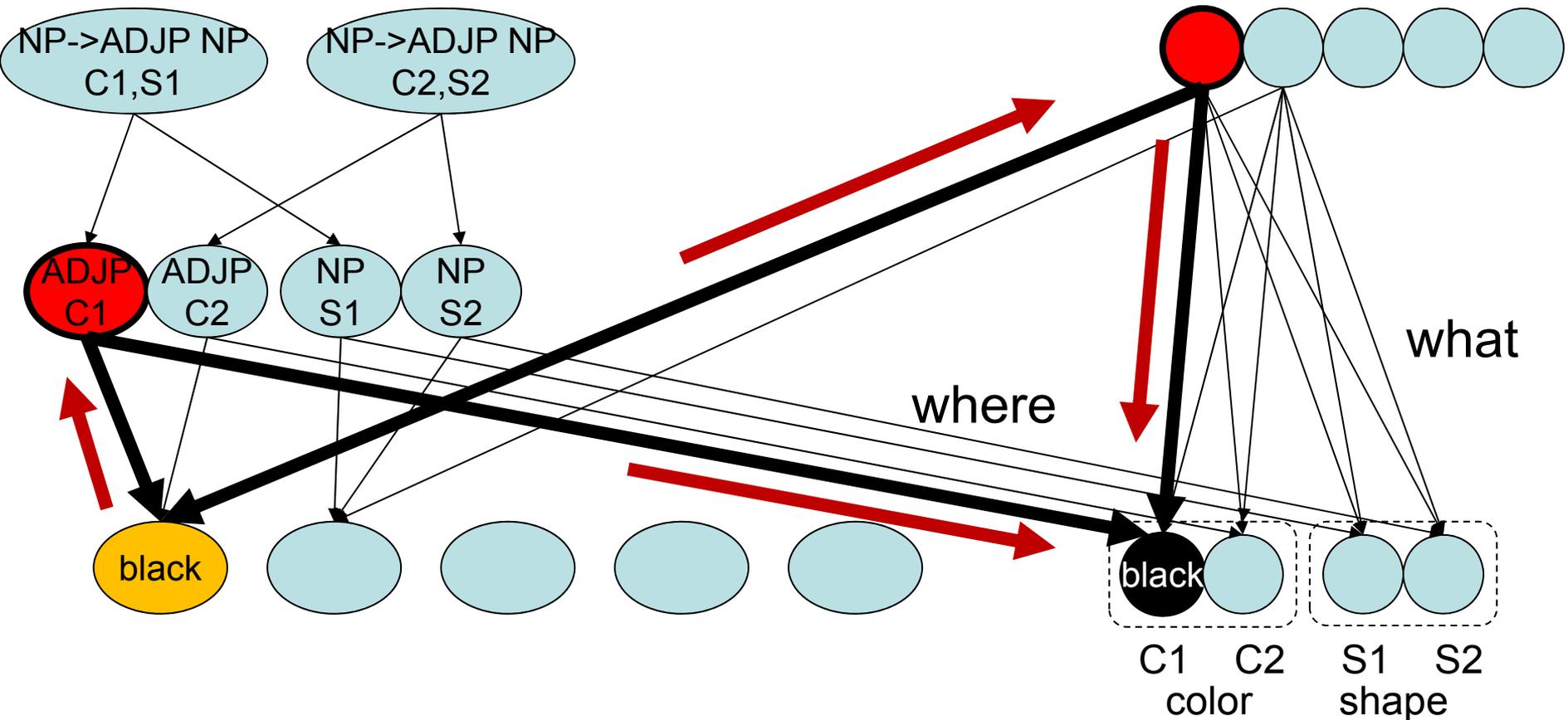
係り受けを処理できる BESOMネット(案)

ノード、エッジをかなり省略してある。



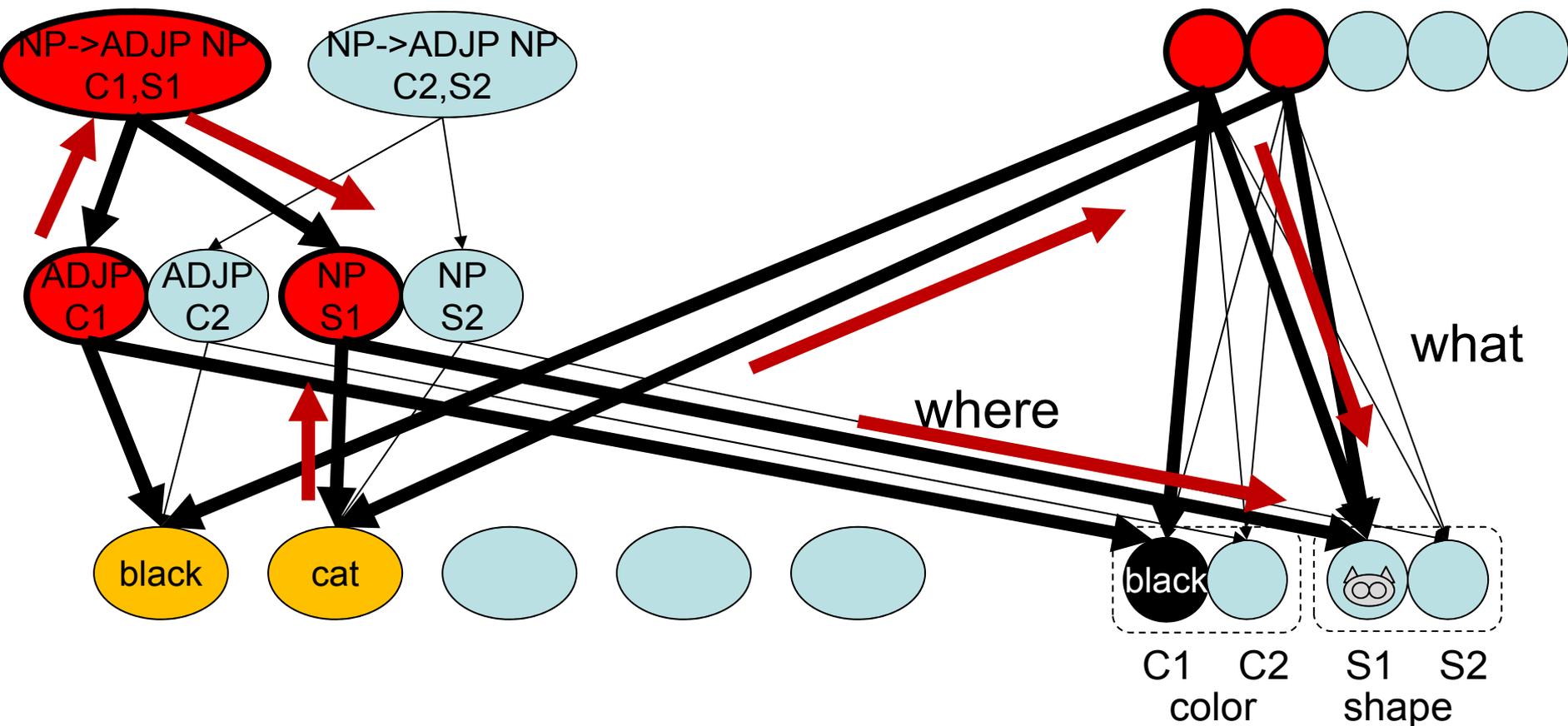
動作(1/2)

単語が与えられると、妥当な解釈のうち1つが非決定的に選択される。(選択にバイアスがあってもよい。)



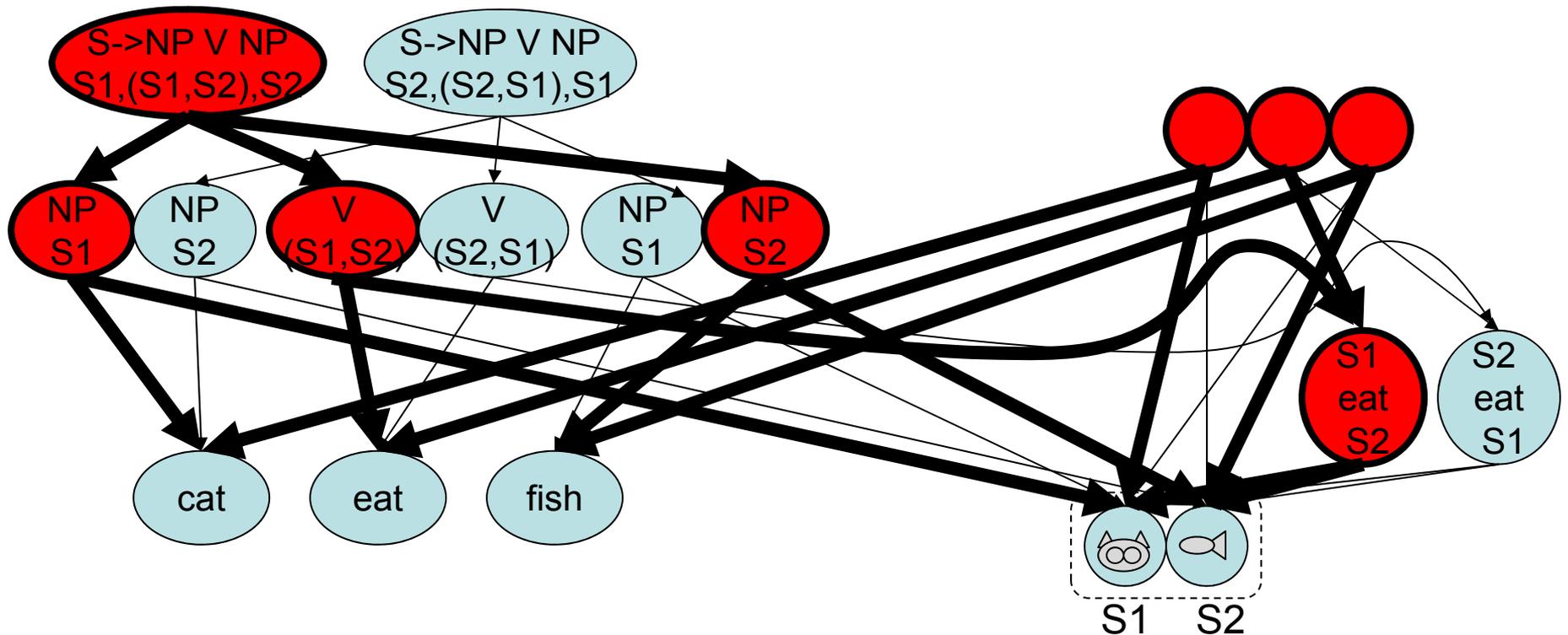
動作(2/2)

オブジェクトファイルの場所の情報が文法ノードを經由して伝播するので、正しい場所に値が書き込まれる。



動詞の扱い

- 動詞の意味とは前運動野や前頭前野の発火パターンであると考える。
- SVO文形の場合、主語と目的語の組み合わせの数 mP_2 個の違う値(発火パターン)を持つと仮定。



補足説明

- 実際には1つの生成規則が1つのノードで表現されるわけではなく、複数の生成規則が複数のノードで「スパーズ符号化」されて表現される。
- ベイジアンネットは逆方向の推論も可能なので、意味から文法にしたがって文を生成することも可能。

このネットワークは自己組織化可能か？

- マクロな構造は生物同様、あらかじめ与えられるものとする。個々のノードの意味はデータから学習可能だろうか。
- 「親ノードは子ノードの値の組を抽象化したデータベース」という原則が全てのノードに当てはまっているので、原理的には自己組織化可能であろう。
- ただし、十分な事前知識(学習のバイアス)を与えないと学習は成功しないだろう。

説明が足りない点

- "NP->ADJP NP: C1,S2" のような間違った係り受けを表現する文法ノードが作られない理由があまり自明じゃない。
 - 学習時に "ADJP:C1" と "NP:S2" が同時発火しないから、というのが理由だが、なぜ同時発火しないかの説明がいる。

ノード数に関する考察

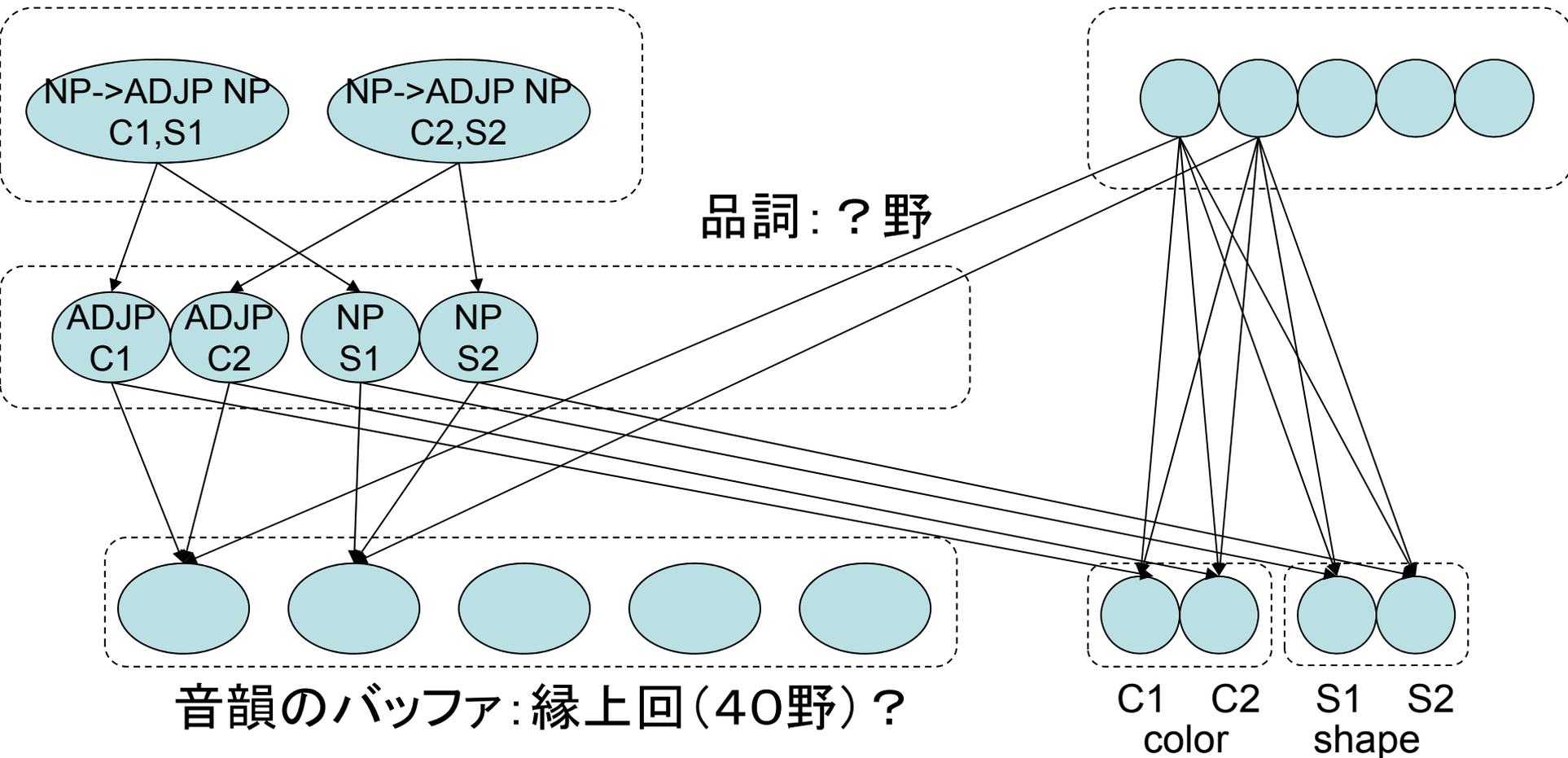
- 自然言語では n は小さいので $O(n^2)$ や $O(n)$ のノード数は許容範囲。
- ニューロン数だけでなく学習時間も $O(n^2)$ 倍になってしまいが、人間の言語習得の遅さを考えると妥当か？
 - なお、計算機上で実現する場合はたまたみ込みニューラルネットで行われる weight sharing で高速化できる。

言語野の神経科学的知見との対応

かなめとなる品詞ノードがある場所は不明。

文法:ブローカー野

単語の意味:ウェルニッケ野



言語理解に必要な BESOM の機能のまとめ

- 学習
 - 抽象化。非線形BSS。
 - オブジェクトファイル。
- 認識
 - explaining away 。
 - where/what の信号源による値の書き込み。
 - 再帰的な結合。
- いずれも BESOM での動作確認はほとんどなされていない。今後の重要課題。

将来の課題

長い文のパイプライン的な処理

- 解析の終わったところは忘却したい。
 - 貴重なオブジェクトファイルを使い回せる。
- ダイナミックベイジアンネット(時間遅れの再帰的入力を持つベイジアンネット)を使えば実現可能か？
- 縁上回(40野)を中心とする損傷で起こる伝導性失語症では、言語理解と発話はできるが文の復唱ができない。
 - パイプライン的処理の機構は保たれている？

認知科学的妥当性の検討

- このモデルは知られている言語に関する現象を説明できるか？できないとすれば、モデルにどのような拡張・制約の追加を行えばよいか？
- このモデルで知られていない現象を予言することはできるか？

より複雑な意味は脳内でどう表現される？

- 部分的な意味構造を海馬にエピソード記憶として保存し、構造の間関係は連想記憶でたどれるようにしておけばよい。海馬はヒープ領域みたいなもの。原理的にはいくらでも複雑な構造が表現できると思う。