2001 年 ベイジアンネットチュートリアル
Tutorial on Bayesian Networks (BN2001)
Tokyo, Japan, July 29, 2001.

人間との対話に基づくロボットの行動知能形成
Behavior Intelligence Formation based on Communication between Human and Robots

稲邑 哲也*
東京大学大学院 情報理工学系研究科／科学技術振興事業団 CREST

Abstract: 日常空間において行動するパーソナルロボットでは，開発者が先天的な知識を与える事は難しく，自律学習を行うにしても何らかの教師や報酬が必要である．本研究では，ユーザがロボットと対話を行う事で，教示と学習の双方の利点を活用して段階的に行動知能を獲得する枠組みの実現を目指す．あるセンサパターンに対して固定的に対応した行動を割り当てる教示法では，動的な環境の変動や，ユーザの不正確な指示に対応する事ができない．そこで，ユーザとロボットの間で交された対話結果を統計的な経験として表現し蓄積する．統計的な表現により，不完全な情報からでも確率的な意思決定が可能となる．この意思決定の枠組みを，Bayesian Network で記述する手法を提案する．この機能を持つ対話学習プラットフォームとして PEXIS と呼ばれるシステムを開発し，パーソナルロボットの行動知能を形成する環境を構築した．具体的に，(1) 対話に基づく自律行動戦略の獲得とユーザ個人への適応化，(2) 意思決定の確信度を評価基準とした適切な対話文の生成，などの応用の可能性を示し，障害物回避タスク，荷物配達タスク，曖昧な指示の元での対象物指定タスクなどを通じて，その実用性を確認した．

1 はじめに

近年，人間が生活する環境に進出し，人間とコミュニケーションを行ないながら，日常生活をサポートするパーソナルロボットに注目が集まっている．ユーザとのインタラクションを行なう事を前提としたロボットは，ユーザとの対話を有効に活用する事によって，様々なブレイクスルーを生み出す事が期待される．そのような側面の中の一つの要素として，ユーザとの対話を経験として蓄積し，記憶や経験に基づいて学習や適応などを行なうアプローチがある．ロボットが使用される環境は，あらかじめ予想する事はできず，前持って知識や自律行動を埋め込む事はできない．このような問題に対して対話経験を活用するアプローチは有効な手段となる．ここで，パーソナルロボットとのインタラクションシステムに必要な要素をまとめてみる．

まず第一に,「パーソナル」という言葉が示すように，ユーザ個人の生活環境や，ユーザ固有の背景知識を考慮した行動，対話を行なう必要性である．すなわち，ロボットがユーザ個人に対して適応し，より親近感を感じる行動を実現する事が望まれる．ユーザの好みや，特殊な言い回しなど,「パーソナル」な情報を扱うためには，しばらく時間をかけてシステムとユーザが対話を行ない，経験を通じてシステムが適応する必要がある．

二つ目の要素として，ロボットが行動するためのインフラストラクチャが整っていない，未知の環境で行動しなければならない点がある．ロボットの開発者は，ロボットが使用される環境に対する事前知識がなく，あらかじめ先天的な知識や自律行動を埋め込む事ができない．ロボットが経験を通じて使用される環境知識を後天的に獲得する必要がある．

第三に，ユーザが代わっても，環境が変動しても，すばやくその変化に追従し，以前に経験した知識を再利用して適切な行動を実現する必要性である．

このような観点から本研究では，ユーザとの対話経験から後天的に環境の知識や，ユーザモデルなどを獲得し適応を行なう，パーソナルロボットのためのインタラクションシステムを提案する．また，いくつかの応用例についての実験結果を通じて有効性を検証する．

*〒 113-8658 東京都文京区本郷 7-3-1．東京大学大学院 情報理工学系研究科 tel: 03-5841-6381, e-mail: inamura@ynl.t.u-tokyo.ac.jp, URL: http://www.ynl.t.u-tokyo.ac.jp/~inamura/index-j.html

2　統計的な経験表現に基づく適応的インタラクション

2.1　Bayesian Networkの導入

何の知識も持っていないロボットに新しい知識を教示する場合，一般的な教示方法 [1] では環境情報やロボットの内部状態に応じて，取るべき行動を一対一対応させるように教示していた．これでは，環境が変動し，獲得したモデルが使用できなくなるとまた初めから教示をし直す事になる．また，人間の教示も完全である保障はなく，人間が教示ミスを行なうと，そのままミスを反映するようなモデルが形成されてしまう．

これは，本来動的に変化する環境情報と取るべき行動の関係が，固定的に表現されていたためと考えられる．この問題を回避するために，対話の結果を直接記録するのではなく，統計を用いて確率的に表現するアプローチを取り，Bayesian Network[2] と呼ばれる統計モデルを導入する．Bayesian Network は複数の事象間の因果関係を確率を用いて表現する推論モデルであり，実際に移動ロボットの地図獲得 [3][4]，コンピュータのエラー診断 [5]，対話システム [6] など，幅広く応用されている手法である [7].

このモデルの利点は，

1. ロボットと人間がインタラクションを重ねた経験から単純な統計計算を用いて，オンラインでの意思決定モデルの獲得が可能になる．
2. ニューラルネットなどの手法と比べてシンボルの扱いに適しているため，言語的な対話の結果を表現しやすく，また事象間の関係をシンボルで説明する事が可能である．
3. ユーザとロボットの間の関係，ロボットと環境の間の関係を同じレベルの表現で記述可能で，ユーザの好みや嗜好などを反映した行動モデルを構築する事が可能である．
4. 意思決定に必要な情報のうち，一部の情報が得られただけで，経験を活用してある程度の精度で意思決定を行う事が可能．

という所にある．

2.2　経験表現の定義

Bayesian Network は有向グラフの一種で，原因となる親ノードと結果となる子ノードのネットワークで構成される．各ノードにはいくつかの命題が割り振られており，ある一つの命題が成立する事でそのノードに対する状態や状況，行動などが表現される．これを命題シンボルと呼ぶ．命題の成立は確率で表現され，このために各ノードには確率変数が存在する．具体的には，図 1 のようにロボットのセンサ情報や，ロボットが取るべき行動，人間から指示された内容，などをノードとする．各ノード間の関係は確率変数を使った条件付き確率で記述される．

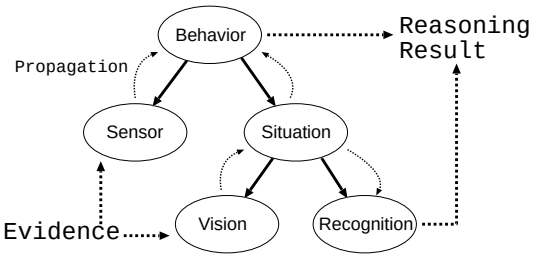

図 1: Bayesian Network におけるノード間の関係と情報の伝播

命題シンボル $b_1, \ldots, b_n$ を持つ，あるノード B に対応する確率変数を $\mathbf{B} = \{b_1, \ldots, b_n\}$ とする．そのノードの上方に接続している親ノードの集合を $\mathbf{A} = \{A_1, \ldots, A_m\}$ とし，それぞれの確率変数の各値の組み合わせによって構成される空間を，$\mathbf{a}_1, \ldots, \mathbf{a}_l$ とした場合，二つのノード間の関係を次の行列

$$CPT_{B|\mathbf{A}} \stackrel{\text{def}}{=} \mathbf{P}(B|\mathbf{A}) = \begin{pmatrix} P(b_1|\mathbf{a}_1) & \ldots & P(b_n|\mathbf{a}_1) \\ \vdots & \ddots & \vdots \\ P(b_1|\mathbf{a}_m) & \ldots & P(b_n|\mathbf{a}_l) \end{pmatrix} \quad (1)$$

で表し，これを CPT (Conditional Probability Table: 条件付き確率テーブル) と呼ぶ．この CPT はすべての親子ノード間に対して計算される．

2.3　対話経験に基づく適応化

上で定義した CPT は固定されたものではなく，ユーザとロボットのインタラクションを通じてユーザに適応して行く．この適応は，観察フェーズ，学習フェーズ，推論フェーズ，対話フェーズ，内省フェーズ，の 5 つの段階に分けられる．

観察フェーズ　基本的にロボットはユーザの指示に従って行動する．その行動の中でロボットは，どういう状況に遭遇したか，どのような指示をユーザが下したか，どのようなセンサ情報を得たか，などを観察する．観察結果は先に述べた命題シンボルに変換され，個人別のデータベースに蓄積される．

学習フェーズ　確率変数 A を持つノード A と確率変数 B を持つノード B の間に存在する因果関係 (CPT) を更新するには，単純に得られた経験の累積頻度度数を用いる．N を観察された経験データの数とし，N_j をノード A の命題シンボル a_j が観測された回数とする．また，a_j が N_j 回観測されたうちにノード B の命題シンボル b_i が観測された回数を n_{ij} とする．この場合，ノード A と B の間に存在する因果関係 $CPT_{B|A}$ の各成分 p_{ij} は

$$p_{ij} = P(B = b_i|A = a_j) = \frac{n_{ij}}{N_j} \tag{2}$$

となり，親ノードの事前確率は

$$P(a_j) = P(A = a_j) = \frac{N_j}{N} \tag{3}$$

となる．これらの「回数」は，ユーザからの指示の回数や，あるセンサ情報を得た回数などに相当する．この計算を全てのリンクと全てのノード間に対して行う．

推論フェーズ　ロボットが自律的に行動する場合，意思決定に必要なノード (一番単純な場合は行動ノード) の各命題が成立する可能性を計算し，もっとも大きい可能性を持つ命題に対応する意思決定を行う．この際システムはなるべく確らしい選択を行うように，証拠となる情報を収集し Bayesian Network に入力する．

ここで，ネットワーク全体に存在する証拠データを E とした場合，ノード X の確率は，親ノードからの影響 E_X^+ と子ノードからの影響 E_X^- の双方の影響を受けて，

$$\mathbf{P}(X|E) = \alpha\mathbf{P}(E_X^-|X)\mathbf{P}(X|E_X^+) \tag{4}$$

と表現される (表記法は [8] を参照の事)．ここで α は正規化係数である．ネットワーク構造が多重木の場合，X の親ノードを $\mathbf{U} = \{U_1, U_2, \ldots, U_m\}$ とし，X の子ノードを $\mathbf{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ として，

$$\mathbf{P}(X|E_X^+) = \sum_{\mathbf{u}} \mathbf{P}(X|\mathbf{u}) \prod_i \mathbf{P}(u_i|E_{U_i/X}) \tag{5}$$

$$\begin{aligned}\mathbf{P}(E_X^-|X) = \quad & \beta \textstyle\prod_i \sum_{y_i} P(E_{Y_i}^-|y_i) \sum_{\mathbf{z}_i} \mathbf{P}(y_i|X, \mathbf{z}_i) \\ & \textstyle\prod_j P(z_{ij}|E_{Z_i j/Y_i})\end{aligned} \tag{6}$$

となる．ただし，β は正規化係数．$\mathbf{Z}_i = \{z_{i1}, \ldots, z_{ij}\}$ はノード Y_i の親ノードのうち，X を除いたノードを意味し，$E_{U_i/X}$ は ノード U_i に関する証拠のうち，ノード X から伝播する証拠を除いたものを意味する．またここでは，$\mathbf{P}(X|\mathbf{u})$ が $CPT_{X|U}$ に対応している．

式 (5) と式 (6) はそれぞれ再帰的な計算となっており，ネットワーク構造全体に渡って情報を伝播させながら計算が進行して行く [2]．ネットワーク構造のあるノードで証拠が観測された場合，その情報が観測されていないノードまで伝播する事によって，目的のノードの確率を計算する事ができる．この計算は，ネットワークがループ構造を持っていない限り，ノードの数に比例したオーダーの計算量で済む事が分かっている．ノード X に対する確率 $\mathbf{P}(X|E)$ は $\mathbf{BEL}(X)$ とも表記され，ノードに属する命題シンボルの確率値を並べたベクトルと解釈される．これをノード X の確信度と呼び，行動決定や対話戦略の選定に使用する．

先に述べた通り，ある程度の経験を積んだ後は，ネットワークの一部の情報が得られただけで残りのノードの確信度を推論する事ができる．

対話 フェーズ　十分に学習が行われたロボットであれば問題は生じないが，観察量が少ない場合には確信度の値が不安定で，間違った行動決定を行なう事もあり得る．このような状態では，ロボットは自律行動を行ないつつユーザからの指示も受け付けられるようにするべきである．最初から完全な自律を目指す事なく，段階的に自律行動を獲得するアプローチである．

このような自律と手動の状態の混在した状況では，どちらが主導権を握るか等，人間とロボットとの間の対話が重要である．ここで，先に述べた確信度を，ユーザとの対話を制御するために使用する．

確信度の高さに応じて 4 つの自律行動レベルを設定し，各レベルに対して，ユーザの指示とロボットの決定が同じか否かの 2 つの状況を設定する．結果，合計 8 つの状況が定義され，それぞれに対話戦略を割り振る．この対話戦略は「ユーザのミスの指摘」「代換案の提案」「質問の生成」「ユーザの指示の確認」の 4 種類が存在する．例えば，ロボットの推論した命題シンボルの確信度が非常に高く，ユーザの指示がロボットの推論結果と違う場合には，ロボットはユーザがミスをしたと判断し，そのミスを指摘する．各命題シンボルには対応する自然言語が定義されており，これを用いて自然な文章を生成するようになっている．具体的な例については第 4 章で述べる．

内省フェーズ　ロボットのセンサなど連続値を扱うノードに対して，シンボルを伴う命題を対応させるために，離散化のための閾値を決定する必要がある．この閾値決定は，ユーザとの対話で得た経験を元に後天的になされるべきであり，開発者があらかじめ決められるものではない．

そこで，過去の対話の経験を利用して，閾値を決定す

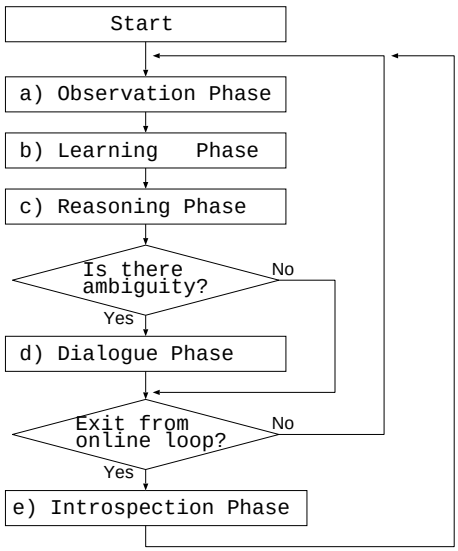

図 2: 対話の蓄積による Bayesian Network の変化の流れ

るために，推論のエラー値と呼ばれるパラメータを導入する．

過去に交された対話の際に，人間から与えられた教示内容と，その瞬間にロボットが推論した命題シンボルとの相違を意思決定モデルのエラーを表す指標として注目する．閾値のパラメータを θ とし，その条件のもとである時刻 t においてロボットが推論したノードの確信度を $\mathbf{BEL}_t(B,\theta)$ とし，そのノードに対してユーザが与えた指示を b_i とする．この時，次のベクトル

$$\mathbf{T}_t = \{T_{t1}, \ldots, T_{tn}\} \tag{7}$$

$$T_{tj} = \{\delta_{ij}\} \qquad (i = 1, \ldots, n) \tag{8}$$

をクロネッカデルタ δ_{ij} を用いて定義し，時刻 t における推論のエラー値を，二つのベクトルの各成分の差の絶対値の和

$$e_t(\theta) = ||\mathbf{BEL}_t(B,\theta) - \mathbf{T}_t|| \tag{9}$$

で表現する．このエラー値を教示が行なわれた全履歴に対して平均を取ったものを，意思決定モデルの推論のエラー値

$$E(\theta) = \frac{1}{N}\sum_{i}^{N} e_i(\theta) \tag{10}$$

とする．N は履歴データの総数である．この値が小さければ，ユーザの教示を良く反映している意思決定モデルであると判断する事ができる．

この意思決定モデルの推論のエラー値は θ の関数であり，閾値の取り得るパラメータ空間の中で，推論のエラー値が最小となる θ を求める事になる．ここで，探索空間が膨大になるため，推論のエラー値 $E(\theta)$ を適合度として，遺伝的アルゴリズム (GA) を用いて探索を行なう．

以上まとめると，このシステムは図 2 のような動作を繰り返しながら，Bayesian Network の状態を更新して行く．

3 PEXIS: the Human-Robot Interaction System

前章の考えに基づき，ロボットとユーザの間のインタラクションを行うための統合システム PEXIS (**P**robabilistic **EX**perience representation based human-robot **I**nteraction **S**ystem :統計的経験表現に基づくヒューマンロボットインタラクションシステム) を開発した [9]．PEXIS はサーバとして実行され，ロボットやユーザは図 3 のようにクライアントとしてこのシステムに接続する形式となっている．

3.1 PEXIS API

ロボットのソフトウェア開発を支援するために，PEXIS API を整備した．この API は意思決定や対話の制御などを行うための関数の集合であり，文字列による通信プロトコルにより，様々な OS や開発言語のロボットプラットフォームに対応可能である．主に次のようなコマンド関数が用意されている．

PEXIS::MakeBayesianNetwork(BN)

Bayesian Network の使用を宣言する．PEXIS では同時に複数のネットワークを登録する事が可能で，BN を ID として，使用するネットワークを指定する．

PEXIS::InputEvidence(n,p,r)

ノード n に観測データされた生データ r を入力し，同時にそれに対応した命題シンボル p を入力する．

PEXIS::QueryReasoningResult(n)

目的のノード n に対して，各命題シンボルが持つ確率，すなわち確信度を求める．

PEXIS::DecisionMaking(n)

ノード n の確信度を利用して，意思決定を行う．二つのシンボル要素に対する確率が同等な場合には，すでに述べたように対話行動がトリガされる．

PEXIS::StoreDataBase(BN, DB)

Bayesian Network BN に対して，現在得られている観測データの状態をデータベース DB に蓄積する．

PEXIS::CptRevisionOnline(BN, DB)

データベース DB に基づいて，Bayesian Network BN に含まれる CPT を更新する．

PEXIS::ThresholdRevisionOffline(BN, DB)

過去の経験 DB となるべく合致するように，Bayesian Network BN の各ノードの閾値をオフラインで調整し，ユーザに適応する．

まず，開発者はタスクに合わせた Bayesian Network を **MakeBayesianNetwork** コマンドでロボットに与える．次に，ユーザの指示と観測されたデータは **InputEvidence** コマンドを用いて Bayesian Network に証拠データとして入力される．ユーザがロボットに対して指示を行っている期間は，**StoreData-Base** と **CptRevisionOnline** によって，学習を促進させ，ロボットが自律行動を行う場合には，**QueryReasoningResult** と **DecisionMaking** コマンドを使って自律行動戦略を決定し，対話の制御を行う．この PEXIS API を用いた具体的な応用例については次章以降で述べる．

3.2 PEXIS のソフトウェア構成

PEXIS は図 3 に示すように，ロボットとユーザの間に位置するインタフェースである．PEXIS は複数のロボット，複数のユーザとの同時接続を可能としており，ロボットの種類やユーザから入力される情報の形態に依存しない．ユーザは音声による発話やキーボード，ジョイスティック，GUI などによって，ロボットに指示を出したり，質問を行なう事が可能である．PEXIS は大きく分類して以下の 4 つのサブシステムから構成され，互いに通信を行いながら全体の挙動を実現している．

(1) 自然言語処理サブシステム

ここでは日本語の自然言語に対して構文解析を行ない，各単語の意味を解析しながら，ユーザの質問に応答するためのデータベース参照コマンドや，推論を行うためのコマンドを生成し，各サブシステムを駆動する．扱う文章はローマ字表記の日本語のテキストであり，音声認識や，キーボード入力から得られたテキストをここで解析する．田中らの TQAS[10] をベースとして，オブジェクト指向 Lisp の一つである EusLisp[11] 上にこのサブシステムを構築した．

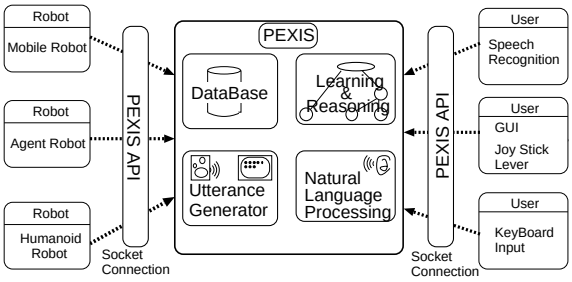

図 3: PEXIS のソフトウェア構成

また，Bayesian Network の各ノードに割り振られている命題シンボルと自然言語の関係の記述もこのサブシステムで行われている．

(2) データベースサブシステム

ここでは，ユーザとロボットの間で交わされた対話の結果や得られたセンサ情報を時系列の経験データとして蓄積する．またロボットが行動を行なわないオフライン時には,「内省フェーズ」で述べたように，蓄積された経験を内省する事によって，ユーザ個人に適応した閾値を決定する．

(3) 学習/推論サブシステム

Bayesian Network による計算を主に担当する．学習時には，センサ，行動，概念等の各ノード間の因果関係を式 (2),(3) を用いて学習し，CPT として表現する．推論時には，式 (4)〜(6) を用いて，経験データからノードの確信度を求める．十分確信度の高い命題シンボルは接続されているロボットに送信され，ロボットが自律行動を行なうトリガとなる．また，最大の確信度が二つあるなど，意思決定が難しい場合には，ユーザとの対話を開始させるトリガをロボットに対して発生する．

(4) ユーザインタフェース部

PEXIS を利用するユーザの人数分用意される．音声認識システムから入力された言語情報を PEXIS に送り，PEXIS から送られてきた発話文章を音声合成でユーザに提示する．また，ジョイスティックレバーや GUI などを扱う機能も持ち合わせている．PEXIS との通信は Socket を用いており，携帯端末と無線 LAN などを使用すれば，ユーザは場所を限定されずにロボットと対話を行なう事ができる．

4 応用例:対話に基づく行動戦略の獲得

4.1 障害物回避タスクへの応用

まず，人間とロボットとの対話に基づいて行動戦略を獲得する例について説明する．自律行動の例題として，移動ロボットにおける障害物回避タスクを選定した．こ

れは古典的なタスクであり，ニューラルネットを用いて学習する手法 [12]，強化学習を用いる方法 [13] など，数多くの研究がなされてきた．しかし，人間とロボットが対話を行なう事によって，行動を獲得する事例についてはほとんど報告がなされていない．対話に基づく行動獲得の例として，[14] があるが，行動のシーケンスを教示する程度にとどまっている．

ロボットが最初に持っている知識はセンサノード $S = \{Near, Middle, Far\}$ と行動ノード $B = \{Forward, Left, Right$ の二種類のノードから構成される，図 4 のような Bayesian Network のみである．センサとモータコマンドのマッピング，すなわち行動決定に関する戦略は全く持たない．

この状態では，ロボットは何の行動も取る事ができない．そのため一番最初の段階では，ユーザがロボットに対して教示を行なう．今回使用したロボットは図 6 左のような仮想環境のロボットである．センサとして 0〜255 までの値を取る距離センサを 8 個持ち，図 5 のような廊下環境で，障害物を回避しながら Start から Goal まで走行する事をタスクとする．また，このロボットは図 6 右に示すようなユーザインタフェース端末に接続され，ジョイスティック操作や，音声による指示によって遠隔操縦する事が可能になっている．人間からの操縦指示は **InputEvidence** コマンドで行動ノードに入力される．

ロボットは教示データを受け取った段階で，その瞬間のセンサデータと共に，**StoreDatabase** コマンドで経験データベースに蓄積し始める．そして **CPTRevisionOnline** コマンドでセンサノードと行動ノードの間にある条件付き確率をオンラインで更新する．そのため，ロボットは少々操縦を行なった後に，ある程度の行動決定を行なう事ができるようになる．この行動決定は Bayesian Network による推論出力，すなわち行動ノードの確信度 (前進，左折，右折，停止のそれぞれの確率値) で表現される．

走行回数の増加と共に確信度が高まって行く様子を図 7に，対話の例を表 1 に示す．この実験では，1 回目の走行は完全にジョイスティックによる教示走行を行なっている．2 回目の走行から自律行動が開始されているが，図 7(1) では前進と右折の二つの行動の確信度が近く，混乱している事がわかる．3 回目の走行からほぼ，どのシー

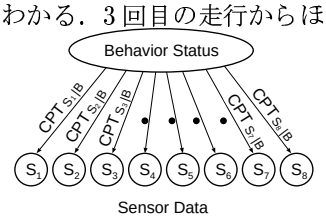

図 4: 障害物回避タスクに使用する Bayesian Network

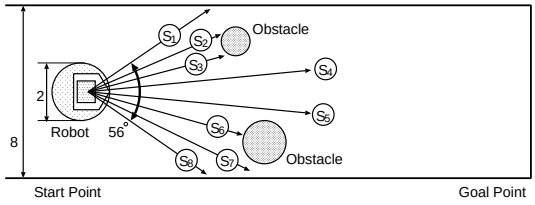

図 5: 移動ロボット上の距離センサの配置

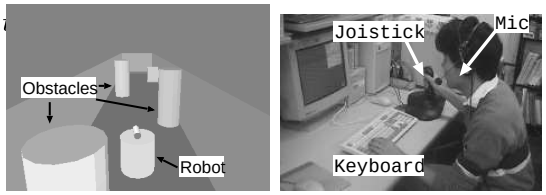

図 6: 仮想環境内の移動ロボットとユーザインタフェース

ンにおいても適切な行動が選択されている事がわかる．また実機ロボットでも同様に行動獲得可能な事を確かめている [15]．

4.2 内省フェーズによる個人への適応

障害物回避タスクでは，センサノードの確率変数を $S = \{Near, Middle, Far\}$ としていたため，連続値のセンサを離散化するのに二つの閾値が必要になる．この閾値を個人に合わせて決定するために「内省フェーズ」による更新を行う．

(1) 障害物が遠くても安全に早めに回避を行なう操縦．(2) なるべく直進を選択し，ぎりぎりまで回避しない操縦．の二種類の操縦を意図的に行なった．閾値があらかじめ決定されている場合は，このような操縦の違いを反映する事はできないが，**ThresholdRevisionOffline** コマンドにより，図 8 のようにユーザの操縦を反映した行動を獲得する事が可能となった．この場合，閾値は全てのノードで共通の値を取る状態からスタートし，(1),(2) それぞれの走行を 5 回づつ行った後で，図 9 のようなセンサノードの閾値を得た．中央のセンサと端のセンサで役割分担が見られている．この事からも適切な閾値がユーザとの対話を通じて獲得されている事が分かる．

5 応用例:個人に適応した語彙の獲得

実環境で作業を行うロボットを考えた場合,「その机の上の赤い本を取って」などと，言語による物体の参照が行われる事が多い．例えば，色の名前を使って発話を行なう場合，同じ色であっても人によって表現のし方が違う場合がある．本章では，このような問題に対して PEXIS を応用し，色の名前と画像処理上の色パラメータの間の関係を対話を通じて後天的に獲得させる例につ

表 1: 障害物回避タスクにおける対話例

ロボット「どうすれば良いか，教えて下さい」
ユーザ　（ジョイスティックレバーで教示）
ロボット「直進かもしれませんが，右折します」
ロボット「左折かも知れませんが，直進します」
ユーザ　「ちょっと左」
ロボット「本当に左折で良いのですか？」
ユーザ　「はい」
ロボット「分かりました」

表 2: 各状況別の対話戦略

The highest certainty factor & Autonomy Level	user's instruction	
	Difference Case	No instruction
80∼100 [%] Autonomy Level 4	Point out the user's mistake	Execute inferred options
65∼79 [%] Autonomy Level 3	Make a suggestion of the inferred options	Execute reasoning results with making suggestions as the inferred options
50∼64 [%] Autonomy Level 2	Obey user's inst-ructions while showing the inferred options	Execute reasoning results while showing inferred options
0∼49 [%] Autonomy Level 1	Obey user's instructions	Wait for user's operations while showing inferred options

いて説明する．

具体例として，図 10 のような上半身が人間型であるロボットを用いて，机の上にある複数の物体から，ユーザが指定した物体を取り上げる「視覚探索タスク」に本システムを応用した．このタスクのために，表 3 に示すような視覚特徴のパラメータと人間がそれを表現する時の名前等のノードを用意し，図 11 のネットワークを構築した．実際には表 4 のような対話を行いながら，目的の物体を特定して行く．具体的な対話の生成法などについては紙面の都合上 [16][17] にゆずる事とし，ここでは色の名前と視覚パラメータの関係の獲得，適応について述べる．

ユーザが「あの赤い物を取って」などの表現をした場合，ユーザが使用した表現が Color-Name(CN) ノードに入力される．ユーザが使用した色の名前を証拠とし，視覚パラメータを意味する Color-Parameter(CP) ノードに対して推論を行なう事により，そのユーザが指示したい物体の視覚パラメータ (Hue と Saturation) を得る．これにより候補となっている複数の物体から，ユーザが指示している可能性の高い物体を選択する．参照物の同定が成立すると，画像処理によって得られた視覚パラメータが CP ノードに入力され，ユーザが使用した表現と共に経験データに蓄積される．このデータベースを元に，確率変数 CP と CN の間の CPT を学習する．

また，個人の色の表現に適応するため，閾値を **ThresholdRevisionOffline** コマンドで修正する．修正の対象となるのは，図 11 下に示す CP ノードの閾値である．Hue と Saturation から構成される空間を 13 個の閾値で，$\{cp_0, \ldots, cp_{12}\}$ のエリアに分割しているため，$\theta = \{\theta_1, \ldots, \theta_{13}\}$ をユーザに対して適応させる事になる．

初期状態として，$CN = \{Red, Blue, Yellow, Green\}$，

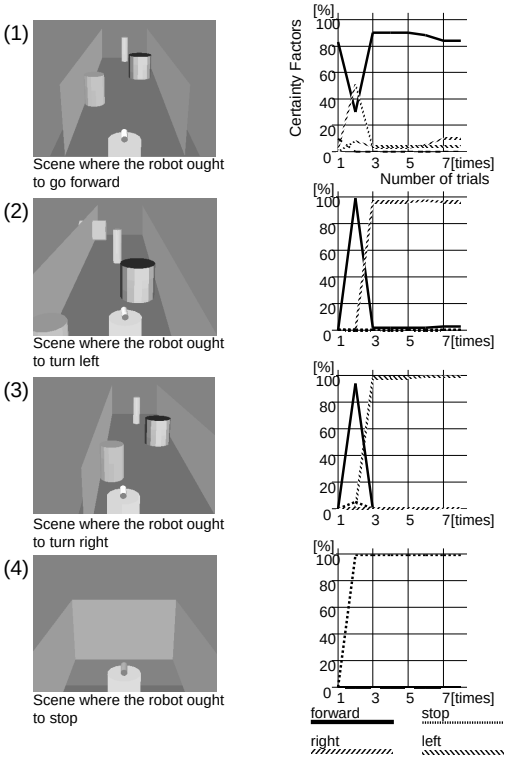

図 7: 仮想環境における学習の進み

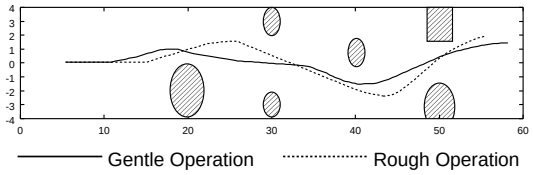

図 8: 異なる操縦を行ったユーザへの適応結果

ノード CP は Hue に対して均等に 12 分割，という状態から，被験者に自由に色の名前を使用して良い条件の元でロボットとのインタラクションを行なってもらった．ユーザが新しい個人特有の色の名前を使用した場合には，CN ノードに新しい命題シンボルを加える事で対応する．約 50 回のインタラクションを行なった後に獲得された色の名前と物理パラメータ間の関係を図 12 に，視覚パラメータの閾値を被験者に適応させた結果を図 13 に示す．

図から分かる通り，Hue のパラメータ空間の中で色の名前は均一に分布しているわけではなく，ある程度違う Hue の値でも同じ色の名前で表現する領域もあれば，わずかな Hue の差で名前が変わる領域も存在する．この事からも，個人に適応した色の表現方法を獲得している事が確認できる．

ここでは色の表現に関して説明したが，「大きい」や「右の」など，大きさや位置についての表現が入力され

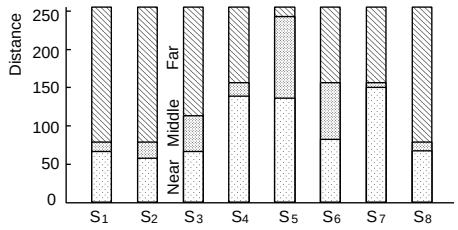

図 9: センサノードの閾値の適応結果

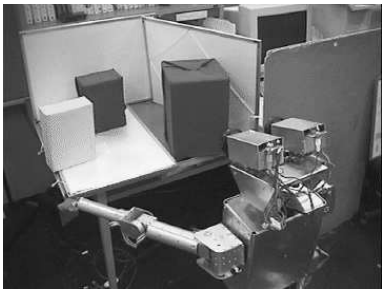

図 10: 曖昧な状況での指示実験

た場合にも，図 3 の中央や右のノードを用いて同様の適応が行われる．

6 関連研究

岡田らは，移動ロボットに対して曖昧な自然言語で教示を行なっても，環境の状況に応じて行動を変動させる学習システムを提案している [18]．しかしながら，このシステムはユーザが常に教示情報を与えなければならず，パーソナルロボットとしての自律行動の条件を満たしているとは言えない．

長尾 [19] は，実世界指向インタフェースの概念を提唱し，ユーザと実環境を結合するインタラクションシステムの重要性を示している．パーソナルロボットにおいても同様の事が言えるが，長尾のシステムはエージェントはユーザの手の中にある携帯端末であり，自律的に行動するロボットを考慮すると，インタラクションの機構は大きく変わる．本システムはエージェントに相当するパーソナルロボットが自律的に実世界の中を行動するために必要な要素を含んでいる．

松井ら [4][20] は事情通ロボットと呼ばれるオフィスサービスロボットを実現している．このロボットも対話や学習機能を持っているが，ユーザ個人への適応については論じられていない．本論文で述べた PEXIS はオフィス環境から，パーソナル環境に進出し，ユーザに適応するためのシステムであると位置付けられる．

従来の「経験を利用する」枠組みは，コンピュータとユーザ間の経験 [21]，ロボットと環境の間の経験 [13]，

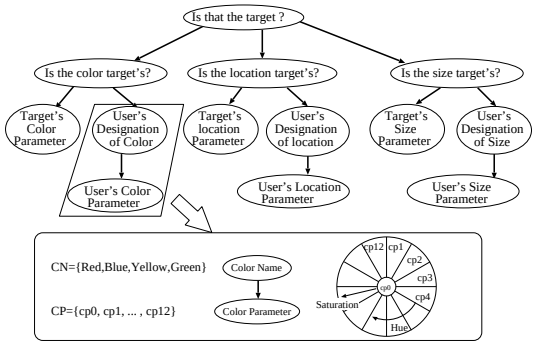

図 11: 視覚探索タスクに使用する Bayesian Network

表 3: 視覚探索タスクで使用される確率変数

Name of nodes	Random Variables
Color-Parameter	$CP = \{cp_0, cp_1, \ldots\}$
Color-Name	$CN = \{Red, Blue, \ldots\}$
Location-Parameter	$LP = \{Lp_0, Lp_1, \ldots\}$
Location-Name	$LN = \{Near, Middle, Far\}$
Size-Parameter	$SP = \{Sp_0, Sp_1, \ldots\}$
Size-Name	$SN = \{Small, Middle, Large\}$

などを扱うものが多かったが，パーソナルロボットの枠組みでは，ロボット，ユーザ，環境の 3 者を同時に扱うモデルが必要となる．本論文で示したシステムは Bayesian Network のシンボル表現を用いて，これら 3 者を扱う事が可能であり，よりパーソナルロボットに適した手法であると言える．

7 結論と展望

実世界でユーザとインタラクションを行ないながら，経験を蓄積し，学習を行ない，知識を獲得する，などのパーソナルロボットに求められる機能を統合するために，統計的に経験を表現し，Bayesian Network を用いて適応的なインタラクションを実現するシステムについて述べた．このシステムはインターネット上のサーバとして立ち上がっているため，複数のユーザに対してサービスを行うロボットや，ネットワークロボットにおける遠隔操縦に基づく行動学習，など応用できる領域は広い．

課題点として，各タスクに応じた Bayesian Network を自動的に構成する方法を確立する事と，少ない回数のインタラクションで最大限の学習を行う効率性の向上がある．前者の問題に対しては，ネットワークの自動構築の大きな指針としては，事象間の相互情報量が有効であると考えている．また後者の問題に対しては，状況に応じて経験に重みを付ける事により，重要な経験と軽視して良い経験を表現し，効率よい学習を行う試みを行っている [15]．

表 4: 曖昧性の解消を行なう対話例

ユーザ	「それを取って」
ロボット	「それはこの，黄色い方ですか？」
	（手を伸ばしながら）
ユーザ	「違います，赤い箱の事です」
ロボット	「それはこの，右の手前の方ですか？」
	（手を伸ばしながら）
ユーザ	「はいそうです」

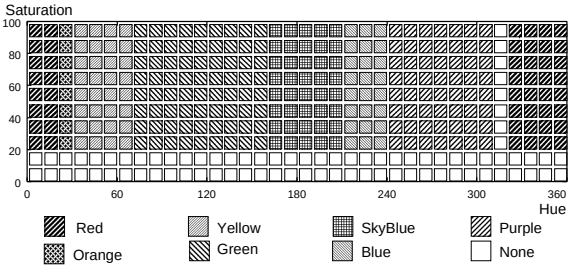

図 12: 獲得された色の名前と視覚パラメータとの関係

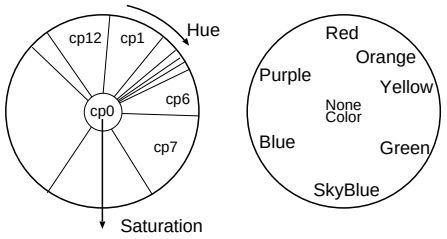

図 13: 獲得されたノード CP の閾値

参考文献

[1] Yasuo Kuniyoshi, Masayuki Inaba, and Hirochika Inoue. Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance. *IEEE Transaction on Robotics and Automation*, Vol. 10, No. 6, pp. 799–822, 1994.

[2] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[3] Kenneth Basye, Thomas Dean, Jak Kirman, and Moises Lejter. A Decision-Theoretic Approach to Planning Perception and Control. *IEEE Expert*, Vol. 7, No. 4, pp. 58–65, 1992.

[4] H. Asoh, Y. Motomura, I. Hara, S. Akaho, S. Hayamizu, and T. Matsui. Acquiring a Probabilistic Map with Dialogue-Based Learning. In *Proceedings of ROBOLEARN-96*, 1996.

[5] David Heckerman, Abe Mamdani, and Michael P. Wellman. Real-World Applications of Bayesian Networks. *Communicaiton of the ACM*, Vol. 38, No. 3, pp. 24–57, March 1995.

[6] Tomoyoshi Akiba and Hozumi Tanaka. A Bayesian Approach for User Modeling in Dialogue Systems. Technical report, Dept. of Computer Science, Tokyo Institute of Technology, 1994.

[7] Eugene Charniak. Bayesian Networks without Tears. *AI Magazine*, Vol. Winter, pp. 50–63, 1991.

[8] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A modern Approach*. Prentice Hall, 1995.

[9] 稲邑哲也, 稲葉雅幸, 井上博允. PEXIS:統計的経験表現に基づくパーソナルロボットとの適応的インタラクションシステム. 電子情報通信学会誌, Vol. J84-D-I, No. 6, pp. 867–877, 2001.

[10] 田中穂積, 元吉文男, 山梨正明. LISP で学ぶ認知心理学 3 ：言語理解. 東京大学出版会, 1983.

[11] Toshihiro Matsui. Multithread Object-Oriented Language EusLisp for Parallel and Asynchronous Programming in Robotics. In *Workshop on Concurrent Object-based Systems, IEEE 6th Symposium on Parallel and Distributed Processing*, October 1994.

[12] Dean A. Pomerleau. *Neural network perception for mobile robot guidance*. Kluwer Academic Publishing, 1993.

[13] Minoru Asada, Eiji Uchibe, and Koh Hosoda. Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development. *Artificial Intelligence*, Vol. 110, No. 2, pp. 275–292, Jun 1999.

[14] Tomohiro Shibata, Yoshio Matsumoto, Taichi Kuwahara, Masayuki Inaba, and Hirochika Inoue. Development and integration of generic components for a teachable vision-based mobile robot. *IEEE/ASME Transactions on Mechatronics*, Vol. 1, No. 3, pp. 230–236, 1996.

[15] Tetsunari Inamura, Masayuki Inaba, and Hirochika Inoue. Acquisition of Probabilistic Behavior Decision Model based on the Interactive Teaching Method. In *Proceeding of the 9th International Conference on Advanced Robotics*, pp. 523–528, 1999.

[16] 稲邑哲也, 稲葉雅幸, 井上博允. 個人に適応した語彙を獲得するロボットとの自然言語対話処理–ヒューマンロボットインタラクションシステム PEXIS の機能その 3–. 第 5 回ロボティクスシンポジア予稿集, pp. 146–151, 2000.

[17] 稲邑哲也, 岡部真人, 稲葉雅幸, 井上博允. タスク達成の確信度に基づくユーザへの質問確認行動の生成. 日本機械学会ロボティクスメカトロニクス講演会 '00, pp. 2A1–76–112, 2000.

[18] 岡田豊史, 開一夫, 安西祐一郎. ロボットコマンド学習システム Acorn-II とその評価. 人工知能学会誌, Vol. 9, No. 6, pp. 882–889, 1994.

[19] 長尾確. マルチモーダル・ヒューマンコンピュータインタラクション–エージェント指向と実世界指向–. 計測自動制御学会誌, Vol. 35, No. 1, pp. 65–70, 1996.

[20] 松井俊浩. おせっかいロボットとも呼ばれる事情通ロボットの計画. *bit*, Vol. 29, No. 12, pp. 4–11, 1997.

[21] Pattie Maes. Agents that Reduce Work and Information Overload. *Communications of the ACM*, Vol. 37, No. 7, pp. 30–40, 1994.