

ベイジアンモデル選択に基づく知覚特徴量を用いた抽象的語意の学習

岩橋 直人¹ 佐藤 健² 麻生 英樹³

¹国際電気通信基礎技術研究所, ²国立情報学研究所, ³産業技術総合研究所

E-mail: ¹naoto.iwahashi@atr.jp, ²ksatoh@nii.ac.jp, ³h.asoh@aist.go.jp

あらまし 実環境中のオブジェクトやそれに加えられる動作についての知覚特徴により直接形成される構造化された基本的概念を基盤にして、知覚特徴から直接的には形成されない抽象的意味を持つ語の学習を行う計算機構を提示する。レキシコンは、知覚特徴に接地した単語の意味および単語間の関連性の情報をもつグラフィカルモデルによって表現される。計算機構は、新しく提示された単語の解釈として可能な語意をその時点のレキシコンモデルに追加することにより拡張されたレキシコンモデルの複数の候補を生成し、その中からベイズ規準に基づいた最適なモデルを選択することで、単語の意味の解釈および学習を行うものである。モデルパラメータの学習には変分ベイズ法を用いた。実験によりこの計算機構が所望の抽象的語意を適切に学習できることが示された。

キーワード 抽象的概念, 知覚, 機械学習, モデル選択, ベイズ規準, 変分ベイズ

Learning Abstract Concepts and Words from Perception Based on Bayesian Model Selection

Naoto IWAHASHI¹ Ken SATOH² and Hideki ASOH³

¹Advanced Telecommunications Research Institute International

²National Institute of Informatics

³National Institute of Advanced Industrial Science and Technology

E-mail: ¹naoto.iwahashi@atr.jp, ²ksatoh@nii.ac.jp, ³h.asoh@aist.go.jp

Abstract We address the computational mechanism that enables machines to learn words for abstract concepts not formed directly from perceived features. The learning is done based on the structured basic concepts that are formed directly from perceived features. The lexicon is represented by a graphical model that includes information about word meanings grounded on perceived features and the relationship between the words. The computation mechanism makes it possible for a machine to interpret and learn a presented new word by selecting one from among the multiple candidates for an extended lexicon, which are generated by adding the word and its possible meanings into the lexicon that has been learned before. Learning of model parameters is done by variational Bayesian methods. Experimental results show that the computational mechanism successfully learned abstract word meanings.

Keyword Abstract concept, Perception, Machine learning, Model selection, Bayesian criteria, Variational Bayes

1. はじめに

センサ技術の進歩とユビキタスコンピューティングの高度化に支えられ、ロボット、ウェアラブル・デバイス、センサ・ネットワークなどの機械が、環境(物理的な周囲の世界)や人の行動を知覚することが可能になってきた。環境から必要な情報を引き出し状況に応じて適応的に処理を行う機械は、近い将来に我々の日常生活を支援してくれる身近な存在になるであろう。

環境から独立している従来の機械に比べて、このような機械において、音声インターフェースがいっそう重要になる。それは、音声インターフェースがハンズフリ

ー・インタラクションを可能にするからだけではなく、発話はその特質として、コミュニケーションに参加する者たちで共有する知識や経験に基づいて意味を伝達することができるからである。

日常のコミュニケーションは、コミュニケーションに参加するものが互いに共有する信念に基づいて成立すると言われる。言語はそうした信念の一部であり、他の共有信念との関連に基づき、意味を伝達するために使用される [Sperber 95]。信念は、主体と環境や他者との認知的な相互作用によって形成されるので、発話の意味はそうした相互作用の中に埋め込まれる。日常生活で経験を共有する人と機械の自然な言語コミ

コミュニケーションを実現するためには、機械が、人と共有する環境を理解し、人と相互に理解し合うための計算機構を持っている必要がある。

言語獲得は、環境理解と相互理解を基盤にして、言語コミュニケーションの能力を向上させようとする相互行為であり、共有経験を基盤とする言語コミュニケーションの基本的な必須要素が明確に見られる認知的活動である。人と機械の共有経験を基盤とする言語コミュニケーションの実現を目指して、筆者らは、ロボットに言語コミュニケーション能力を学習させる計算機構の研究を行ってきた[岩橋 03]。

図1のように、人とロボットがぬいぐるみや箱などのオブジェクトを用いながら音声によるコミュニケーションを行う。ロボットは、初期状態ではオブジェクトやその動かし方についての概念とそれを指示する単語、および文法などの言語信念を持たない。人がオブジェクトを提示しながらそれに対する単語を発話して教えたり、人とロボットのそれぞれが、テーブルの上の一つのオブジェクトを動かすように要求する発話を行い、聞き手がその発話に従って行動したり、といったインタラクションを行う。このようなインタラクションを通して、ロボットは、発話を状況に応じて適切に生成・理解し、行動できる能力を学習する。たとえば、「小さい エルモ 赤い箱 のせて」という発話だけでなく、「のせて」といった省略された発話も、人とロボットが共有している経験と発話が行われた時の状況に基づいて、理解できるようにする。

しかしながら、これまでの研究で学習できた単語は、色、形、大きさによって表現されるオブジェクトの知覚特徴に関する概念、および人の動作によるオブジェクトの軌道の概念を指示するものであった。これらの概念は知覚情報から直接形成されるものである。しかしながら、我々が日常のコミュニケーションで使う単語には、知覚情報から直接には形成されない抽象レベルの高い概念を指示するものがある。

本論文では、そのような抽象レベルの高い概念を指示する単語を学習する計算機構を提示する。このような単語は、知覚情報から直接形成される基本的概念からなるレキシコンを基盤にして学習される。このレキシコンはグラフィカルモデルによって表現され、上記のインタラクションを通じて学習される。このグラフィカルモデルの構造をベイズ規準に基づいて拡張することで、抽象的な語意を学習する。

本論文の構成は次のとおりである。次節で抽象レベルの高い概念とは何かについて述べる。3節でそのような概念の形成の基盤となる構造化された基本的概念を表すグラフィカルモデルについて述べる。4節で抽象的概念を指示する単語の学習の計算機構を提示す

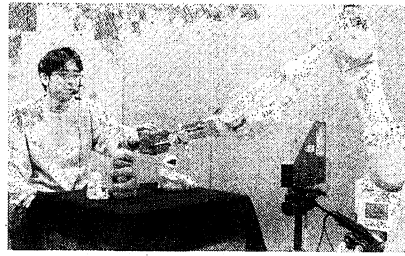


図1 人とロボットのインタラクション

る。5節で検証実験結果を示し、6節で本手法の一般的な知的処理への展開などについて考察する。

2. 抽象レベルの高い概念とは

Lakoff ら[Lakoff 87]によれば、抽象的な概念構造は、知覚や運動によって直接形成される基本的な概念の比較的単純な構造から、概念化と呼ばれるプロセスを経て形成されるという。この概念化の能力として、物理的な領域における構造から抽象的領域における構造へメタファー的に写像する能力などが関与していると考えられている。概念の全体全体構造は、各部分の相互作用により決定される。個々の抽象的な概念は、独立に形成されるのではなく、複雑な全体構造の中に埋め込まれた概念間の関係によって成立する。

ここで、抽象的な意味をもつ語の特殊性が明確にみられるチンパンジーのシンボル学習の実験[Savage-Rumbaugh 86]について簡単に紹介しよう。それはチンパンジーに抽象レベルの異なる二種類の単語記号(レクシグラム)を学習させようというものであった。第一の種類のレクシグラムは、オレンジ、パン、ケーキ、鍵、硬貨、棒などの個々の品物を表すものである。第二は、これらの品物の機能で、具体的には、『食物』、『道具』を表すものである。機能を表すレクシグラムは、対応付けられる概念が、各品物の知覚特徴から直接的に決定されるものではないという点で、個々の品物を表すレクシグラムよりも抽象レベルが高いと言える。最初、チンパンジーは、各品物をそれぞれの機能を表す容器に入れるように訓練された。この訓練後、訓練で使われなかった品物についても、高い正解率で正しい機能を表す容器に入れることができた。その後、段階的に訓練刺激を変化させて、各品物、および機能を表す各容器に対応するレクシグラムを学習させた。最終的な訓練は、各品物を表すレクシグラムを正しい機能を表すレクシグラムと対応づけさせることであった。訓練は成功し、この訓練後、訓練に使われなかった品物に対するレクシグラムをも、機能を表すレクシグラムに正しく対応付けることができるようになってい

た。

この実験は、チンパンジーがレクシグラムと対象の物的・空間的連合を学習できるだけでなく、その連合に基づいて、さらにレクシグラム間のシンボリックな参照関係をも学習する記号能力を持っていることを示している。ここで我々が注目するのは、抽象的概念である機能を表すシンボルが、各品物の知覚特徴によって直接形成される概念と、『食べる』という行為の概念、およびこれらの関係性からなる、事前にチンパンジーが持っていた構造化された概念全体を基盤にして学習されることである。

次節以降、チンパンジーのこのような抽象的な概念を指示する語の学習能力を機械で実現するための計算機構について述べる。まず、次節で、事前に学習され、抽象的語意の学習の基盤となるレキシコンを表すグラフィカルモデルについて述べる。

3. 基本概念からなるレキシコン

事前に学習されるレキシコンのグラフィカルモデル(図2)は、三種類の言語的情報を有している。オブジェクトに加える動作の概念とそれを指示する単語 W_A 、オブジェクトの知覚特徴を表す概念とそれを指示する単語 W_T , W_L 、および、これら二種類の単語の間の依存関係 $P(W_A | W_T, W_L)$ である。

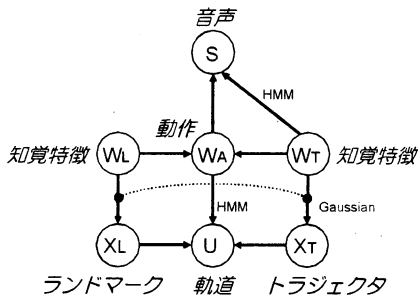


図2 基本概念からなるレキシコンを表すグラフィカルモデル

3.1. 動作を表す概念と単語

動作の概念は、動かされるオブジェクト(トラジェクタ)と基準点として働くもの(ランドマーク)の位置関係の変化によって表される。トラジェクタが同じ軌道で動いても、ランドマークの位置が違えば、異なる動作の概念になることがある。たとえば、図3のようにぬいぐるみ動かされるとき、この動きは、動かされているぬいぐるみをトラジェクタとし、真中に置かれているぬいぐるみをランドマークとみなせば、『飛び越えさせる』という概念の実現例と解釈され、右側の箱をランドマークとみなせ

ば、『のせる』という概念になる。

動作の概念は、動作を指示する単語、トラジェクタ・オブジェクトの位置、およびランドマーク・オブジェクトの位置が与えられた時、トラジェクタ・オブジェクトの軌道 U に関する確率密度関数 $p(U | W_A, X_T, X_L)$ を表すHMMによって与えられる。ここで、 U は位置(2次元)ベクトル時系列で表現される。 $W_A = [w_{A,1}, \dots, w_{A,m}]'$ は、動作を指示する単語の生起を表す確率ベクトルで、 $w_{A,i}$ は i 番目の単語を表す。 X_T, X_L は、それぞれトラジェクタ・オブジェクトとランドマーク・オブジェクトの位置情報を含む画像特徴量ベクトルである。動作を指示する単語の音声特徴量時系列の確率密度関数 $p(S|W_A)$ はHMMで表される。これらのHMMは、オブジェクトを動かしながら単語を発話することの繰り返しによって与えられる; 音声、軌道、および情景中のオブジェクトの位置の組の集合を用いて学習される [羽岡 00]。

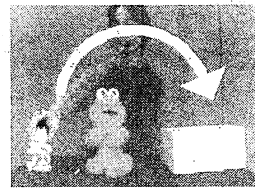


図3 動作の例

3.2. オブジェクトの知覚特徴を表す概念と単語

オブジェクトの知覚特徴を表す概念を指示する単語は、オブジェクトの知覚特徴全体を指示するもの(例えば、『エルモ』、『バナナ』)と、オブジェクトの知覚特徴の一部(色、形、大きさの属性、例えば、『赤い』、『はこ』)を指示するものからなる。これらの単語の生起を表すベクトルを $W_G = [w_{G,1}, \dots, w_{G,m_g}]'$ とおく。オブジェクトの画像特徴量は、輪郭(複素フーリエ係数、8次元)、色($L^*a^*b^*$ 、3次元)、大きさ(面積、1次元)、位置(2次元)からなる。各語が指示するオブジェクト概念を表す画像特徴量ベクトル X の確率密度関数 $p(X | w_{G,i})$ は、ガウス分布で表される。この分布を表すパラメータは、Normal-Wishart 分布で表現される。

図2の中で、 $W_T = [w_{T,1}, \dots, w_{T,m_g}]'$ と $W_L = [w_{L,1}, \dots, w_{L,m_g}]'$ は、それぞれトラジェクタ・オブジェクトとランドマーク・オブジェクトを指示する語の生起を表すベクトルで、それぞれの要素は W_G の各要素に対応する。これは、 W_G が表すすべての単語が、トラジェクタ・オブジェクトとランドマーク・オブジェクトのどちらの指示にも使うことができることを意味する。また、 $p(X_T | W_T)$ と $p(X_L | W_L)$ は、 $p(X | W_G)$ とパラメータを共有する。

これらの概念と単語は、人がオブジェクトを提示しながら単語を発話することで与えられる、音声特徴量ベ

クトル時系列 S と画像特徴量ベクトル X_T の組を用いて学習される。この学習では、ある音声 S が別と音声と同じ単語を発声したかどうかの判断が問題となるが、音声単語と画像カテゴリーの相互情報量を用いる手法 [Roy 99][金 01] や、判断の曖昧性を教示者とのインタラクションで解消する手法 [Iwahashi 04] により、効率的な学習が可能となっている。なお、図2で、 $p(S|W_A, W_T)$ は、音声 S が W_A, W_T の要素すべてが表す単語集合の中どれか一つの単語のみから生成されるという拘束付きで、 $p(S|W_A, W_T) = p(S|W_A) + p(S|W_T)$ と表される。

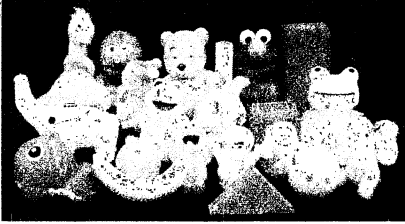


図 4 学習に用いたオブジェクトの例

3.3. オブジェクトと動作の関係

オブジェクトの知覚特徴を指示する単語と動作を指示する単語の関係は $P(W_A | W_T, W_L)$ によって表現される。この条件確率は、これは、 $P(W_A | W_T, W_L) = P(W_A | W_T) \cdot P(W_A | W_L)^{\alpha}$ と因子分解され、それぞれの確率を表すパラメータの分布は、Dirichlet 分布で表される。 $P(W_A | W_T, W_L)$ と $P(X_T | W_T)$ と $P(X_L | W_L)$ を用いて、トラジェクタ X_T およびランドマーク X_L が観測された時に、活性化される動作単語を計算することができる。 $P(W_A | W_T)$ と $P(W_A | W_L)$ は、次節で述べる抽象的語意の学習で、重要な役割を担うことになる。

4. 抽象的語意の学習の計算機構

抽象的語意の学習は、前節で述べた基本概念の単語の学習と同様に、オブジェクトの画像特徴量と音声特徴量の組を入力することで行われるものとする。

チンパンジーの実験では、『食物』と『道具』という品物の機能を指示するシンボルが学習されたが、ここで学習の対象とする機能は、各動作においてトラジェクタとなるもの、およびランドマークとなるものである。例えば、『何かにのせるもの』『何かをのせられるもの(台)』『何かに近づけるもの』『何かを近づけられるもの』『持ち上げられるもの』などである。これらの機能を指示する単語が学習できれば、例えば、『上げられるもの。台近づけて』という発話の意味を、『持ち上げられる』機能をもつオブジェクトを、『何かがのせられる』機能をも

つオブジェクトに近づけて』であると理解できるようになる。

学習のために、まず、前節で示したグラフィカルモデル(図2)に、トラジェクタ機能を指示する単語の生起を表す確率ベクトル $W_{FT} = [w_{FT,1}, \dots, w_{FT,m_a}]'$ と、ランドマーク機能を指示する単語の生起を表す確率ベクトル $W_{FL} = [w_{FL,1}, \dots, w_{FL,m_a}]'$ が加えられる(図5)。これらベクトルの要素は、動作を指示する単語の生起を表すベクトル $W_A = [w_{A,1}, \dots, w_{A,m_a}]$ の各要素に対応したものである。 $P(W_{FL} | W_L)$ と $P(W_{FT} | W_T)$ は、それぞれ $P(W_A | W_L)$ と $P(W_A | W_T)$ と、パラメータの Dirichlet 分布を共有する。この共有が、知覚特徴単語と動作単語の関係を、オブジェクト概念単語間の関係に写像し、抽象レベルの高い概念であるオブジェクト機能を表すことを可能にしている。

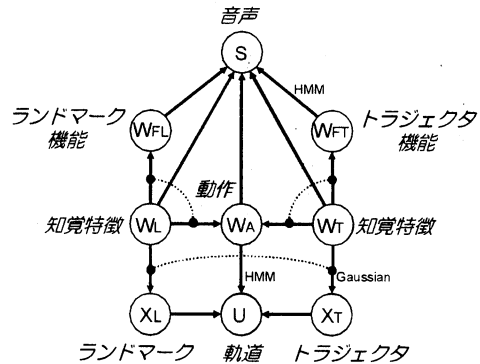


図 5 抽象的概念を指示する単語を含むレキシコンのグラフィカルモデル

学習では、オブジェクトの画像特徴量と新しい単語の音声特徴量が与えた時、その単語がオブジェクトの知覚特徴を指示する単語であるか、または、オブジェクトの機能のどれかを指示する単語であるか、わからないことが問題である。これらの概念を指示する単語をレキシコンに追加すると、それぞれ異なった構造のグラフィカルモデルが得られる。知覚特徴を指示する単語とする場合は、確率ベクトル W_T と W_L に新しい要素が追加される。トラジェクタ機能を指示する単語の場合は、新しい単語が W_{FT} の要素のうちの一つに対応付けられ、ランドマーク機能を指示する単語の場合は、 W_{FL} の要素のうちの一つに単語が対応付けられる。これらの構造のうちの一つを選択することが、新しい単語の意味を解釈することになる。こうして、単語の意味の学習をモデル選択の問題として定式化できる。

学習法の基本的な性質を調べるために、音声の代わりに単語のラベルを与えるようにした。こうすることで、

3.2節で触れた、入力音声単語を分類する問題を分離することができる。学習データ $D = \{D_X, D_L\}$ は、提示オブジェクトの画像特徴量セット $D_X = \{x^1, \dots, x^N\}$ と単語ラベルセット $D_L = \{1, \dots, N\}$ からなる。ここで、 x^i と 1^i はそれぞれ i 番目に提示されたオブジェクトの画像特徴量とそれに対応する単語ラベルである。以下、学習データ中のすべてのサンプルに、同じ単語ラベルが与えられた場合の学習アルゴリズムを示す。異なる複数のラベルが与えられた場合は、このアルゴリズムを各ラベルごとに適用すればよい。

画像特徴量 x は、語意の解釈の違いにより、 X_T または X_T の観測値とみなされる。生成されるグラフィカルモデルは、画像特徴量 x を X_T の観測値とみなし語意を知覚特徴概念としたモデル $M_{g,x}$ を X_T の観測値とみなし語意を動作のトラジェクタ機能としたモデル $M_{FT,1}, \dots, M_{FT,m_a}$ 、 x を X_L の観測値とみなし語意を動作のランドマーク機能としたモデル $M_{FL,1}, \dots, M_{FL,m_a}$ の合計 $(1 + 2m_a)$ 個である。データが入力されるごとにこれらのモデルを生成・評価し、語意が判断される。アルゴリズムは次のとおりである。

学習アルゴリズム

1. 初期化 $n \leftarrow 1$
2. $n = N$ となるまで3から8を繰り返す。
3. 画像特徴量 x^n 受け取る。
4. $(1 + 2m_a)$ 個の候補モデルを作る。
5. データセット $\{x^1, \dots, x^n\}$ を用いてこれらのモデルのパラメータを変分ベイズ法 [Attias 99] により学習する。
6. 学習された各モデルに対して、自由エネルギーを計算する。
7. 自由エネルギーが最大となるモデルを選択し、語意を判断する。
8. $n \leftarrow n + 1$

ここで、トラジェクタ機能を語意とする場合の自由エネルギーは次式で与えられる。

$$\mathcal{F}[q(Z, \theta)] = \left\langle \left\langle \log p(X, W_{FT}, W_T | \theta) \right\rangle_{q(W_{FT}, W_T)} \right\rangle_{q(\theta)} \quad (1) \\ - \left\langle \log q(W_{FT}, W_T) \right\rangle_{q(W_{FT}, W_T)} - KL(q(\theta), p(\theta))$$

ここで、 Z は非観測変数、 θ はモデルパラメータベクトルである。 $p(\theta)$ は θ の事前確率密度関数、 $q(\cdot)$ は変分ベイズ法で推定された事後確率密度関数を表す。ランドマーク機能を語意とする場合は、(1)式の中の W_{FT} と W_T をそれぞれ W_{FL} と W_L に置き換えたものとなる。

5. 実験

初期モデルとして、知覚特徴概念を指示する30単語(知覚特徴全体を指示する24単語, 知覚特徴の一部を指示する6単語)と、動作を指示する4単語(『のせる』, 『近づける』, 『飛び越えさせる』, 『持ち上げる』(これだけはランドマークがない))からなる図5で示されるモデルを学習した。 $P(W_A | W_T)$ と $P(W_A | W_L)$ の学習の際、各動作に対してトラジェクタおよびランドマークとしたオブジェクトは実験者によって選択された。それぞれのオブジェクトの集合ができるだけ形や色の類似性によって区別できないように選択したが、『のせる』のランドマーク機能(『何かかのせられるもの(台)』)のように大きくて四角い形によって特徴付けられるオブジェクト集合も作られた。また、なるべく各オブジェクトが多くの機能に対応しないようにしたが、異なる動作のトラジェクタ機能とランドマーク機能の間でいくつかのオブジェクトが共通に使われた。

抽象語意学習のための学習データとして、8種類の知覚特徴概念を指示する単語と4種類の動作のトラジェクタ機能を指示する単語、3種類の動作のランドマーク機能を指示する単語のラベルが与えられた画像特徴量セットを用いた。この画像特徴量セットは、オブジェクトのカメラからの位置や照明条件を複数設定して記録した大きなデータセットの中からランダム選択したものである。それぞれの単語は、入力データ列の中で8サンプルずつ連続して与えられた。動作を指示する単語が4個であるから、学習アルゴリズムの中で9個の候補モデルが作られる。

実験の結果、すべての単語が所望の概念として学習された。知覚特徴とトラジェクタ機能とランドマーク機能を指示する単語それぞれ一つずつについて、入力サンプル数を増やしていった時の、アルゴリズム中で生成される候補モデルの自由エネルギーの変化を図6に示す。図中、トラジェクタ機能モデルとランドマーク機能モデルについては、それぞれの候補の中で最も自由エネルギーが高くなったモデルのみの値が示されている。入力サンプル数が増えるに従い、所望の概念と他の概念が自由エネルギーの値によって、はっきりと区別できるようになってくるのがわかる。これらの結果では、画像特徴量を2サンプル与えただけで、すでに正しい概念クラスの選択を行うことができている。他の単語についてもほぼ同ような結果となった。

6. 考察

実験結果は、非常に少ない学習サンプルで所望の

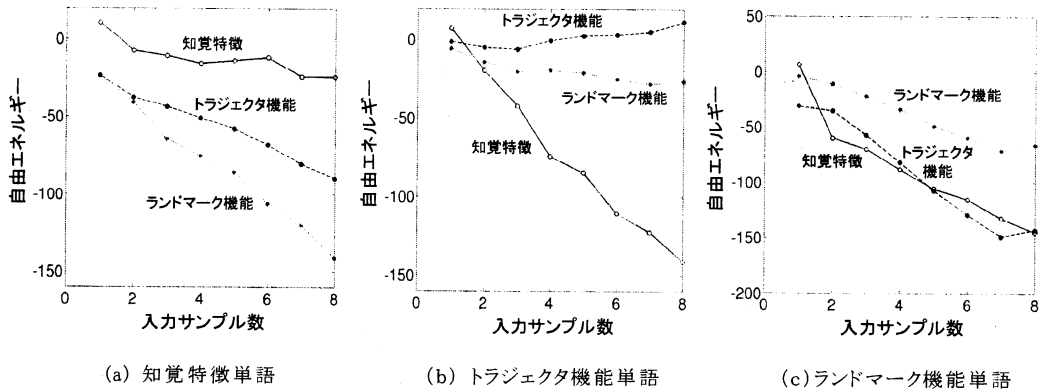


図 6 入力サンプル数の増加による自由エネルギーの変化

概念クラスの学習を可能にすることを示している。これは、事前に学習されたレキシコンモデルに表現されている知覚特徴単語と動作単語の関係が、抽象的な語意の学習に有効に利用できていることを意味する。この計算機構は、語意学習の課題に限らず、一般的な知的処理に適用できる基本原理を用いている。それは、事前の知識モデルを拡張して、新たな観測データの可能な解釈を組み込んだ知識モデルを複数生成し、その中から最適なモデルを選択することで、観測データの解釈と学習を行うときに、事前の学習で得られている変数間の関係に基づいてモデルを拡張することである。提示した計算機構は、限定された領域のスケールの小さな課題を解くものであったが、この原理をもって広い領域の大きなスケールの課題に適用できるようになれば、メタファーに基づく理解の計算機構の実現に繋がられるかもしれないと考えている。

なお、実験は、なるべく各オブジェクトが一つだけの機能に対応するようにした設定で行われており、これはこの学習課題にとって都合の良い条件である。今後、さまざまな条件で評価する必要がある。

7. まとめ

知覚情報から直接形成される基本概念を基盤として形成される抽象レベルの高い概念について述べ、そのような抽象的概念を指示する単語を学習するための計算機構を提示した。基本概念からなる構造化されたレキシコンのグラフィカルモデルを、基本概念を表す単語どうしの関係を用いて拡張することで、新しい語の解釈を表す複数のグラフィカルモデルを作り、その中からベイズ基準に基づき、最適なモデルを選択する。

謝辞 本研究は、国立情報学研究所共同研究「高次元環境知覚データにおける情報構造の発見的認識に関する研究」による研究助成を受け実施したものである。

文献

- [Attias 99] Attias, H.: Inferring Parameters and Structure of Latent Variable Models by Variational Bayes, Proc. 5th conference on Uncertainty in Artificial Intelligence,
- [羽岡 00] 羽岡哲郎, 岩橋直人: 言語獲得のための参照点に依存した空間的移動の概念の学習, 電子情報通信学会技術研究報告 PRMU2000-105, 2000.
- [岩橋 03] 岩橋直人: ロボットによる言語獲得: 言語処理の新しいパラダイムを目指して, 人工知能学会誌, 18(1), 49-58, 2003.
- [Iwahashi 04] Iwahashi, N.: Active and unsupervised learning of spoken words through a multimodal interface, Proc. 13th IEEE Workshop Robot and Human Interactive Communication, 437-442, 2004.
- [金 01] 金景柱, 岩橋直人: 知覚情報の統合に基づく階層構造を有する音声単位の獲得, 日本音響学会春季発表会講演論文集 1, 99-100, 2001.
- [Lakoff 87] Lakoff, G.: Women, fire, and dangerous things—What categories reveal about the mind, University of Chicago Press (1987), 池上嘉彦, 河上誓作訳: 認知意味論—言語から見た人間の心, 紀伊國屋書店, 1993.
- [Roy 99] Roy, D.: Learning words from sights and sounds - A computational model, Ph.D thesis, MIT.
- [Savage-Rumbaugh 86] Savage-Rumbaugh, E.: *Ape Language—From Conditional Response to Symbol*, Columbia Univ. Press, 1986.
- [Sperber 95] Sperber, D. and Wilson, D.: *Relevance (2nd Edition)*, Blackwell, 1995.