

# invisible な迷路タスクを用いた人の学習過程の可視化による 継続的学習の支援

Visualizing mental learning processes with invisible mazes for assisting continuous learning

山口 智浩\*<sup>1</sup> 竹森 孝樹\*<sup>1</sup>  
Tomohiro Yamaguchi and Kouki Takemori

高玉 圭樹\*<sup>2</sup>  
Keiki Takadama

\*<sup>1</sup> 奈良工業高等専門学校  
Nara National College of Technology

\*<sup>2</sup> 電気通信大学  
The University of Electro-Communications

This paper reports our learning support system for a human learner to visualize his/her mental learning processes with invisible mazes for continuous learning. The objective of this research is to bring the learning ability of the learning agent close to that of a human. To fill in the missing piece of reinforcement learning whose learning process is mainly behavior change, we add two mental learning processes, awareness as pre-learning process and reflection as post-learning process. To observe mental learning processes of a human, we propose a new method for visualizing mental learning processes with invisible mazes consisting of invisible walls which are perceived as a sign that is the number of walls in the neighborhood.

## 1. はじめに

本論文は、強化学習の枠組みを拡張した、人と学習エージェント共通の継続的な学習過程モデルの設計[Yamaguchi 2014]について述べる。強化学習を拡張する学習前後の過程として、気づき、振り返りの2つの心的過程を追加し、継続的学習タスクを通して両者を可視化する。まず迷路内の invisible な壁への遭遇・回避行動の有無によって気づき過程を可視化する。次に invisible な壁の気配からの気づきを記述する meta-action によって振り返り過程を可視化する。実験結果から継続的学習タスクの解発見数が振り返りコストと強い負の相関を持ち、解発見数上位者ほど効率よく振り返りを行っていることが示唆された。

## 2. 継続的な学習過程モデル

本節では継続的学習[Sessa 2006][Smita 2012][Yamaguchi 2013]、気づき、振り返りの理論的背景について述べる。

### 2.1 継続的学習と強化学習との比較

Fig.1に人間の継続的学習と学習 agent の強化学習との学習過程の比較を示す。Fig.1(a)は Buckler によって提案された継続的改善を達成するための学習過程モデル[Buckler 1996]である。その主な特徴は各過程が円状で終りがなくである。しかしながら各過程は心的過程であり、アルゴリズム化されていない。

一方、Fig.1(b)に学習 agent の強化学習過程を示す。まず学習過程全体は Fig.1(a)と異なり、直線状で始まりと終わりがある。

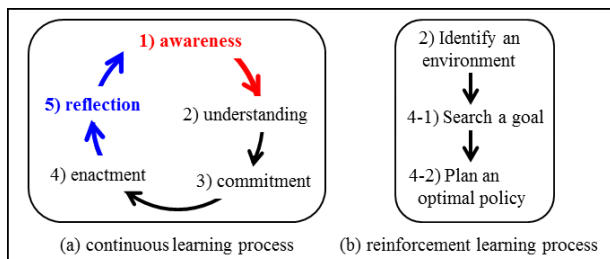


Fig. 1: A comparison of learning processes

その理由は、既存の強化学習の目的が人間の設計者から与えられた学習目標に対する最適解の発見だからである。次に各過程を両者で比較すると、Fig.1(a)の 2) understanding, 4) enactment がそれぞれ Fig.1(b)の 2)と 4-1), 4-2)の過程に相当する。Fig.1(a)の過程 1), 3), 5)は Fig.1(b)にない。これらは学習モデルやアルゴリズムを変更するメタ学習的役割を持ち、継続的学習の本質的過程である。うち、3) commitment 過程は標準的強化学習アルゴリズムには含まれず、設計者が学習目標を設定する。Fig.1(a)の 1) awareness, 5) reflection 過程については次節で説明する。

### 2.2 学習過程における気づきと振り返り

本節では、継続的学習過程における本質的な過程である気づきと振り返り[Kravcik 2013]について述べる。気づき(Awareness)と振り返り(Reflection)の共通点は、将来の改善のための情報となる学習者の経験に注目することである。両者の違いは、気づきが知覚に関連し、振り返りが行動(行為または知覚と行為)に関係することである。

#### (1) 学習前過程としての気づき(Awareness)

学習に関する気づき[Marton 1997][Reinhardt 2011]には様々な意味がある。本研究では気づきの定義を“近い将来の状況の予測に関する知覚”[Kurapati 2012]とする。学習前過程において、気づきは行動変更(behavior change)のきっかけという重要な役割がある。そこで我々は、気づきの具体的な意味を“将来の成功と失敗との間での区別困難な不完全知覚を区別する必要性”とする。理論的モデルとして、上述の区別困難な知覚は不完全知覚状態[Kaelbling 1998]によって生じると仮定する。

#### (2) 学習後過程としての振り返り(Reflection)

学習に関する振り返り[Boud 1985]には様々な意味がある。本研究では振り返りの定義を“将来の行動を導くための経験の意識的な再評価”[Krogstie 2012]とする。学習後過程において、振り返りは行動変更や学習結果からの何らかの意味の創造という重要な役割がある。そこで我々は、振り返りの具体的な意味を、“学習者を将来の失敗でなく成功へ導く規則・習慣に向けての説明・解釈の生成”とする。

## 3. 迷路モデルによる継続的な学習過程の設計

本節では学習環境と継続的学習タスクのための設計概念

[Yamaguchi 2014][山口 2014]について述べる。本研究では Fig.1 の継続的学習過程のモデル化として、学習環境を一つの目標地点と迷路を構成する壁がそれぞれ invisible である迷路モデルによって定式化[Yamaguchi 2013]する。継続的学習タスクは、迷路内をくまなく歩く迷路巡回タスクとし、学習者にとって当初は気づかない様々な学習目標を発見するよう設計する。

### 3.1 継続的学習過程の流れ

Fig.2 に本論文で提案する継続的学習過程の流れを示す。この過程は三重のサイクル[Yamaguchi 2014]で構成される。内側のサイクルを trial と呼ぶ。trial は、開始状態から目標状態との遭遇(迷路巡回タスク条件を満たした場合は成功、他は失敗)または壁と呼ばれる障害との遭遇(失敗)までの遷移系列で定義される。このサイクルで、点線で囲まれた mental processes は、外部から観測困難な心的過程で、Fig.1(a)の 1)から 5)の各過程に相当する。学習者はタスクが成功または失敗するまで、action と気づき(awareness)を含む心的過程を繰り返す。

中間のサイクルを achievement と呼ぶ。achievement は、迷路巡回(maze sweeping)タスクの学習で定義される。このサイクルで、学習者は壁や目標との遭遇による trial 終了時に reflection でその trial を振り返る。その後、進行中の trial が未成功の場合には、開始状態に遷移してタスクを再開(restart)し、trial が成功の場合には achievement を終了する。外側のサイクルは継続的学習サイクルである。進行中の achievement を成功または give up すると学習者は次の achievement に挑戦できる。

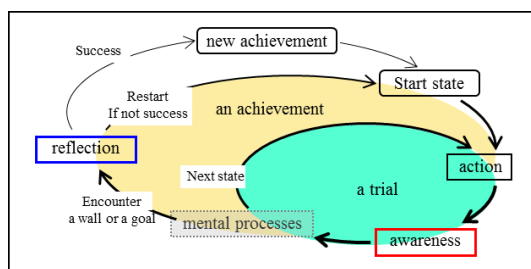


Fig. 2: The flow of the continuous learning process

### 3.2 invisible 迷路モデルによる学習環境の設計

人の学習者のための学習環境の設計として、開始状態から目標状態に至る格子迷路モデル[Yamaguchi 2013]を用いる。迷路モデルを用いる理由は、試行錯誤の過程を通した解の道筋を見出す例として、両者を対応させやすい例だからである。迷路モデルは状態集合、壁の気配値(sign), 遷移と壁、action 集合、meta-action の 5 つの要素で定義される。

Fig.3 に 2 次元格子迷路の構成を示す。周囲の壁が長方形である迷路を単純迷路と呼ぶ。Fig.3 (a) に開始状態 S, 目標状態 G を持つ 3×2 の単純迷路の表示イメージを示す。迷路モデルの各正方形を状態と呼ぶ。迷路モデルの状態間の遷移は、ある状態が上下左右 4 近傍の状態と連結(壁なし)か非連結(壁あり)かで定義され、Fig.3 (b) の有向グラフで表される。action 集合は、(学習者が実行可能な)上下左右 4 方向いづれかの移動を表す {up, right, left, down} の 4 種類の action で定義される。

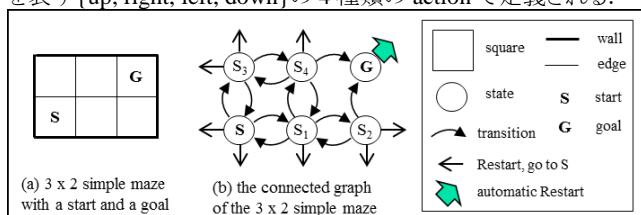


Fig. 3: The structure of a 3 x 2 simple maze

次に 3.1 節で述べた achievement で行う迷路巡回(maze sweeping)タスクについて述べる。迷路巡回タスクは、2 次元迷路内に固定された開始状態 S と目標状態 G において、迷路内の全状態をちょうど1回ずつ訪問する S から G への経路(解と呼ぶ)を求めるタスクとして定義される。

移動先の状態に遷移するとその状態を訪問済とする。但し、巡回タスクでは迷路訪問済状態への後戻り(backtrack)はできない。壁に衝突する action は、開始状態 S へ遷移して trial を restart する。目標状態 G への遷移時には、迷路巡回タスク達成の判定を行い、未達成時には自動的に restart する。

### 3.3 気配値(sign)による気づき過程の可視化

本節では、迷路内の invisible な壁の存在を示唆する気配値を学習者の知覚として提示し、壁に対する衝突・回避行動の有無によって気づき過程を可視化する方法について述べる。

ある状態の気配値(sign) [Yamaguchi 2014]は、その状態での 4 近傍における壁の数 {0,1,2,3,4} で定義される。学習者は、invisible な壁を直接知覚できない代わりに、各状態の気配値を invisible な壁の存在として知覚でき、知覚済の気配値は、trial 終了まで迷路の訪問済状態に表示される。Fig.4 に invisible な壁で構成された単純迷路の例を示す。Fig.4 (a) に 4×4 の invisible な単純迷路を示す。点線は、trial 中、表示されない invisible な壁の位置を表す。Fig.4 (b) に遷移系列の例を示す。Fig.4 (c) に迷路内での訪問済の状態での気配値の分布を示す。

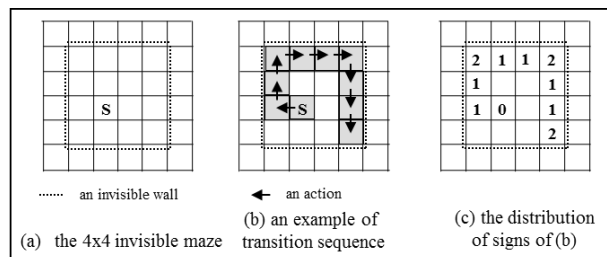


Fig. 4: An example of a simple maze with invisible walls

### 3.4 気配値に基づく不完全知覚状態とその解決

本節では、まず invisible な壁の気配値が 0 以外の場合に不完全知覚状態であることを述べ、次に学習者が自身の視野拡大に気づくと不完全知覚状態が解決でき、未知壁の予見が可能となることを説明する。

気配値 0 は周囲に壁がない(完全知覚)状態で、気配値 1,2,3 が 4 近傍での壁の位置が特定できない不完全知覚状態である。単一の不完全知覚状態を解決するには、学習者は移動して視野を拡大し、近傍状態の気配値を知る必要がある。invisible な壁を衝突前に予見するための重要な気づき(学習目標)は、Fig.4 に示す周囲の壁が長方形である単純迷路の場合、まず近傍状態間での気配値系列 01 および invisible な壁との位置関係に気づくことである。次に重要な点は、“適切な特徴に注目した視野拡大によって invisible な壁が見えるようになる”ことである。Fig.4 (c)の単純迷路の例では、2種の気配値 {01, 12} と壁(!, #)によるパターン {01!, 12#} に着目すると、未知壁!および未知な角(corner)#を予見できる。周囲の壁形状が長方形以外の凹凸、壁の分岐や孤立、島などを持つ非単純迷路の場合では、壁配置の状況に応じた気配値の分布と壁位置との規則性に気づくことができれば、未知壁の予見が可能となる。

以下に未知壁の予見のための学習目標系列を示す。

- subgoal1: invisible な壁を見つける(位置を特定する)
- subgoal2: 壁周辺の気配値の分布に注目する
- subgoal3: 未知壁を予見する気配値の規則性を見出す。

### 3.5 meta-action による振り返り過程の可視化

本節では, invisible な壁の気配からの気づきを記述する meta-action によって学習者の心的学習過程を表出し, 振り返り過程を可視化する方法について述べる. meta-action は, マーカーを迷路地図上の状態に置く操作として定義され, 学習者の心的学習過程を表現し要約するのが目的である.

振り返りマーカーは, 緑, 赤, 青 の 3 種類あり, 各 Trial の最後に行う振り返り過程を表出するために用いる. これらのマーカーは色とマーカー数で区別される. 緑マーカーは 1 つ, 赤マーカー, 青マーカーは 4 つずつとする.

trial 終了時の振り返り過程で, 振り返りマーカーを迷路地図上の任意の状態に置くことができる. 典型的な使い方は, 緑マーカーは目標状態を見つけた時の印を, 黄色マーカーと赤マーカーは発見した壁の位置および迷路の四隅の位置を表す場合である. 地図上に印をつけた振り返りマーカーは, ある achievement での一連の trial 中, 迷路地図として表示される.

## 4. 実験

本章では, invisible 迷路系列からなる学習環境による継続的学習タスクの被験者実験[山口 2014]について述べる. 気づきに関する先行研究[藤本 2004]では, 答えの直接教示より間接教示の方が学習の定着率が良いことがわかっている. この知見を参考に, 教示での学習タスクに関する直接的な説明を少なくして, 実験条件を設定した. 主な実験目的が, 被験者の継続的学習の長さや振り返り過程の可視化であるため, 被験者には実験を行うための操作方法を教示し, ステージ総数や迷路形状, 用意した色マーカーの役割, 使い方等は教示していない.

### 4.1 invisible 迷路による継続的学習タスク設定

#### (1) 学習環境

被験者の継続的学習の長さを見積もるため, 全 54 ステージからなる迷路系列を用意した. 各ステージは単一の invisible 迷路でそれぞれ 2 つ以上の解を持つ. 解の総数は,  $54 \times 2 = 108$  個である. ステージ系列は, 難易度順となるよう, 前半の 8 ステージは単純迷路を状態数の昇順に並べ, 9 ステージ目以降は非単純迷路を状態数の昇順に配置した. 用いた非単純迷路は, 周囲の壁形状が長方形以外の凹凸を持ち, 同一形状の迷路では後半のステージの迷路内に独立した壁を配置している.

#### (2) meta-action による振り返り過程の可視化

trial 終了時の振り返り過程において, 3 種の色マーカー(緑 1, 赤 4, 青 4)を用意した. 色マーカーの役割については説明せず, 事後アンケートで各色マーカーの役割を決めたか, 決めていた場合にはその使い方を調査した.

#### (3) 継続的学習タスク

実験前の被験者への教示の概要を下記に列挙する.

- 「スタートマスからゴールマスに至る経路の中で最大長の経路」を解と呼ぶ.
  - 解を 2 つ見つけることがタスクの目的である.
  - 実験を続けたくないと思えば終了してよい.
  - マスに表示される数字は隣接する壁の枚数である.
  - 迷路中での操作方法
  - reflection 過程での説明, マップとマーカーの操作方法
- “ここでは迷路での挑戦を振り返り, 気づいたことを記録することができます. (中略) マップにはマークを置くことができます. 次の挑戦に備えて, 挑戦時に気になったセルや注意したい場

所にマークを移動させて下さい. マークには, Red(4 つ), Blue(4 つ), Green(1 つ) の 3 種類があります”(以下操作方法の説明)

#### (4) 測定項目

主な測定項目は, 被験者が挑戦したステージ数, ステージの挑戦回数, 壁との衝突回数, 解の発見数, 実験時間, achievement 時間, reflection 時間, reflection 回数である. reflection 過程では, 被験者が作成した迷路地図を画像形式で保存した. 実験時間の構成を式(1)に示す.

$$\text{実験時間} = \text{achievement 時間} + \text{reflection 時間} \cdot (1)$$

#### (5) 継続的学習タスクでの仮説

- reflection 時間が長い程, 継続的学習が長くなる.

## 4.2 実験結果

18 才から 22 才前後の高専男子学生 8 名を被験者として 4.1 節で説明した同一条件を用いて学習実験を行った. 実験の謝礼は, 継続的学習タスクへ影響を与えないように, 実験時間に関係なく, 安価なお菓子(うまい棒 1 本)とした. 以下で実験結果の概要を述べる.

- 継続的学習の長さ

- (1) 実験時間: 最短の被験者で 48 分間実験を行った. 残りの 7/8 名の被験者が 60 分以上(60~105 分間)実験を続けた.
- (2) 挑戦ステージ率と解発見率: 6/8 名の被験者が半数以上(30~50 ステージ)のステージに挑戦し, かつ半数以上(60~100 個)の解を発見した.

- 継続的学習における実験時間の分析

- (3) reflection 時間の長さは, 解発見数と弱い負の相関がある. (相関係数 -0.28)
- (4) achievement 時間は, 解発見数と強い正の相関がある. (相関係数 0.94)
- (5) 壁との衝突回数は, 解発見数と強い正の相関がある. (相関係数 0.88)

実験結果(1)(2)より, 大半の被験者が実験タスクに 60 分以上取り組み, 半数以上の解を発見し, 結果的に invisible 迷路が学習タスクの継続性に貢献したと考えられる.

これに対し, 実験結果(3)は, 著者らにとって予想外であった. 事後アンケートの回答および実験状況の分析から, 解発見数上位 2 名 (No.8,6) の被験者は, 主に序盤のステージの reflection 時間が長く, 中盤以降は一部のステージでのみ reflection 時間を使っていることがわかった. 一方, 解発見数最小 (No.7: 16 ステージで解 20 個を発見) の被験者は, reflection 時間が 73 分間で, 8 名中 2 番目に長い. これらから, 解発見数は reflection 時間(量)でなく reflection の質に依存することが予想される. reflection の質を評価する指標として reflection コストを式(2)で定義し, 分析を行った.

$$\text{reflection コスト} = \text{reflection 時間} / \text{解発見数} \cdot (2)$$

Table 1 に reflection コストを含めた実験結果を示す. 相関分析の結果から, reflection コストは, 解の発見数と強い負の相関(相関係数 -0.78)が, 壁との衝突回数と中程度の負の相関があった(相関係数 -0.68). つまり, 今回の継続的学習タスクにおいて, 解発見数上位の被験者は, 解発見数下位の被験者と比べて要領よく reflection を行い (reflection コスト小), かつ壁との衝突を許容する(壁との衝突回数大)ことで, 多くの解を発見していたことが示唆される.

測定項目 被験者番号	挑戦 ステージ の数	解の 発見数	挑戦 回数	衝突 回数	Achievement 時間 [sec]	Reflection 時間 [sec]	Reflection コスト
8	50	100	596	403	2647	3053	30.53
6	50	98	716	470	2238	1962	20.02
4	43	86	540	352	2521	3419	39.76
1	35	70	228	120	1755	4545	64.93
3	31	61	319	201	1634	3286	53.87
5	30	60	250	116	948	2652	44.20
2	17	34	211	153	735	2145	63.09
7	16	20	73	33	557	4423	221.14
Average	34	66	367	231	1629	3186	67.19

Table 1 Experimental results

### 4.3 議論

#### (1) reflection コストの意味と上達支援

reflection コストの意味と使い方について議論する. reflection コストを振り返りスキルの上達度とみなすと, 上達するにつれてコストが下がり, 楽に解けるようになっていく. それゆえ, 多くのステージで解を発見し, 継続的学習が長くなったことが予想される. 今後の課題として, 学習タスク中の reflection コストの逐次的算出を予定している. これにより, reflection コストの大きい被験者を学習早期に発見し, 振り返りスキルの初級者として選択的に支援できる. 一方, reflection コストの小さい被験者に対しては, 上級者と位置づけ, タスクの難易度調整および後述する新たな学習目標の設定を検討する.

#### (2) 振り返りスキルの個人差に基づく対話的上達支援へ向け

reflection コストの個人差の要因と有効な使い方について議論する. reflection コストの大きい下位被験者は, マーカーの配置数が多く, かつ配置規則の一貫性が弱い傾向があるのに対し, reflection コストの小さい上位被験者は, reflection map でマーカーを置かずに頭の中だけで迷路タスクを解いているステージが中盤以降多い. つまり, reflection コストだけでなく, マーカーの使用数や置き方を行動観察法[松波 2013]で分析すると, 振り返りスキルの上達度を精度良く見積もれると期待される. 一方, 今回実験した学習環境および学習タスク系列だけでは, 振り返りスキルの個人差を吸収しきれなかったことが予想される. そこで, 振り返りスキルの個人差に基づく適応的上達支援および対話的上達支援として, 以下の項目を検討予定である.

- ・振り返りスキルの個人差に基づく上級者と初級者との分類
- ・振り返りスキルに応じた気配値提示による適応的上達支援
- ・上級者初級者ペアによる振り返りスキルの対話的上達支援
- ・行動観察法による reflection 過程・スキルの分析と細分化

### 5. 結論

本論文では, 強化学習の枠組みに気づき, 振り返りの 2 つの心的過程を追加した三重のサイクルで構成された継続的学習過程モデルを提案し, invisible 迷路タスクを用いて気づきの可視化, 振り返りの可視化を実現した. invisible 迷路系列を用いた継続的学習タスクの被験者実験結果の分析から, 継続的学習タスクの長さを見積もる解発見数が振り返りコストと強い負の相関を持ち, 解発見数上位者ほど要領よく振り返りを行っていることが示唆された.

今後の課題は, 継続的学習条件を明らかにするための比較条件の検討, reflection コストに基づく学習タスク中の振り返りスキルの上達度のオンライン見積もりの実装と検証, 振り返りスキルの個人差に基づく適応的・対話的上達支援の検討である.

### 謝辞

本研究は JSPS 科研費 23500197 の助成を受けた.

### 参考文献

- [Boud 1985] Boud, D., Keogh, R., and Walker, D. "Reflection: Turning Experience into Learning", Routledge (1985)
- [Buckler 1996] Buckler, B. "A learning process model to achieve continuous improvement", The Learning Organization, Vol. 3, No. 3, pp.31-39, MCB University Press (1996)
- [藤本 2004] 藤本 武司, 砂山 渡, 山口 智浩, 谷内田 正彦: "視線行動の可視化による着眼スキル伝達支援", 人工知能学会論文誌, Vol. 19, No. 3, pp.174-183 (2004)
- [Kaelbling 1998] Kaelbling, L.P., Littman, M.L. and Cassandra, A.R. "Planning and acting in partially observable stochastic domains", Artificial Intelligence, Vol.101, pp.99-134, Elsevier (1998)
- [Kravcik 2013] Kravcik, M., et al.(Eds.), Proceedings of the 3rd Workshop on Awareness and Reflection in Technology Enhanced Learning (ARTEL2013), <http://ceur-ws.org/Vol-1103/> (2013)
- [Krogstie 2012] Krogstie, B., Prilla, M. "Tool support for reflection in the workplace in the context of reflective learning cycles", Proceedings of the 2nd Workshop on Awareness and Reflection in Technology-Enhanced Learning (ARTEL 2012), pp.57-72, <http://ceur-ws.org/Vol-931/> (2012)
- [Kurapati 2012] Kurapati, S., et al. "A Theoretical Framework for Shared Situational Awareness in Sociotechnical Systems", Proceedings of the 2nd Workshop on Awareness and Reflection in Technology-Enhanced Learning (ARTEL 2012), pp.47-53, <http://ceur-ws.org/Vol-931/> (2012)
- [Marton 1997] Marton, F. and Booth, S. "Learning and Awareness", Routledge (1997)
- [松波 2013] 松波 晴人 "「行動観察」の基本", ダイアモンド社 (2013)
- [Reinhardt 2011] Reinhardt, W. and Christian, M. "Awareness in Learning Networks", Proceedings of the The PLE Conference 2011 (ARTEL2011), pp.12-20 (2011)
- [Sessa 2006] Sessa, V. I., and London, M. "Continuous Learning in Organizations: Individual, Group, and Organizational Perspectives", Psychology Press (2006)
- [Smita 2012] Smita, J., and Trey, M. "Facilitating continuous learning: re-view of research on individual learning capabilities and organizational learning environments", the annual meeting of the AECT International Convention, Louisville (2012)
- [Yamaguchi 2013] Yamaguchi, T., Takemori, K. and Takadama, K. "Modeling a human's learning processes toward continuous learning support system", in Interdisciplinary Mechatronics, M. K. Habib, J. Paulo Davim, (eds.), pp.69-94, Wiley-ISTE (2013)
- [Yamaguchi 2014] Yamaguchi, T., Takemori, K., Takadama, K. "Visualizing mental learning processes with invisible mazes for continuous learning", Proceedings of 16th International Conference, HCI International 2014, June, (2014) to appear
- [山口 2014] 山口 智浩, 竹森 孝樹, 高玉 圭樹: "Invisible な迷路タスクを用いた人の学習過程の可視化", 計測自動制御学会, 第 41 回知能システムシンポジウム (2014) to appear