

情報検索技術を補うための知的ネットワークエージェントの設計と試作

著者	石川 尚志, 石本 一生, 植田 和憲, ISHIKAWA Naoyuki, ISHIMOTO Issei, UEDA Kazunori
雑誌名	情報処理学会研究報告. DSM, [分散システム/インターネット運用技術]
巻	2007
号	38
ページ	43-48
発行年	2007-05-10
その他のタイトル	The network agent's to supplement information retrieval technology design and trial
URL	http://hdl.handle.net/10173/292

情報検索技術を補うための 知的ネットワークエージェントの設計と試作

石川 尚志[†] 石本 一生[†] 植田 和憲[†]

[†] 高知工科大学工学部 〒782-8502 高知県香美市土佐山田町宮ノ口185

あらまし 近年のインターネットの普及により、様々な情報やサービスをネットワークを介して利用できるようになった。しかし、非常に多岐にわたる情報やサービスの中から必要なものを見つけ出すためには時間と労力を消費してしまう。そこで本報告では、この問題を解決するために知的ネットワークエージェントを用いた効率的なWeb情報検索システムを提案する。このシステムでは、ユーザの代わりとなる知的ネットワークエージェントによる各種のフィルタリングにより、ユーザはあまり手間をかけることなく必要な情報を手に入れることができる。

キーワード 知的ネットワークエージェント, Web 検索, フィルタリング

The network agent's to supplement information retrieval technology design and trial

Naoyuki ISHIKAWA[†], Issei ISHIMOTO[†], and Kazunori UEDA[†]

[†] Faculty of Engineering, Kochi University of Technology Miyanokuchi 185, Tosayamada-tyou, Kochi, 782-8502 Japan

Abstract In recent years, various information and services have been available on the Internet. However, it usually takes very long time to find information or service we need. To solve this problem, we propose a web information search system based on concept of the intelligent network agent. In this system, the intelligent network agents execute the several filtering instead of users and users can get information they need.

Key words Intelligent Network Agent, Web Search, filtering

1. ま え が き

近年、インターネットが普及してきたことによって、非常に多くの情報やサービスを利用することができるようになってきている。しかし、情報量の増加は、情報の分散や巨大化の原因ともなった。そのため情報の検索は、存在するものを探し当てるといった側面から多数ある候補の中から最も適したものを選択するという側面へと変化してきている。言い換えれば、情報を検索する際に MECE (Mutually Exclusive, Collectively Exhaustive: 完全網羅) から非 MECE へとパラダイムシフトが起こっているということである。このことは、現在広く用いられている検索エンジンである Google などの検索結果の表示方法を見てもうかがい知ることができる。このように、Web上の情報の検索においては、検索条件に合致する全ての Web ページを網羅するのではなく、ユーザが必要としていると推測される Web ページを得ることができるようなフィルタリング手法の重要性が高まっている。また、普遍的な基準による推薦ページのフィルタリングのみならず他のユーザの嗜好などを反

映してフィルタリングを行う協調フィルタリングという手法にも注目することができる。

フィルタリングは大きく、レーティング方式、ブラックリスト方式、ホワイトリスト方式、キーワード/フレーズ方式の4つに分けることができる。レーティング方式とは、一定基準で格付け(レーティング)を行い、情報受信者がレーティング結果を利用してフィルタリングを行う手法である。ブラックリスト方式は必要としない、あるいは閲覧したくない情報についてのキーワードなどをブラックリストとしてデータベース化し、検索結果からそれらを除くというものである。逆に、ホワイトリスト方式は必要とする、あるいは閲覧したい情報についてのキーワードなどをホワイトリストとしてデータベース化するものである。協調フィルタリングはこの中のレーティング方式に属するフィルタリング手法である。しかし、協調フィルタリングは人間の情報収集行動から興味・関心・意図といった意識をデータ化したものやフィルタリングで得られた結果、そして他のユーザの嗜好情報などを用いるため、フィルタリングのために必要な情報が膨大となってしまうため、フィルタリングを行

うための各種オーバーヘッドが大きくなることが考えられる。この問題を解決するため、近年注目を集めているエージェント技術 [1], [2] を用いる。協調フィルタリングを行うためにネットワーク上に分散して存在するユーザの行動を観察し、個別にモデリング (modeling) を行う必要があるが、これを集中型サーバ・クライアントシステムで行うためにはある程度の規模のシステムが必要である。そこで、エージェントを用いて分散した情報の交換を行うことによって、個々が持つ情報量を抑えることができる。

エージェントとは、一般的に『代理人』や『動作主体』という意味を持つ。例えば旅行案内業者では、動作の主体が人間であり、客の希望に応じて航空機やホテルの予約などを行う。すなわち、ユーザが時間不足、知識不足、経験不足のため自分で行うことのできないタスクあるいは行いたくないタスクを、ユーザの代わりに行う動作主体であるといえる。また、エージェントは着目する視点によって様々な名前がつけられている。例のエージェントでは動作主体に注目した場合、ヒューマンエージェントと呼ぶことができ、仕事の内容に注目した場合、アシスタントエージェントと呼ぶことができる。本報告で提案するシステムで用いるエージェントは知的な処理を自律的に行う知的エージェントと、ネットワーク移動機能を持つモバイルエージェント [3] を併せたエージェントであり、知的ネットワークエージェントと呼ばれる [2], [4]。

本報告では、知的ネットワークエージェントの行うユーザの嗜好などを基にした協調フィルタリングによりユーザが必要とする情報を推薦するためのフィルタリングシステムを提案する。提案においては、協調フィルタリングを行うために必要な多くの情報を分散して蓄積し、またそれらを適宜交換することができるような知的ネットワークエージェント、それらが活動できる環境 (プラットフォーム) などの設計を行う。

次章で、本研究で基にする協調フィルタリングを始めとする推薦精度を向上させるための関連技術について述べる。3 章以降では、本報告で提案するシステムの詳細な説明を行う。

2. 既存の情報検索手法

情報量が膨大となった WWW システムから推薦する Web コンテンツを選択する手法に関する研究が数多く行われている。その中の一つに情報検索の精度の向上を目指した適合性フィードバック (relevance feedback) を用いた手法がある [5]。これは、Web コンテンツの被推薦者によって行われる有効、あるいは有効でないという評価に基づいてキーワードに重みを持たせ、その情報をシステムにフィードバックすることで検索精度の向上を図る手法である。しかし、システムの利用者が検索結果に対して評価を行う必要があることから、ユーザインタフェースに制約の多い携帯端末などの評価のための操作が行いにくい環境を想定する場合は必ずしも有効ではない。

また、協調フィルタリング (Collaborative Filtering, CF) [6] に関する研究も多数行われている。これは、各ユーザの嗜好情報を蓄積し、他のユーザの類似する嗜好情報を用いて推論を行う方法である。協調フィルタリングに用いられる嗜好情報の取

集方法には、ユーザの評価付けによる明示的なものと、システムの操作履歴 (例えばブラウザの閲覧履歴) などを利用する暗黙的なものがある。ユーザの行動を収集した膨大な暗黙的データから、ユーザの関心を特定するためにデータマイニングを応用することが期待されている。

一方で、WWW 上の情報リソースにメタデータを付与し、それらの意味的關係をオントロジとして明示化することによって意味的な識別を可能とすることを目的とした研究が活発に行われている。これらの研究の一つである Folksonomy [7] は、folks (人々の) と taxonomy (分類学) を組み合わせた造語で「みんなの分類」という意味を持ち、Web ページのようなテキスト形式のファイルに対して有効である。Folksonomy では、コンテンツを見たユーザによって、コンテンツに当てはまるような属性 (タグ/tags) と呼ばれるキーワードを付加する。タグはどのような内容でも良く、複数付加しても構わない。このようにして多数のユーザによって付加されたタグの情報を共有して集約することにより、分類情報を構築することを可能としている。しかし、多数のユーザが個々にタグを付加していくため、タグの表記のゆれ (意味は同じであっても表記方法が違う単語など) や内容の正当性評価など改善すべき点が多くある。

3. システム概要

私たちが Web ページを検索するためには、検索エンジンへアクセスし、キーワードを入力することによって情報を検索することができる。しかし、既存の検索エンジンのインターフェースは、あくまでもパーソナルコンピュータのユーザインタフェース環境を想定して設計されているため、網羅性や一貫性には優れるが、必ずしも必要な情報へ容易にたどり着けるとは限らない [8]。

そこで本報告では、ユーザの嗜好を協調フィルタリングの暗黙的なものとして利用することによって、ユーザが必要とする情報を推薦するためのフィルタリングシステムを提案する。

嗜好を反映した協調フィルタリングを知的ネットワークエージェントに行わせることによって情報量の巨大化の問題を解消することができると思われる。

今回提案するエージェントシステムの概要図を図 1 に示す。本システムのエージェントはエージェントプラットフォーム上で動作する。エージェントプラットフォームはエージェントが動作するための場所である。エージェントはプラットフォーム上に生成され、プラットフォーム間を移動しつつ、ユーザから与えられた要求を解決する。各プラットフォームは任意のコンポーネント (エージェントの処理群) を持っている。図 1 では、Web 検索 API コンポーネントを持つサーチプラットフォーム、エージェントの休憩所またはエージェント間の会話を行うコンポーネントを持つウェイトプラットフォーム、エージェントを生成・受け取りのみを行う 3 種類のプラットフォームを示している。多人数のユーザが様々なプラットフォームを起動することによって、プラットフォーム同士が繋がり、P2P 的なネットワーク上にエージェントソサエティを形成することが可能である。また、ウェイトプラットフォームやサーチプラットフォー

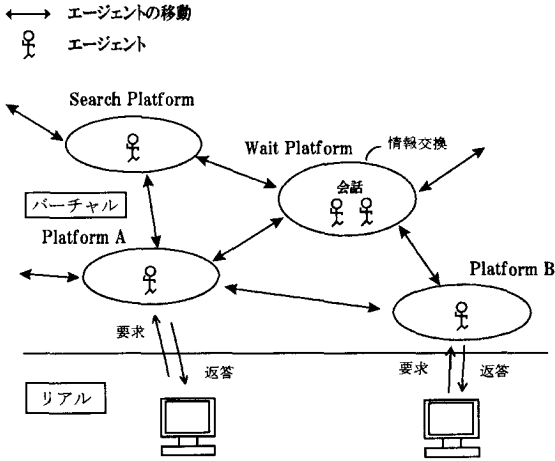


図1 システム概略図

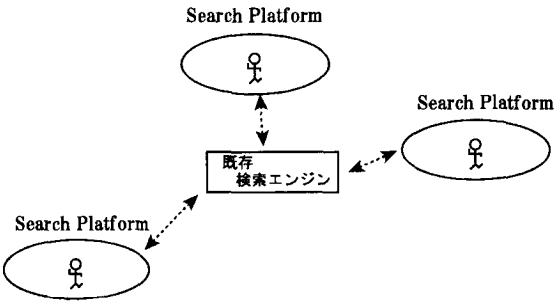


図2 システム概略図

の様なエージェントが計算処理を行うようなプラットフォームにはエージェントの行動記録が残され、これを用いて他ユーザの検索傾向の統計に利用することが可能である。

本システムで、サーチプラットフォームを用いるが、これは検索エンジンではない。本システムでは、図2に示すとおり、既存の検索エンジンを使って検索を行う。この検索結果からユーザの嗜好を反映した協調フィルタリングを行うことで推薦Webページをユーザに渡すことができる。

4. 提案システムの試作と結果

前章で述べたシステムを構築するための、試作としてプラットフォーム上で、エージェントが移動してコンポーネントを使って自律的に処理を行って帰ってくるプログラムを作成する。

知的エージェントやネットワークエージェントにはモビリティ、自律性、知性を実現させなくてはならないので、Javaで実装されるのが一般的である。モバイルエージェントのようにネットワーク上を移動するシステムは環境に依存しないように動作させる必要があるが、Javaは現在使われているほぼすべてのOS (Operating System) で利用可能であり、Javaの実行環境を用意することによって、どのような環境でも動作させることが可能である。また、自律したプログラムを実現するためには長時間の実行や待機する必要になるが、これもJavaではプ

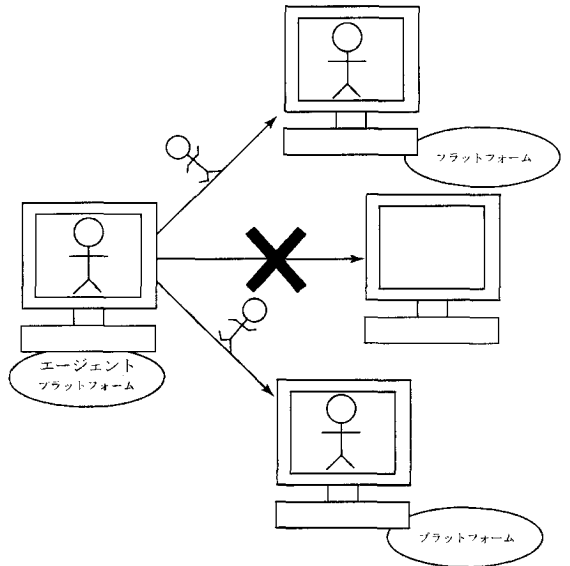


図3 エージェントの行動条件

ロセスまたはスレッドとして実現可能である [9]。

図3に記すようにエージェントはプラットフォームの中のみ活動することができる。複数のプラットフォームを用意することによってエージェントはプラットフォーム間を移動でき、移動先のプラットフォーム上で作業することができる。また、プラットフォームには特有のコンポーネントがあり、それをエージェントに与えることによってエージェントは様々な処理を行うことができる。

エージェントの動作の流れについて説明する (図4)。2章で述べた通り、エージェントとは、ユーザの様々な事情により自分自身で行うことが困難なタスクをユーザに代わって行う動作主体である。エージェントは、ユーザから検索要求を受け取ると、この検索要求を遂行するためにデフォルトサーバプラットフォームに移動する。移動が終了すると、このプラットフォームのコンポーネントを受け取って仕事を行う。仕事が正常に終了すれば、ユーザプラットフォームに戻り、ユーザに結果を報告して次の要求が来るまで待機する。もし、正常に終了しない場合は、別プラットフォームに移動して再実行を行う。しかし、移動先サーバプラットフォームでエージェントがどこにいるのか把握できるように、エージェントは移動元サーバプラットフォームに自分のIDと、どこへ移動したのか履歴を残して行く。タスクを全て終了したエージェントがユーザプラットフォームに戻って来る際にユーザプラットフォームが閉じられている可能性があるが、その場合、エージェントはデフォルトサーバプラットフォームに待機し、ユーザプラットフォームが再起動するのを待つ。ユーザプラットフォームが再起動すれば、ユーザプラットフォームに戻るよう指示される。ユーザプラットフォームに戻ったエージェントはユーザからの検索要求を待つ。

ユーザプラットフォームについて説明する。ユーザプラット

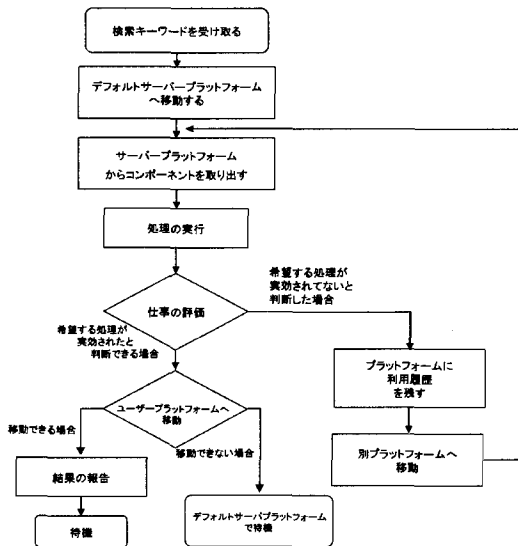


図4 エージェントのライフサイクル

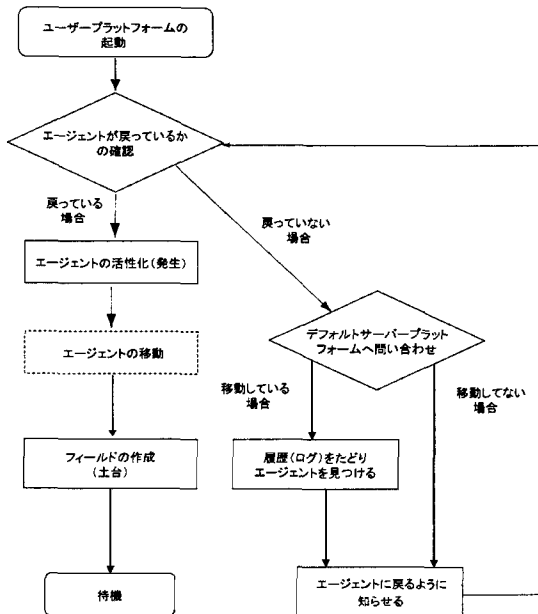


図5 ユーザプラットフォームのフローチャート

フォームはユーザ側で動作するエージェント用の活動環境である。このプラットフォームはユーザによって立ち上げられる。このプラットフォームの大きな役割はエージェントの発生または管理、別プラットフォームへ移動したエージェントの待ち受けのみである。このプラットフォームの仕事はエージェントが前回起動時に戻ってきているのかということ履歴から参照すること、エージェントが戻ってくるののためにフィールドを用意することである。

図5に記すように、ユーザプラットフォームは起動されるとエージェントがユーザプラットフォームに戻っているか確認

する。もし戻っているのであれば、エージェントを活性化(発生)させる。エージェントがユーザプラットフォームから出て行くとフィールドを作成し、エージェントが戻ってくるのを待つ。戻ってくると、エージェントを活性化した状態に戻る。次にユーザプラットフォームを起動してエージェントが戻ってきていなければデフォルトプラットフォームに問い合わせる。デフォルトプラットフォームにいない場合、デフォルトプラットフォームに残された履歴を参照してエージェントを探していく。エージェントが見つければエージェントに戻るよう指示して、エージェントが戻って来るまで待つ。帰るよう指示して一定の時間が過ぎるともう一度エージェントが戻っているかの確認をする。

サーバプラットフォームについて説明する。サーバプラットフォームはサーバ側で動作するエージェント用の活動環境である。このプラットフォームは基本的に常時稼働している。サーバプラットフォームもユーザプラットフォームと同じように、エージェントが移動してくるののためにフィールドを用意する。エージェントが移動してくるとコンポーネントを渡し、活動環境を提供する。また、エージェントがユーザプラットフォームに戻れなくなった場合、エージェント滞在時間を越えるとエージェントを削除する。

図6に記すように、起動された時にエージェントを待ち受けるためのフィールドを作成する。エージェントが移動してきたら、エージェントが行動できるように活性化する。次に、エージェントがこのプラットフォーム上で作業を行うためのコンポーネントを渡し、エージェントが作業を終了すれば作業領域を閉じる。また、エージェントにデフォルトサーバプラットフォームと指定されている場合、エージェントの滞在時間を確認し、一定時間以上居るのであれば削除を行う。

以上のことを考慮し、システムのコーディングを行い動作確認を行った。本試作システムはエージェントの自律的な行動、モビリティの動作確認を行うもので、エージェントにキーワードを渡すとエージェントはサーバプラットフォームに移動し、検索結果が得られたら帰還し、結果をユーザに返すプログラムである。

本研究で提案したエージェントシステムの結果は図7と図7となった。まず、ユーザプラットフォーム側処理結果(図7)の(1)~(6)について説明する。

- (1) エージェントを活性化させて、エージェントを起動
- (2) エージェントからキーワードを入力するように促され、キーワード「Agent」を入力
- (3) エージェントがキーワードをもって、サーバプラットフォームに移動
- (4) エージェントに絞り込み処理をさせた結果を表示
- (5) (2)と違ったキーワードの入力
- (6) 以前検索したキーワードを反映して、調べたいキーワード「Network」とのAND検索を行い絞り込みをした

本研究では以上のような結果となった。しかし、検索結果の絞り込みについては新しいアルゴリズムを考えた訳ではないのでアルゴリズムについては今後の研究で行う。

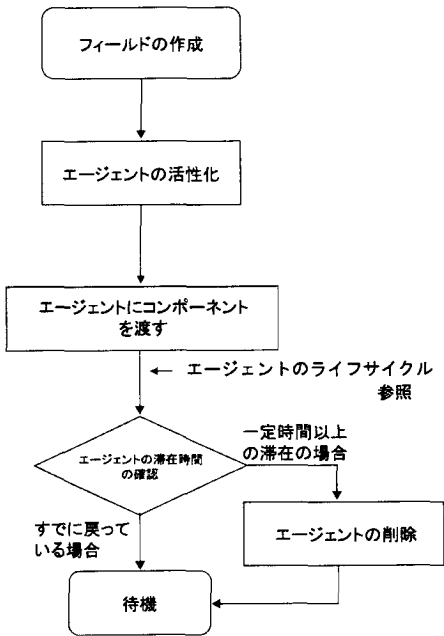


図6 サーバプラットフォームのフローチャート

5. まとめ

近年、インターネットが普及してきたことによって、非常に多くの情報やサービスを利用することができるようになってきている。しかし、インターネットで公開されている膨大な情報の中からユーザの欲する情報にマッチしたものを捜し出すにはある程度の経験や知識が必要となる。これを利用する技術はまだ確立されている状況ではなく、また万人にとって使いやすいシステムを開発することは非常に困難である。本研究で取り上げた知的ネットワークエージェントとは、ネットワーク上を移動する移動する機能を持ったネットワークエージェントと、知識を持ってユーザをサポートすることのできる知的エージェントとしての二つの側面を持つエージェントである。この知的ネットワークエージェントはネットワーク上を移動しながらユーザから与えられた要求に対し自らの判断でタスクを処理していく。このとき、ユーザは要求を達成するために必要なタスクに関しての専門的な知識や経験を必要としない。本研究ではこの利点を Web 検索システムに応用できると考え、そのための知的ネットワークエージェントシステムの開発を進めてきた。

本報告では Web 検索を対象とし、知的ネットワークエージェントにユーザと絞り込みを検索機構との仲介役をさせることでユーザの知識・経験に依存しないシステムを提案、実装した。これによりユーザが必要とする情報やサービスへのアクセスを高速化、また効率化図った。しかし、今回はエージェントが検索した検索結果に対して、ユーザの要求にどの程度マッチしたのかを評価していないため、エージェントの仕事に対しての信頼性に関して確認を得ることができなかった。

今後は実験により確認出来なかったエージェントの検証結果

Machine City:PlatformInformation: IP Address:192.168.12.124

Agent started...
I am a Agent.

My name is Smith Information

Please Input Keyword

Start Server Platform

Wait for Own Agent...

Agent

Smith:SimpleMove: Let's move to /192.168.12.7:5000

Result:

--- Search Result ---

Title: Agent 4.2 - Newsreader + Email

Uri: http://www.forteinc.com/agent

Title: intelligent agents and bots - AgentLand 2005

Uri: http://www.agentland.com/

Title: Agent Newsreader & Premium Newsgroups

Uri: http://www.forteinc.com/

I am a Agent.

My name is Smith Information

Please Input Keyword

Network

Smith:SimpleMove: Let's move to /192.168.12.7:5000

Result:

--- Search Result ---

Title: SourceForge.net: Network Agents

Uri: http://sourceforge.net/projects/networkagent/

Title: The Network Agents

Uri: http://www.thenetworkagents.com/

Title: e-PRO Internet Real Estate Agent Network

Uri: http://www.ireagents.com/

図7 実験結果

における信頼度について検証を行う。また、今回は知的ネットワークエージェントシステムをコマンドラインプログラムにより実現したが、これではユーザにとって直感的な操作できないため、利便性の向上を目的として GUI (Graphical User Interface) を追加することも必要である。さらに、本提案では Web 検索を対象としたシステムであったが、検索以外の機能についても今後検討していく。

文 献

- [1] 荒屋真二, 人工知能概論 -- コンピュータ知能から Web 知能まで. 共立出版, 2006.
- [2] 本位田真一, 飯島正, 大須賀昭彦, エージェント技術. 共立出版, 2000.
- [3] "Aglets," (参照 2007-02-20). http://www.research.ibm.com/trl/aglets/index_e.htm.
- [4] "Plangent," (参照 2007-02-20). http://www2.toshiba.co.jp/rdc/plangent/index_plangent_j.htm.
- [5] 原田昌紀, 清水奨, "WWW 検索システムにおける不特定多数の操作履歴の活用," (参照 2007-04-10). <http://www.ingrid.org/~harada/publications/SIGDPS81/dps9702.pdf>.
- [6] 福原知宏, "協調フィルタリングに関する研究動向," (参照 2007-04-12). <http://www.race.u-tokyo.ac.jp/>

~/fukuhara/Research/paper/98/cofil.pdf.

- [7] 西村 悟, 千野 晋平, 三木 光範, 廣安 知之, “Folksonomy ～タグで繋がるみんなの分類～,” (参照 2007-04-10) . <http://mikilab.doshisha.ac.jp/dia/research/report/2005/0911/001/report%20050911001.html>.
- [8] 長沼武史, 菊池悠, 稲村浩, 倉掛正治, “タスク知識に基づくモバイルユーザ支援システム,” 情報処理学会, 2004.
- [9] 新谷虎松, *Java* による知能プログラミング入門. コロナ社, 2006.