

個人支援エージェントにおける Preference Elicitation 最適化に基づく競合調整機構の試作 A Preliminary Implementation of a Negotiation-based Cooperation Mechanism for User Assistance Agents and Humans

大石翔*1
Sho Oishi

福田直樹*2
Naoki Fukuta

*1静岡大学情報学部情報科学科

Department of Computer Science, Shizuoka University

*2静岡大学大学院情報学領域

College of Informatics, Academic Institute, Shizuoka University

In this paper, we present our initial idea and an overview of an automated negotiation mechanism that allows agents to request information about their preferences as well as having some assistances from real-world humans. To avoid bothering users to ask about their preferences, we also introduce a mechanism and its prototype implementation to predict “when to ask”.

1. はじめに

何らかのタスクを遂行することで、個人支援を行うようなソフトウェアエージェントの能力は様々なものになると考えられる [Oishi 16a]。このようなエージェントを用いたアプリケーションの例としては、P2P ネットワーク上に分散されたファイルの収集を手助けするようなアプリケーションや [Fukuta 12]、インターネットオークションサイトにおいて、協調してアイテムを検索、適切な入札を行うようなアプリケーションが考えられる [Ito 00]。このような能力は、エージェントに許されているリソースや、エージェントを利用しているユーザの状況によって様々なものとなる [Oishi 16a]。それゆえ、タスクを遂行するためには、ユーザとエージェントの状況を考慮し、ソフトウェアエージェントと人が可能な能力を適切に組み合わせて利用することは非常に重要である。

このような課題を解決するために、複数のエージェントの適切な振る舞いを強化学習に基づく手法を用いて学習させることが提案されている [Arai 00]。また、既存のサービスやエージェントの構成のセマンティクスを表現し、効率的に処理することは解決すべき重要な課題である [Sirin 04]。しかしながら、それぞれのタスクは独立して構成されている場合においても、いくつかのリソースを共有している場合においては、タスクが競合する可能性がある。

[Oishi 16a] において、個人支援を目的としたエージェントを個人支援エージェントと定義している。個人支援エージェントは、動作している環境およびハードウェアから制約を受け、単一のエージェントでは、全てのタスクを遂行することが困難になる場合がある。この課題を解決するために、我々は複数の個人支援エージェントを動作可能なシステムの設計及び試作を行った [Oishi 16a]。本システム上では、複数の個人支援エージェントが、それぞれ異なるハードウェアおよび実行環境で動作する。能力オントロジーに基づき全てのエージェントの能力に関する情報を共有する機構を用意し、エージェントはそれぞれの実行環境の制約を考慮したタスク遂行が可能である [Oishi 16b, Oishi 16a]。[Oishi 16b] においては、個人支援エージェントの能力をどのようにタスク遂行の過程として利用するかについても提案し、能力オントロジーを利用することで、エージェントの能力間の暗黙的な関係に基づく様々な能力を参照することが可能である。一方で、エージェントは、タス

クに合わせたノルムを機械学習手法を用いて学習することも可能である [Riemsdijk 15]。学習したノルムは、タスク遂行過程で行動を決定するための手助けをし、不適切なエージェントの行動を制限する。

しかし、これらのノルムによる制約は、複数のエージェントが全てのタスクを同時に協調しながら遂行する場合を考慮していない。特定のタスクにおいては、エージェントのノルムと現実世界の制約により、エージェントのみの協調的タスク遂行では遂行が困難である一方で、人であれば容易に遂行可能である場合がある [Oishi 16c]。

本論文では、自動交渉に基づき、エージェントがユーザの好みを探りながら現実世界の人との協調を実現することで、エージェントのタスク遂行競合を調整する機構の試作について述べる。エージェントがユーザの好みについて尋ねることでユーザの邪魔することを避けるために、[Chen 16] で提案されている仕組みに基づく「いつ尋ねるか」を予測する機構とその試作について述べる。

2. システムの試作

2.1 個人支援エージェントシステムの実装

我々は、個人支援エージェントが柔軟に様々なハードウェアシステム上で動作可能なシステムを試作している [Oishi 16a]。本節では、[Oishi 16a, Oishi 16b, Oishi 16c] において提案した本システムの概要について述べる。図 1 に、動作している個人支援エージェントの外観を示す。本プラットフォーム上では、いくつかの個人支援エージェントが異なる環境で同時に動作することができる。図 1 において、いくつかのエージェントは仮想マシンとして動作している Ubuntu Linux OS で動作し、他のエージェントはホスト OS である OS X 上で動作している。本システムは 3 つのソフトウェアモジュールから構成されている。個人支援エージェントの動作環境において、個人支援エージェントを管理するクライアントサイドエージェントマネージャと、サーバサイドで動作し、異なる動作環境間でのエージェント間通信を仲介するエージェントコミュニケーションマネージャおよび個人支援エージェントである。図 2 に本システムの構成を示す [Oishi 16b]。個人支援エージェントはクライアントサイドエージェントマネージャの機能を利用することで、エージェントコミュニケーションマネージャを介して、他の動作環境上のエージェントと通信することが可能である。本プラットフォームにおいてエージェントは、他の動作環境上

Personal Assistant Agent

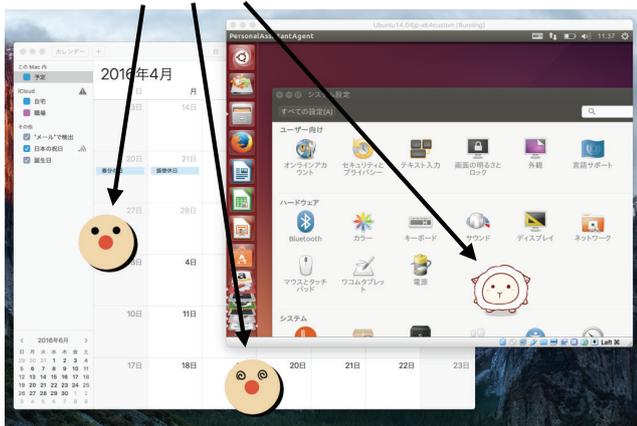


図 1: 個人支援エージェントの動作例 [Oishi 16a]

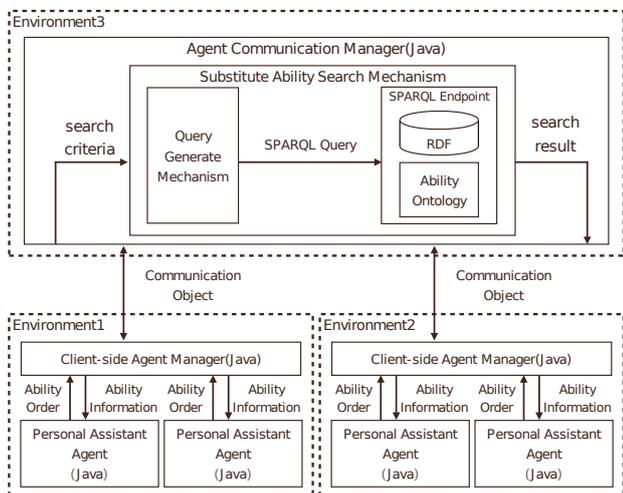


図 2: 個人支援エージェントシステムの構成 [Oishi 16b]

のエージェントとの協調的な動作によってユーザを支援することが可能である。協調的なタスク遂行のために必要な能力の検索とその能力の利用を可能とするために、エージェントコミュニケーションマネージャ上には、SPARQL クエリを活用する機構を準備している。

個人支援エージェントはバックグラウンドでユーザの支援を行うため、ユーザは、タスク遂行時のエージェントの状態を知ることが困難になる可能性がある。タスク遂行時のエージェントの状態を、ユーザに対して視覚的に提供するために、本プラットフォームではそれぞれの個人支援エージェントの GUI として図 1 のような外観を準備している。図 1 では、一部のエージェントが外観を変化させることで、ユーザにタスク遂行に失敗したことを知らせている。

エージェントには、ユーザから直接教師信号を得ることで、強化学習を用いて、環境からいつ行動すべきかを学習する機構を用意している。この機構を用いることで、ユーザはエージェントがユーザにとって望ましくない動作をした場合に、エージェントをクリックするだけで教師信号を与えることができる。

2.2 個人支援エージェントの強化学習機構

本システム [Oishi 16a] において、個人支援エージェントはユーザの状態を知るために、動作している OS から情報を参照する。エージェントが参照する情報の例としては、エージェントと並列して動作しているプロセス名のリストとメモリ使用率及び CPU 使用率が挙げられる。本プラットフォームでは、エージェントはこれらの情報にアクセスするために ps, top といった unix コマンドを利用する。エージェントは、これらの情報から推定したユーザの現在の状態と予測した将来の状態に基づく最適な行動を、エージェント内部の学習器を用いて学習することができる [Oishi 16a]。

現在の実装では、強化学習アルゴリズムの 1 つである Q-Learning [Sutton 98] を拡張したアルゴリズムを実装している。Q 値は実環境で学習させたものではなく、強化学習を行う環境シミュレータを用意し、シミュレータ上で再現した環境で学習させたものとする [Oishi 16a]。エージェントは、シミュレータ上でユーザの作業状況のモデルから状態を推定、行動選択を行い、報酬を改めてユーザの作業状況から得ることで学習を行う。さらに、ユーザは図 1 の GUI を介して、エージェントに教師信号として負の報酬を与えることで学習することも可能である。シミュレータ上で用いるユーザの作業状況のモデルは、マルコフ決定過程に従い、学習可能なものであると仮定している。エージェントの行動は、エージェントが内部に持つ能力ライブラリに含まれる能力から決定される。各エージェントが持つ能力ライブラリは、エージェントの持つタスクに合わせた能力かつ、そのエージェントの動作環境で処理可能な能力で構成される。

しかし、これらは「いつ実行するか」と「なにを実行するか」を特定の環境から学習するものであり、ユーザの嗜好やそれをユーザに尋ねるための仕組みは準備されていない。

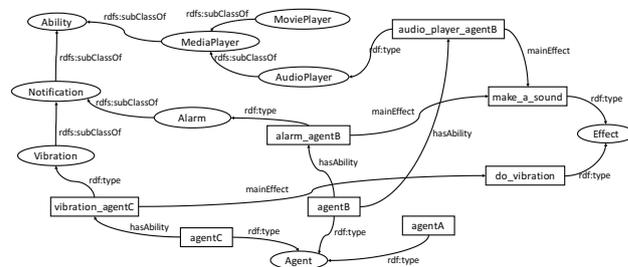


図 3: 能力オントロジーの例

2.3 能力検索機構の実装

エージェントは必要な能力に関する情報をオントロジーに基づき取得可能とするために、本システム内部で SPARQL エンドポイントを扱う機構を準備している [Oishi 16b]。SPARQL^{*1} は W3C に勧告された標準クエリ言語であり、ストアデータに対する意味的な検索を実現するためにの OWL ベースのオントロジーの定義を受け付ける。本機構では、エージェントは SPARQL エンドポイントに対して能力に関する情報を参照するクエリを用いて、検索条件を満たす能力のリストを得ることができる [Oishi 16b]。[Oishi 16b] の実装では、検索条件は必要とされるハードウェアや能力からなり、[Oishi 16c] において、ユーザがインセンティブされて行う行動を扱うこと

*1 <https://www.w3.org/TR/sparql11-query/>

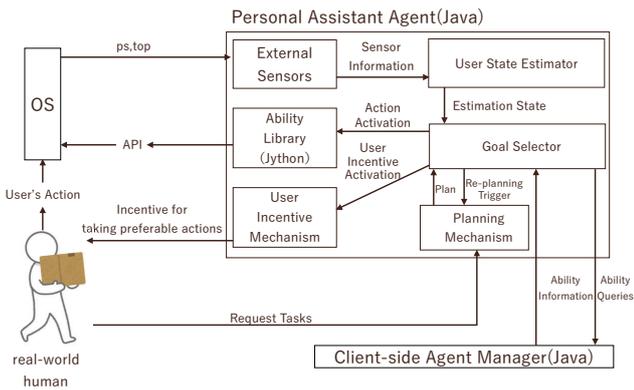


図 4: 拡張した個人支援エージェントの構成 [Oishi 16c]

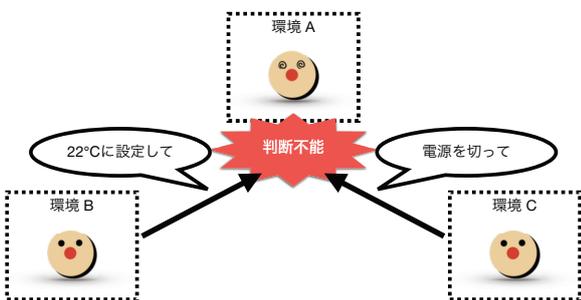


図 5: タスク遂行を行うエージェント間の競合

ができるよう拡張した。図 3 に能力オントロジーの例を示す [Oishi 16b]。エージェントは、いくつかのタスクを遂行することでユーザを助けるために必要な能力オントロジーを用いて、モデル化可能であると仮定している。本システムは、このようなモデルに基づき、エージェントが十分な能力のインスタンスを探索するための支援を行う [Oishi 16b]。

2.4 人とエージェントにおける協調的タスク遂行

我々は、エージェントがタスクを遂行するために他のエージェントの能力を活用するとともに、現実世界の人に対して、個人支援エージェントでは遂行が困難なタスクの遂行をリクエストする機構を試作している [Oishi 16c]。本機構では、人の行動を含むタスク遂行を扱うために、拡張した能力オントロジーに対して、SPARQL エンドポイントを介して特定の人物へタスクの遂行を頼むことが可能なエージェントを検索することが可能である。能力は、エージェントのタスク遂行にとって望ましい結果に加えて様々な副作用があることを考慮すべきである。図 4 に、本機構を含む、拡張した個人支援エージェントの構成 [Oishi 16c] を示す。

本機構では、異なる目的を持つエージェント間での競合は考慮されていない。本論文では、実世界のタスク遂行における競合を解決するための、自動交渉技術と [Chen 16] の手法に基づく Preference-elicitation [Lu 11] を組み合わせた機構の試作について述べる。

3. 交渉に基づく人とエージェントの協調

本節では、自動交渉に基づく競合調整機構の動作例を図 6 のシナリオに従い述べる。図 6 のシナリオでは、タスク遂行競合

の例として、異なる選好を持つ人が混在する部屋の温度を調節するというタスクを想定する。部屋には、ユーザ A とユーザ B がいるとする。部屋の温度はエアコンによって、 20°C に保たれている。ユーザ A は、エアコンの設定温度を 22°C に設定したいと考えている。一方で、ユーザ B は現在の室温に満足しているが、そろそろ換気をしたいと考えている。エアコンを操作するリモコンはユーザ B の手元にあり、個人支援エージェントは直接エアコンの操作が行えないものとする。ユーザ A は、個人支援エージェント A に温度を 22°C に設定するように指示し、ユーザ B は、個人支援エージェント B に部屋を換気したいのでエアコンの電源を切るように指示する場合について考える。個人支援エージェント A と B のタスクである「エアコンの設定温度を 22°C に設定する」と「エアコンの電源を切る」は競合しており、同時に行うことは困難であると考えられる。そこで本機構では、個人支援エージェント A と B の間で交渉を行い、この競合を調整する。エージェントは、温度設定及び時間がどの程度ユーザにとって好ましいか判断するための知識を持たず、必要に応じて Preference-elicitation [Lu 11] を行う。Preference-elicitation に基づく交渉では、エージェントは、交渉中の論点に対する選好を公開することで交渉を進める。



図 6: 2 者間でエアコンを共有する場合の動作例

本機構において、交渉は複数者間交渉プロトコルの 1 つである Stacked Alternating Offers Protocol に従って行われ、この例では 2 者間の交渉であるが、実際には 3 者以上での交渉が可能である。交渉を始める前にエージェントは、ユーザの効用空間についての知識を持たず、Offer された提案の効用を計算するための知識を持たないものとする。そこで、本機構では、[Chen 16, Baarslag 15] の手法に基づき、必要に応じてユーザに対して効用を尋ねながら行動を選択、交渉を行う。この例では、論点をエアコンの温度と時間とし、提案を { 設定温度, 設定温度を保つ時間 (分) } と表記する。図 7 に、この例における交渉の概観を示す。

まず、エージェント A は Offer を選択し、提案 $\{22^{\circ}\text{C}, *(\text{時間指定なし})\}$ をエージェント B に提案する。この提案に対して、エージェント B はユーザ B に対してこの提案の効用を尋ね、この例における効用は 0.1 とする。エージェント B は提案を破棄し、新たな提案 $\{0^{\circ}\text{C}, 10\}$ を Offer する。このシナリオでは、10 分後からユーザ A は外出する予定があるとする。エージェント A は、ユーザ A に対して提案 $\{0^{\circ}\text{C}, 10\}$ の効用を尋ね、この例における効用は 0.1 とする。エージェント A は提案 $\{0^{\circ}\text{C}, 10\}$ を破棄し、新たな提案 $\{22^{\circ}\text{C}, 10\}$ を Offer する。この提案は 10 分間 22°C にエアコンを設定するという提

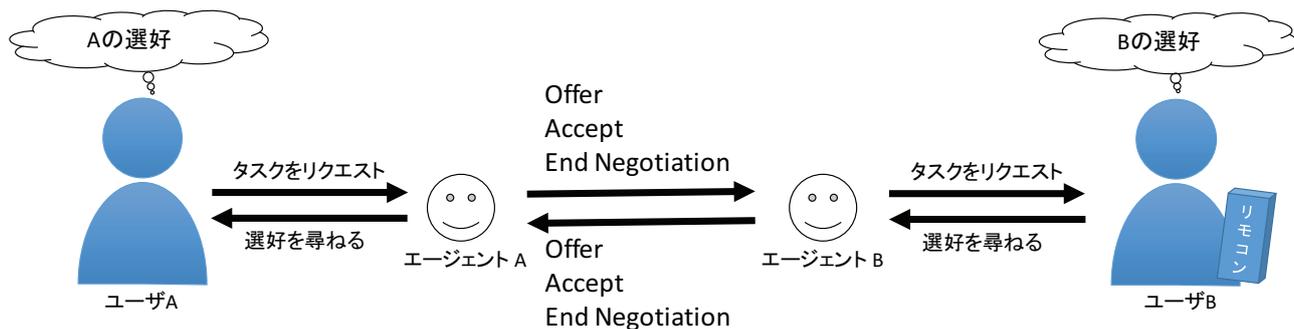


図 7: 本機構における 2 者間交渉の概観

案である。この例では、ユーザ A は 10 分後に外出するものとする。この提案に対して、エージェント B はユーザ B に対して効用を尋ね、この例では、この提案に対するユーザ B の効用値を 0.6 とし、エージェント B は Accept を選択するものとする。結果、エージェント A のタスク「エアコンの設定温度を 22°C に設定する」は「10 分間エアコンの設定温度を 22°C に設定する」となり、エージェント B のタスク「エアコンの電源を切る」は「10 分後エアコンの電源を切る」となり、タスクを実行するタイミングが調整され、競合が解消される。この例では、交渉を行うことで、タスク遂行を行うタイミングをずらすことによる競合解消を実現したが、交渉の際にどのような論点に関して調整することで競合を解消できる余地があるかを事前にどこまで想定できるかは、今後の課題となる。例えば、論点として席の配置を加え、部屋の中でも席によっては温度が異なると仮定した場合、席の配置を変えることで競合を解消できるかもしれないが、この仮定をあらかじめエージェントの中に組み込んでおくべきかどうかについては、必ずしも自明ではない。

参考文献

- [Arai 00] Arai, S., Sycara, K., and Payne, T. R.: Multi-agent reinforcement learning for planning and scheduling multiple goals, in *Proceedings Fourth International Conference on MultiAgent Systems*, pp. 359–360 (2000)
- [Baarslag 15] Baarslag, T. and Gerding, E. H.: Optimal Incremental Preference Elicitation during Negotiation, in Yang, Q. and Wooldridge, M. eds., *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pp. 3–9, AAAI Press (2015)
- [Chen 16] Chen, S., Baarslag, T., Zhao, D., Chen, J., and Shen, L.: A Polynomial Time Optimal Algorithm for Robot-Human Search under Uncertainty, in Kambhampati, S. ed., *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*, pp. 819–825, IJCAI/AAAI Press (2016)
- [Fukuta 12] Fukuta, N.: A Mobile Agent Approach for P2P-based Semantic File Retrieval, *Journal of Information Processing*, Vol. 20, No. 3, pp. 607–613 (2012)
- [Ito 00] Ito, T., Fukuta, N., Shintani, T., and Sycara, K.: BiddingBot: a multiagent support system for cooperative bidding in multiple auctions, in *Proceedings Fourth International Conference on MultiAgent Systems*, pp. 399–400 (2000)
- [Lu 11] Lu, T. and Boutilier, C.: Robust Approximation and Incremental Elicitation in Voting Protocols, in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume One, IJCAI'11*, pp. 287–293, AAAI Press (2011)
- [Oishi 16a] Oishi, S. and Fukuta, N.: A Cooperative Task Execution Mechanism for Personal Assistant Agents Using Ability Ontology, in *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pp. 664–667 (2016)
- [Oishi 16b] Oishi, S. and Fukuta, N.: Toward a Flexible Ability Selection Mechanism for Personal Assistant Agent using Ontology Reasoning, in *2016 IEEE/WIC/ACM International Conference on Web Intelligence Workshops (WIW)*, pp. 92–95 (2016)
- [Oishi 16c] Oishi, S. and Fukuta, N.: Toward a User Incentive Mechanism to Accomplish Unworkable Tasks for Agents, in *2016 IEEE International Conference on Agents (ICA)*, pp. 125–126 (2016)
- [Riemsdijk 15] Riemsdijk, van M. B., Dennis, L., Fisher, M., and Hindriks, K. V.: A Semantic Framework for Socially Adaptive Agents: Towards Strong Norm Compliance, in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS '15*, pp. 423–432, Richland, SC (2015), International Foundation for Autonomous Agents and Multiagent Systems
- [Sirin 04] Sirin, E., Parsia, B., Wu, D., Hendler, J., and Nau, D.: HTN Planning for Web Service Composition Using SHOP2, *Web Semant.*, Vol. 1, No. 4, pp. 377–396 (2004)
- [Sutton 98] Sutton, R. S. and Barto, A. G.: *Introduction to Reinforcement Learning*, MIT Press, Cambridge, MA, USA, 1st edition (1998)