

対話戦略最適化としての発話理解確信度の能動学習

Active Learning of Confidence Measure Function as Dialogue Strategy Optimization

杉浦 孔明*1
Komei Sugiura岩橋 直人*1*2
Naoto Iwahashi

*1(独) 情報通信研究機構

National Institute of Information and Communications Technology

*2(株) 国際電気通信基礎技術研究所

ATR

This paper proposes a method for learning confidence measure function in the framework of active learning. A robot's utterance such as "Raise Kermit" and a user's response to it is viewed as collecting training samples. For training the confidence measure function, the method selects the optimal utterances based on expected log loss reduction. Bayesian logistic regression is used for the learning. We evaluated the method with an object manipulation dialogue task in which a robot was supposed to manipulate object as the user said. The experimental results revealed that the proposed method reduced the risk of motion execution failure.

1. はじめに

高齢化社会の到来とともに、生活環境で人間を支援するロボットへの期待が高まっている。生活支援ロボットにとってユーザとのコミュニケーション機能は極めて重要であるが、現状の対話処理技術は必要なレベルに全く到達していない。

さらに、従来の対話技術には安全性上の観点から大きな問題がある。それは、ユーザの発話の意味が適切に理解されずに、ロボットが予期しない動作を行ってしまう危険性があることである。

本研究では、この危険性を減少させることを目的とする。具体的なタスクとして、物体操作対話タスクを扱う。物体操作対話タスクとは、ユーザが発話によりロボットにオブジェクトを操作させるタスクを指す。物体操作対話タスクにおいて、ユーザの発話の意味が適切に理解されるためには、(1) 言語による動作参照、(2) 言語によるオブジェクト参照、における曖昧性を解消する必要がある。

(1) の曖昧性解消に対して多くの研究 (例えば [山肩 04, Roy 02]) が行われてきたものの、(1)(2) をともに解消する研究は今までにない。一方我々は、実世界にグラウンドした動作のイメージをユーザとロボットが共有する手法 LCore を提案している [Iwahashi 07]。さらに [杉浦 09] では、発話理解確信度を最尤推定に基づいて学習する手法を提案し、ユーザの発話が曖昧である場合に確認発話を行って、動作失敗率を減少させることができることを示した。

本論文では、これらの手法を拡張し、能動学習の枠組みにより確信度を学習する手法 LCore-AL を提案する。今まで我々が提案してきた手法では、ユーザの発話を受けてロボットが行動し、その結果をもとに確信度の学習を行っていたため、学習フェーズで動作失敗 (負事例) を行わせて、過学習を防ぐ必要があった。しかし、学習フェーズにおける動作失敗であってもハードウェアの故障などを引き起こす可能性があるため、ユーザの予期しないような動作失敗は減少させることが望ましい。これに対し本手法では、確信度関数の学習に必要な発話をロボットに選択させ、その発話を受けてユーザが物体を操作する*1。発話の選択には、能動学習 [渡辺 05] の一種である

Expected Log Loss Reduction (ELLR) [Roy 01] を用いる。

提案手法の独自性は、(1) 対話システムへの能動学習の導入、(2) 確信度の学習に Bayesian Logistic Regression (BLR) [Genkin 07] を用いること、の2点である。これまで一般の対話システムでは、ユーザの発話を受けてシステムが応答を生成し、ユーザが発話を理解するモデル (ユーザモデル) を学習させることは可能であったものの、システムの発話によりユーザモデルを学習させることは困難であった。一方、本手法では、ロボットとユーザが実世界にグラウンドした対話を行うため、ロボットが「ユーザが用いる語彙」を用いてユーザに動作応答させることができる。このような特徴から、対話戦略として能動学習を用いることが初めて可能になった。

2. LCore における発話理解

LCore [Iwahashi 07] では、マルチモーダル入力から学習されたユーザモデルを用いてユーザの発話を理解する。本論文では、音声・画像・動作などの各モダリティに対応するユーザモデルを信念モジュールと呼ぶ。また、(1) 音声、(2) 動作、(3) 視覚、(4) 動作-オブジェクト関係、(5) 行動コンテキスト、の5つの信念モジュールを統合したユーザモデルを共有信念 Ψ と呼ぶ。

2.1 物体操作対話タスク

図1に、本研究で扱う物体操作対話タスクを行うロボットの例を示す。いま、ユーザがロボットに対して、オブジェクト1(バーバブライト)をオブジェクト2(赤い箱)にのせるように発話で指示したとする。このときのユーザが意図する軌道を白線で表す。本手法では、ロボットの動作生成手法として、[Sugiura 07] で提案した隠れマルコフモデル (HMM) に基づく手法を用いる。この手法では、物体操作軌道がトラジェクタ (動かされるオブジェクト) と参照オブジェクトとの相対軌道としてモデル化される。参照オブジェクトとは、動作の基準となるオブジェクトのことを指し、トラジェクタそのもの、あるいはランドマーク (トラジェクタの動きの基準となるオブジェクト) から選択される。図1の場合、トラジェクタはオブジェクト1、参照オブジェクトはオブジェクト2である。

連絡先: 杉浦孔明, 京都府相楽郡精華町光台 2-2-2

*1 本手法では、ユーザが「動作失敗」することがあるが、ユーザによる物体操作の失敗で、物を壊す等の被害が発生することは (ロボッ

トが操作する場合に比べて) 稀である。



図 1: 物体操作対話の例

2.2 共有信念モデルに基づく発話理解

ユーザの発話 s は、トラジェクタを表す文節 W_T 、ランドマークを表す文節 W_L 、動作を表す文節 W_M からなる概念構造 $z = (W_T, W_L, W_M)$ と対応づけて解釈される。例えば、図 1 に示すシーンにおいて、ユーザが「バーブプライト、赤い箱のせて」と発話したとする。このとき、正しく分割された文節は以下のようになる。

W_T : バーブプライト, W_L : 赤い箱, W_M : のせて

ただし本手法では、 s に含まれる動詞の活用形は全て命令形であり、音声認識時に助詞を扱わないこととする。また、ランドマークを必要としない動作概念では、 $z = (W_T, W_M)$ である。

いま、シーン O において発話 s が与えられたとしよう。 O は、カメラ画像中の全オブジェクトの画像特徴量および位置を表す。 O において可能な動作の集合 A は以下により与えられる。

$$A = \{(i_r, i_l, C_V^{(j)}) \mid i_t = 1, \dots, O_N, i_r = 1, \dots, R_N, j = 1, \dots, V_N\} \\ \triangleq \{a_k \mid k = 1, 2, \dots, |A|\}, \quad (1)$$

ここに、トラジェクタのインデックスを i_t 、参照オブジェクトのインデックスを r 、 O 中のオブジェクトの数を O_N 、動作を表す単語数を V_N 、動詞 $C_V^{(j)}$ に対して可能な参照オブジェクトの数を R_N とする。従って、物体操作対話タスクでは、 s に対し正しい a_k を選択することが求められる。

各信念モジュールを以下のように定義する。まず、音声信念 B_S を発話 s に対する z の条件付き確率の対数として表す。視覚信念 B_I は、オブジェクト i の視覚特徴量 $\mathbf{x}_I^{(i)}$ に対する確率モデルの対数尤度である。同様に、動作-オブジェクト関係信念 B_R は、オブジェクト (i, j) の視覚特徴量に対する確率モデルの対数尤度である。 a_k に対する最尤軌道を $\hat{\gamma}_k$ とすると、動作信念 B_M は、トラジェクタ i の位置 $\mathbf{x}_p^{(i)}$ が与えられたうえでの $\hat{\gamma}_k$ に対する L の対数尤度で表される。行動コンテキスト信念 B_H は、コンテキスト $\mathbf{q}^{(i)}$ のもとでの、指示対象としてのオブジェクト i の適切さ (スコア) を表す。コンテキストの例としては、「オブジェクト i が把持されている」、「直前に操作された」などが挙げられる。

以上より、共有信念関数 Ψ を、各信念モジュールの重み付き和として定義する。

$$\Psi(s, a_k, O, \mathbf{q}^{(i)}) = \max_z \left\{ \begin{aligned} &\gamma_1 \log P(s|z) \\ &+\gamma_2 \log P(\hat{\gamma}_k | \mathbf{x}_p^{(i)}, \mathbf{x}_p^{(i)}, C_V^{(j)}) \\ &+\gamma_3 \left(\log P(\mathbf{x}_I^{(i)} | W_T) + \log P(\mathbf{x}_I^{(i)} | W_L) \right) \\ &+\gamma_4 \log P(\mathbf{x}_I^{(i)}, \mathbf{x}_I^{(i)} | C_V^{(j)}) \\ &+\gamma_5 \left(B_H(i, \mathbf{q}^{(i)}) + B_H(i_r, \mathbf{q}^{(i_r)}) \right) \end{aligned} \right\}, \quad (2)$$

ここに、 $\mathbf{x}_p^{(i)}$ はオブジェクト i の位置、 $\gamma = (\gamma_1, \dots, \gamma_5)$ は、各信念に対する重みを表す。 γ の学習には、Minimum Classification Error 学習 [Katagiri 98] を用いる。 Ψ により、発話 s と行動 a_k の対応の適切さを評価することができる。

3. 発話理解確信度の能動学習

3.1 統合確信度による発話理解確率のモデル化

本節では、まず共有信念関数に基づく曖昧性の尺度について説明する。

前節の共有信念関数を用いると、コンテキスト q 、シーン O 、発話 s が与えられたときの最適行動 \hat{a}_k は以下で得られる。

$$\hat{a}_k = \operatorname{argmax}_{a_k \in A} \Psi(s, a_k, O, \mathbf{q}) \quad (3)$$

行動 a_j と、最適行動 $\hat{a}_k (k \neq j)$ のマージンを以下の関数 d により定義する。

$$d(s, a_k, O, \mathbf{q}) = \Psi(s, a_k, O, \mathbf{q}) - \max_{j \neq k} \Psi(s, a_j, O, \mathbf{q}) \quad (4)$$

いま、最大値の次に大きい値を与える行動を a_l とする。式 (4) より、最適行動 \hat{a}_k に対するマージンは \hat{a}_k と a_l の共有信念関数の値の差であることがわかる。よって、 \hat{a}_k に対するマージンが 0 に近ければ、発話 s は \hat{a}_k と a_l を指示する発話として同程度に適した表現であるといえる。逆に、マージンが大きい場合には、 \hat{a}_k の方が s の指示する行動として適している。従ってマージン関数は、行動 \hat{a}_k を指示する発話としての s の曖昧性の尺度として用いることができる。

ここで、マージンを用いて \hat{a}_k に対する確信度を得ることを考える。音声認識の分野では、認識結果に対する確信度を導入することにより、発話を棄却するか否かを制御する研究が行われている [Kawahara 98]。また、確信度は、動作認識や音声認識において新規動作や未登録語の検出にも用いられている。

提案手法では、統合確信度関数 $f(d)$ をシグモイド関数を用いて以下のように定義する。

$$f(d; \mathbf{w}) = \frac{1}{1 + \exp(-(w_1 d + w_0))} \quad (5)$$

ここに、パラメータ $\mathbf{w} = (w_0, w_1)$ である。この $f(d)$ により、 d のもとで発話が正しく理解される確率をモデル化する。

3.2 統合確信度関数の学習

マージンと正解ラベルを学習サンプルとして、ロジスティック回帰により $f(d; \mathbf{w})$ のパラメータ \mathbf{w} を推定することを考える。学習サンプル集合を入力 d_i と教師信号 u_i の組として以下のように与える。

$$\mathbb{T}^{(N)} = \{(d_i, u_i) \mid i = 1, \dots, N\}, \quad (6)$$

ただし、 u_i は 0 または 1 の 2 値であるとする。

いま、入力 d_i を与えたときの出力 $f(d_i)$ を、入力 d_i のもとで教師信号 u_i が 1 である確率の推定値であると考え。本手法では、BLR [Genkin 07] を用いて \mathbf{w} の推定を行う。 $w_j (j=0, 1)$ の事前分布として、平均 m_j 、分散 τ_j のガウス分布を用いる。

$$P(w_j | m_j, \tau_j) = \mathcal{N}(w_j, \tau_j) = \frac{1}{\sqrt{2\pi\tau_j}} \exp \frac{-(w_j - m_j)^2}{2\tau_j} \quad (7)$$

3.3 ELLR に基づく発話の選択

本来、確信度の学習に用いる教師データは、ユーザの発話を受けたロボットの応答から得られる。一方、確信度関数をユーザとロボットが共有していると仮定すれば、ロボットの発話に対するユーザの応答を教師データとみなすことができる。つまり、シーン O において全ての可能な言語表現に対してマージ

ンを計算すれば、特定の損失基準を最小化するマージンが選択できる。さらに、このマージンに対応する発話を行ってユーザの動作応答を観測すれば、そのマージンに対するラベル（ユーザが正しい動作を行ったか否か）を獲得できる。このように、学習者が訓練集合に加えるサンプルを選択できる枠組みを能動学習、または最適実験計画などと呼ぶ [渡辺 05]。本手法のように、サンプル候補が有限集合である場合は、pool-based な能動学習と呼ばれる。

本手法では、ELLR [Roy 01] に基づき、確信度関数の学習に有効な発話をロボットに選択させる。いま、データ集合 $\mathbb{T}^{(N)}$ を用いて訓練された確信度関数を $\hat{f}^{(N)}(d)$ と表すとする。このとき、対数損失 $L(\mathbb{T}^{(N)})$ を以下で定義する。

$$L(\mathbb{T}^{(N)}) = \sum_{i=1}^N \{ \hat{f}^{(N)}(d) \log \hat{f}^{(N)}(d) + (1 - \hat{f}^{(N)}(d)) \log(1 - \hat{f}^{(N)}(d)) \}$$

シーン O で可能な発話を $V = \{v_j | j = 1, \dots, |V|\}$, v_j から得られるマージンを e_j とする。 V は、学習済みの単語を用いて生成できる、可能な単語列集合である。ただし、単語列の長さに制約を設ける。このとき、発話 v_j を選択することに対する期待対数損失 $E(\mathbb{T}^{(N)}, e_j)$ は以下ようになる。

$$E(\mathbb{T}^{(N)}, e_j) = \hat{f}^{(N)}(e_j) L(\mathbb{T}_+^{(N+1)}) + (1 - \hat{f}^{(N)}(e_j)) L(\mathbb{T}_-^{(N+1)}),$$

$$\mathbb{T}_+^{(N+1)} \triangleq \mathbb{T}^{(N)} \cup \{e_j, 1\},$$

$$\mathbb{T}_-^{(N+1)} \triangleq \mathbb{T}^{(N)} \cup \{e_j, 0\}$$

つまり、ラベル付けされていない入力 e_j に対して、ラベルが 0 または 1 であった場合の損失を求め、 N サンプル時の確信度（確率）で重み付ける。ELLR では、上記の $E(\mathbb{T}^{(N)}, e_j)$ を最小化する v_j を選択する。

4. 実験

4.1 設定

実験に用いたロボットシステムを図 1 に示す。ロボットシステムは、7 自由度のロボットアーム（三菱重工製 PA-10）、4 自由度のロボットハンド（Barrett Technology 製 BarrettHand）、マイククロフォン、ステレオカメラ（Point Grey Research 製 Bumblebee 2）、視線表出ユニットからなる。

オブジェクトに関する画像特徴や座標は、ステレオカメラから得られた画像から抽出される。カメラのフレームレートを 30[frame/sec] とし、解像度を 320×240 とした。画像特徴量として、色 3 次元 ($L^*a^*b^*$)、形状 3 次元を用いる。あらかじめ、[Iwahashi 08, Sugiura 07] で提案した手法により、表 1 に示す単語を学習させた。

提案手法を評価するために、2 種類の実験を行った。すなわち、(1) 確信度関数の能動学習、(2) 事前分布としての学習結果の評価、である。実験 (1) の目的は、学習に必要なサンプルを調査することである。また、実験 (2) の目的は、能動学習により得られた確信度関数を、事前知識として用いることの有効性評価である。

実験 (1) では、まず、図 1 に示すように、被験者をロボットとオブジェクト（ぬいぐるみ）が置かれたテーブルをはさんで対面させた。次に、提案手法を用いてロボットに発話を行わせ、被験者にオブジェクトを操作させた。「ロボットによる発話、被験者による動作、ロボットによる正解・不正解の評価」を 1 つのエピソードとして、30 エピソードまで確信度関数の学習を行わせた。ハイパーパラメータを $m_0 = 0, m_1 = 1, \tau_0 = \tau_1 = 100$



【状況】 オブジェクト 1 が直前に操作された
R: プーサン カーミット とびこえさせて
U: (動作実行: オブジェクト 3 を持ってオブジェクト 2 の上を飛び越えさせる)

図 2: 学習フェーズにおける対話例。

と設定した。生成する単語列の長さは、 W_T, W_L が三語以内、 W_M は一語であるとした。

実験 (2) では、最適行動 \hat{a}_k のマージン d と \hat{a}_k が正解か不正解かを表すラベル u の組 $\langle d, u \rangle$ から確信度関数を学習させる。実験 (2) で用いる訓練およびテストデータは以下のように収集した。実験 (1) と同様の実験環境で、被験者にロボットにオブジェクトを操作させるための発話を行わせ、カメラ画像と音声で 60 セット収録した。得られた画像・音声セットに、ユーザが意図した行動を正解としてラベル付けした。収録データのうち半数の 30 個を訓練集合、残りの 30 個を評価集合とした。実験 (1) で得たパラメータを事前分布として訓練を行った場合と、実験 (1) で用いた無情報事前分布に近い分布 ($m_0 = 0, m_1 = 1, \tau_0 = \tau_1 = 100$) を用いて訓練を行った場合のテストセット尤度を比較する。

表 1: 実験で用いた学習済みの単語
オブジェクトを指示する単語 C_N

オブジェクトを指示する単語 C_N			
アカイ	アオイ	ミドリ	キイロイ
オレンジ	マルイ	オオキイ	グローバー
チイサイ	エルモ	カーミット	チュートトロ
プーサン	ハコ	バーバズー	バーバブライト
動作を指示する単語 C_V			
のせて	はなして	ちかづけて	とびこえさせて
あげて	さげて	まわして	

4.2 結果 (1): 確信度関数の能動学習

統合確信度関数の能動学習に関する定性的結果を図 3 に示す。図における (a) から (c) の曲線は、それぞれ訓練サンプル数 10, 20, 30 のときの BLR による回帰結果を表す。図より、サンプル数 30 において収束していることがわかる。

図 2 に、学習フェーズにおける被験者 (U) とロボット (R) の対話例を示す。このとき、可能な動作とオブジェクトの組は合計 45 種類であった。この中で、「プーサンカーミットとびこえさせて」という単語列でオブジェクト 2 とオブジェクト 3 を表現するものが、対数損失を最小化する発話であった。このときの最適マージンは 13.4 であった。

図 4 は、エピソード数（ロボットの発話数）に対する、選択されたマージンの変化を示す。図において、点線はユーザの動作が失敗したエピソードを示す。図より、ユーザの動作失敗の次のエピソードでは、より大きいマージンを持つ（曖昧性の低い）発話を選択されていることがわかる。

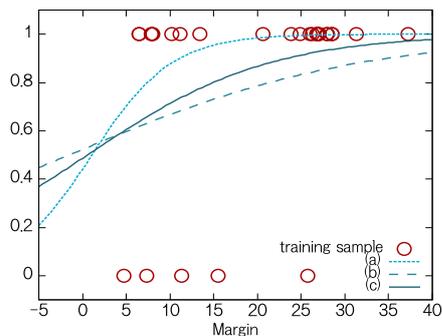


図 3: 統合確信度関数の学習結果. ○は訓練サンプルを表す.

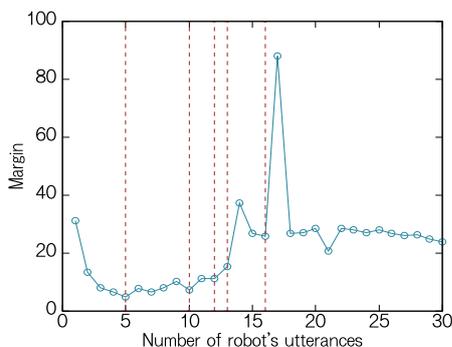


図 4: 選択された発話の-margin の変化

4.3 結果 (2): 事前分布としての有効性評価

図 5 は、各手法のテストセット尤度を比較したものである。図において、(1) は能動学習を事前知識とする提案手法、(2) は事前知識を用いない通常の事前分布を用いた手法、の結果を表す。図には、10 種類の異なる訓練集合と評価集合を用いた場合の平均を示した。図より、特に学習の初期において、提案手法が (2) を上回ることがわかる。

さらに、提案手法による尤度が、3 サンプル時においてほぼ収束しているのに対し、(2) による尤度が収束するには 10 サンプル程度必要である。訓練集合中の負事例の割合は 18% であったので、収束までの負事例数の期待値は、提案手法が 0.54 であるのに対し、(2) では 1.8 となる。このことから、能動学習で得られた学習結果を事前分布に用いることにより、学習が収束するまでにロボットが動作を失敗する回数の期待値を減少できることが示唆される。

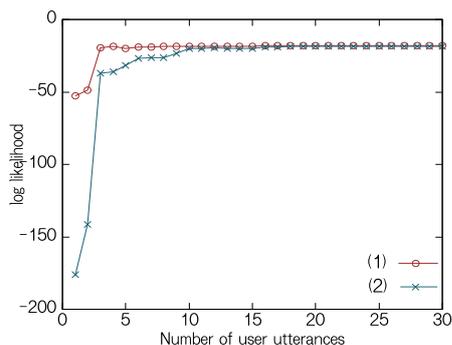


図 5: テストセット対数尤度

5. おわりに

生活支援ロボットが日常環境に導入されるためには、ユーザとの安全・安心なインタラクションを実現する必要がある。本論文では、学習の一部をユーザが代替することにより、学習フェーズにおけるロボットの失敗動作回数を削減する手法 LCore-AL を提案した。本手法では、能動学習の枠組みにより確信度の学習に有効な発話を選択される。音声合成されたこれらの発話と、発話に対するユーザの動作応答から、「ユーザの発話に対する確信度関数」が学習される。このようにして学習された確信度関数は、[杉浦 09] で提案した手法と同時に用いることにより、実行フェーズでの動作失敗率についても減少させることが可能である。

謝辞

本研究の一部は、日本学術振興会科学研究費補助金 (基盤研究 (C) 課題番号 20500186) および立石科学技術振興財団による研究助成を受け実施したものである。

参考文献

[Genkin 07] Genkin, A., Lewis, D., and Madigan, D.: Large-scale bayesian logistic regression for text categorization, *Technometrics*, Vol. 49, No. 3, pp. 291–304 (2007)

[Iwahashi 07] Iwahashi, N.: Robots That Learn Language: Developmental Approach to Human-Machine Conversations, in Sanker, N., et al. eds., *Human-Robot Interaction*, pp. 95–118, I-Tech Education and Publishing (2007)

[Iwahashi 08] Iwahashi, N.: Interactive Learning of Spoken Words and Their Meanings through an Audio-Visual Interface, *IEICE Transactions on information and systems*, Vol. 91, No. 2, p. 312 (2008)

[Katagiri 98] Katagiri, S., Juang, B., and Lee, C.: Pattern recognition using a family of design algorithms based upon the generalized probabilistic descent method, *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2345–2373 (1998)

[Kawahara 98] Kawahara, T., Lee, C., and Juang, B.: Flexible speech understanding based on combined key-phrase detection and verification, *IEEE Transactions on Speech and Audio Processing*, Vol. 6, No. 6, pp. 558–568 (1998)

[Roy 01] Roy, N. and McCallum, A.: Toward optimal active learning through sampling estimation of error reduction, in *Proceedings of 18th International Conference on Machine Learning*, pp. 441–448 (2001)

[Roy 02] Roy, D.: Learning visually grounded words and syntax for a scene description task, *Computer Speech and Language*, Vol. 16, No. 3, pp. 353–385 (2002)

[Sugiura 07] Sugiura, K. and Iwahashi, N.: Learning object-manipulation verbs for human-robot communication, in *Proceedings of the 2007 workshop on Multimodal interfaces in semantic interaction*, pp. 32–38 (2007)

[杉浦 09] 杉浦孔明, 岩橋直人: 物体操作タスクにおける発話理解確信度の推定に基づく発話と動作の生成, 第 36 回知能システムシンポジウム, pp. 29–34 (2009)

[山肩 04] 山肩洋子, 河原達也, 奥乃博, 美濃導彦: 音声対話システムにおける物体指示のための信念ネットワークを用いた曖昧性の解消, *人工知能学会論文誌*, Vol. 19, No. 1, pp. 47–56 (2004)

[渡辺 05] 渡辺澄夫, 萩原克幸, 赤穂昭太郎, 本村陽一, 福水健次, 岡田真人, 青柳美輝: 学習システムの理論と実現, 森北出版株式会社 (2005)