

個人支援エージェントにおけるユーザの負担を考慮した Preference Elicitation 最適化手法の検討

A Preliminary Approach on Better Preference Elicitation with Timing Optimization for Personal Assistant Agents

大石翔^{*1}
Sho Oishi

福田直樹^{*2}
Naoki Fukuta

^{*1}静岡大学大学院総合科学技術研究科

Graduate School of Integrated Science and Technology, Shizuoka University

^{*2}静岡大学学術院情報学領域

College of Informatics, Academic Institute, Shizuoka University

本論文では、個人支援エージェントにおけるユーザの負担を考慮した Preference Elicitation 最適化手法の検討について述べる。ユーザの作業のバックグラウンドで動作するソフトウェアエージェント間で交渉を行う場合において、事前に交渉で扱われる論点に対するユーザの Preference に関する知識をエージェントに組み込むこととあわせて、Preference Elicitation を適切なタイミングで行い、ユーザから Preference に関する知識を抽出することで、他のユーザとの競合をより望ましい形で解決できる可能性について検討する。Preference Elicitation を行う際には、エージェントはユーザに対して直接 Preference を尋ねることが増加することはユーザの負担になると同時に、プライバシー情報を含む Preference が意図せず開示されてしまう場合への対処方法の検討について述べる。本論文では、ユーザの好ましいタイミングで Preference について尋ねるための行動タイミングを、QS-Learning を用いて事前にエージェントに学習させることで、ユーザの負担を考慮した Preference Elicitation を行う機構の設計について述べる。

1. はじめに

我々は、エージェントがユーザのバックグラウンドで動作しタスクを遂行することで、ユーザの作業の支援を行うエージェントを個人支援エージェントと定義し、試作している。個人支援エージェントはステークホルダーであるユーザの目的に合わせて適切な能力を持ったエージェントであり、ユーザが任意の機能を実装可能である。個人支援エージェントの例としては、ユーザにとって好ましい情報を SNS 上から取得し、ユーザに適切な形で通知するようなエージェントや [Dinakar 15, Oishi 17a], ユーザのスマートフォン上で動作して他者からの連絡にたいしてユーザの状況に応じた形で着信を通知するエージェントが考えられる [Ajmeri 17]。エージェントの持つ能力は動作する環境に依存している場合もあり、一部の資源を複数のエージェントが共有している環境において、エージェントのタスク遂行が困難になることが考えられる [Oishi 17b]。タスク遂行の競合を解消するために、我々は個人支援エージェント間で交渉を行うことで競合を解消する機構を試作している。

エージェントがユーザのバックグラウンドで交渉を行う場合、交渉の論点に対するユーザの Preference に関する知識をエージェントに組み込む必要がある。試作している個人支援エージェントは、タスク競合の解消の結果として、よりユーザに望ましいタスク遂行方法を見つけるために、交渉の過程で必要に応じて Preference に関する知識をステークホルダーであるユーザに尋ねることを許可している。Preference Elicitation においては、エージェントが Preference に関する知識を得るためにステークホルダーに対して尋ねる回数が増えることはユーザの負担になり、質問内容、回数およびタイミングを最適化することは重要な課題である。Baarslag らはエージェントが過去の交渉から曖昧性のある効用に関する知識を持っている場合において、ユーザに尋ねるコストを考慮して質問の内容を最適化する手法を提案している [Baarslag 15]。Baarslag ら

の手法では、交渉の過程でユーザから Preference に関する知識を尋ねた上で交渉を行なった場合に得られる効用の期待値と尋ねなかった場合に得られる効用の期待値を比較することでユーザが被る Preference Elicitation のコストを最小化する。我々が試作している個人支援エージェントでは、Baarslag らの手法を拡張し、エージェントが Preference Elicitation の際にユーザに対して尋ねる質問内容とタイミングを最適化する機構を試作している。次節において、Preference Elicitation を行うタイミングを最適化するために事前に個人支援エージェントに行う学習について述べる。

2. ユーザの負担を考慮した Preference Elicitation タイミングの最適化

本節では、ステークホルダーであるユーザの負担を考慮した Preference を尋ねるタイミングの学習について述べる。ユーザの状況を考慮したタイミングで Preference Elicitation すべき状況の例を、図 1 に示す。

図 1 の状況において、アリスは喫茶店でラップトップを開いてレポートを書いており、たかしはアリスに連絡をしたいことがある。アリスとたかしの持つコンピュータ上では個人支援エージェントがそれぞれ動作している。アリスのいる喫茶店には他にも客があり、アリスは自身のエージェントにレポートを書く邪魔をしないよう指示している (①)。たかしはアリスに早急に連絡したいことがあり、自身が行なっている作業のバックグラウンドでアリスに連絡することをエージェントに指示している (①)。たかしをステークホルダーとするエージェントは、アリスに早急に指示を出すためにアリスのスマートフォンから最大の音量で通知するという手段を選ぶ (②)。アリスのスマートフォン上で動作しているアリスをステークホルダーとするエージェントは、アリスの邪魔をしないというタスクを遂行しているため、たかしのエージェントとの間でタスクの遂行が競合する。たかしとアリスの個人支援エージェントは競合を解消するために交渉を行い、お互いのタスクが遂行できるよ

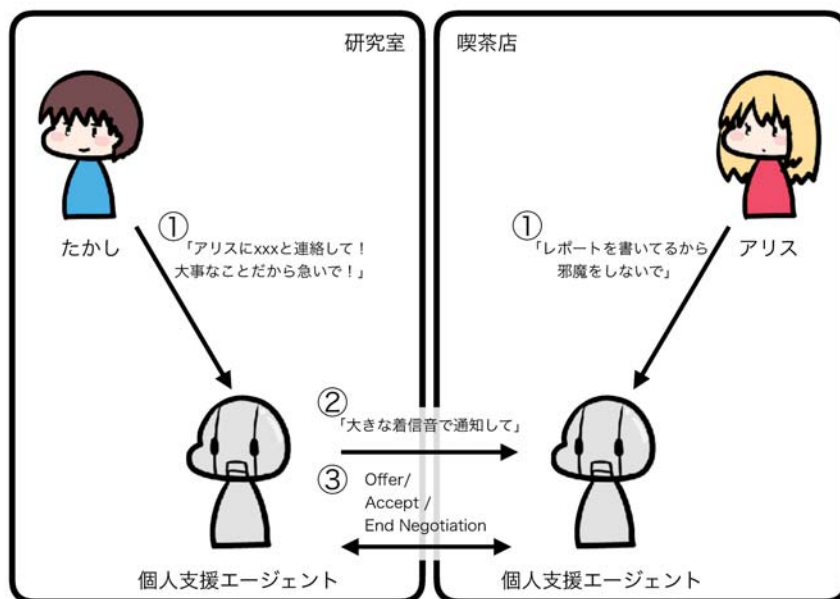


図 1: 個人支援エージェントの交渉支援の例

うな妥協案を探すことが可能である (③). 交渉の結果導き出される妥協案の例としては, アリスを邪魔しないことを優先してアリスがレポートを書き終えるのを待って通知する場合や, たかしからの連絡の緊急性が優先されバイブレーションを用いて通知する場合が考えられる. ソフトウェアエージェントである個人支援エージェントはこのような妥協案を見つけるために, 交渉を行う過程で必要に応じて Preference Elicitation を行い, それぞれのステークホルダーから Preference に関する知識を得る必要がある. アリスのエージェントはアリスに対して質問を行うが, 質問にはアリスのプライバシーに関する情報が含まれる場合に, Preference Elicitation を行うことでアリスのプライバシーが周囲の喫茶店の客に漏洩してしまう可能性がある. Preference が複数のユーザのプライバシーを含む場合において, プライバシーの漏洩を防ぐために, ステークホルダー以外の人やエージェントへの影響を考慮して個人支援エージェントが Preference Elicitation を行うか判断する必要がある.

この課題を解決するために, 本研究では個人支援エージェントに強化学習を用いて事前にエージェントに最適な Preference Elicitation を行うタイミングの学習を行う. 我々の試作している個人支援エージェントは, ユーザを支援するために行動したことが原因となり, ユーザの邪魔をしてしまうことを避けるために事前に強化学習を用いて行動タイミングを学習する機構を準備している [大石 16a, 大石 16b, Oishi 16]. 本研究ではこの機構を拡張し, 個人支援エージェントに複数のステークホルダーのもつコンテキストを考慮した Preference Elicitation を行うタイミングを学習させることを狙う.

Q-Learning[Sutton 98] のような強化学習手法では, エージェントが取りうる可能性のある全ての状態に対して最適な行動を網羅的に探索する必要がある, エージェントに効率的に探索を行わせ学習の収束を早めることは重要な課題である. Q-Learning を拡張し, 学習の収束が高速化された手法としては, Rosenfeld らの提案した QS-Learning[Ariel Rosenfeld 17] が

あり, エージェントが取りうる状態のうち, 類似した状態をまとめて1つの状態として扱うことでエージェントの探索範囲を狭めることで学習の収束を早めることが示されている. 本研究では, [大石 16a, 大石 16b, Oishi 16] の強化学習用環境シミュレータを拡張し, エージェントに QS-Learning を用いて Preference についてユーザに尋ねるタイミングを学習させるための強化学習環境シミュレータを試作している.

本シミュレータでは, 学習エージェント, ステークホルダーモデル及びエージェントが動作する環境を想定した環境モデルを入力に学習を行なうことが可能である. それぞれが事前に準備したクラスを継承し実装することが可能であり, 学習エージェントは, 同時に動作する複数のエージェントの行動とステークホルダーの行動が影響を受ける環境モデルから事前に定義した観測情報のみを受け取り, 学習を行なう. 学習エージェントには, 学習の中で取りうる行動と観測する情報を定義し, ステークホルダーモデルにはステークホルダーの行動と対応するエージェントの行動に対する報酬関数を定義する必要がある. 学習を高速化するために, 観測情報から異なる状態を類似の状態として判断する関数を学習エージェントに与えることで, QS-Learning を用いて学習を行なうことも可能である. 環境モデルには個人支援エージェントが観測可能な情報源のモデルを組み込む必要があり, ステークホルダーモデルに定義した行動による観測情報の変化とステークホルダーの行動に関わらず学習のステップが進むことによって起きる値の変化を近似した関数として定義する.

3. おわりに

本論文では, 個人支援エージェントにおけるユーザの負担を考慮した Preference Elicitation 最適化手法の検討について述べた. ユーザの作業のバックグラウンドで動作するソフトウェアエージェント間で交渉を行う場合において, 事前に交渉で扱われる論点に対するユーザの Preference に関する知識をエージェントに組み込むこととあわせて, Preference Elicitation

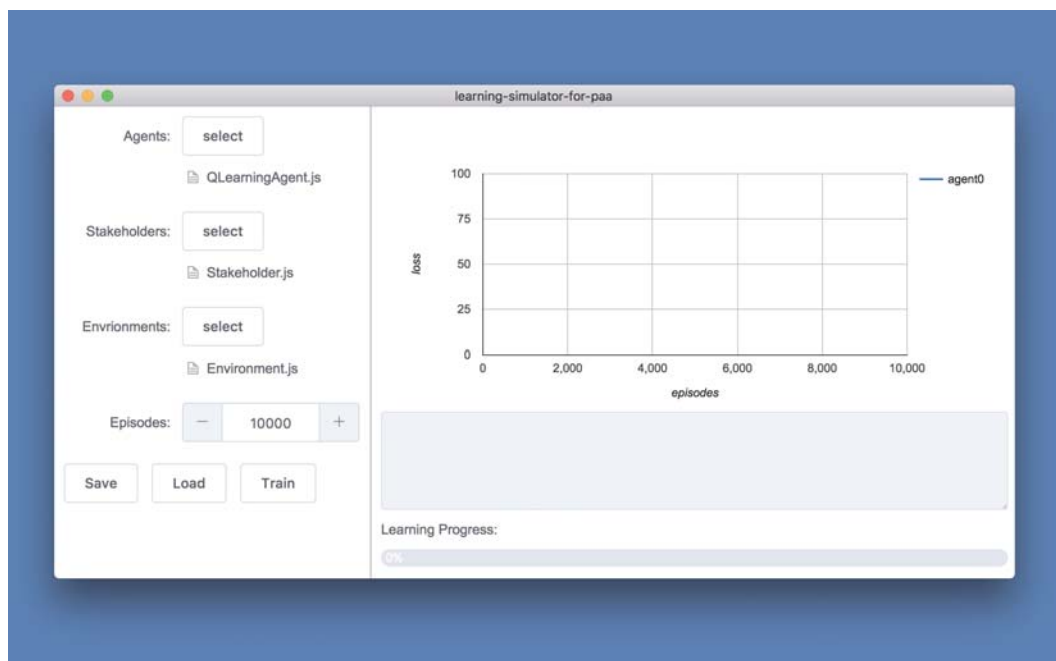


図 2: シミュレータの外観

を適切なタイミングで行い、ユーザから Preference に関する知識を抽出することで、他のユーザとの競合をより望ましい形で解決できる可能性について検討した。Preference Elicitation を行う際には、エージェントはユーザに対して直接 Preference を尋ねることが増加することはユーザの負担になると同時に、プライバシー情報を含む Preference が意図せず開示されてしまう場合への対処方法の検討について述べた。本論文では、ユーザの好ましいタイミングで Preference について尋ねるための行動タイミングを、QS-Learning を用いて事前にエージェントに学習させることで、ユーザの負担を考慮した Preference Elicitation を行う機構の設計について述べた。

参考文献

- [Ajmeri 17] Ajmeri, N., Murukannaiah, P. K., Guo, H., and Singh, M. P.: Arnor: Modeling Social Intelligence via Norms to Engineer Privacy-Aware Personal Agents, in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS '17*, pp. 230–238 (2017)
- [Ariel Rosenfeld 17] Ariel Rosenfeld, S. K., Matthew E. Taylor: Leveraging Human Knowledge in Tabular Reinforcement Learning: A Study of Human Subjects, in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pp. 3823–3830 (2017)
- [Baarslag 15] Baarslag, T. and Gerding, E. H.: Optimal Incremental Preference Elicitation During Negotiation, in *Proceedings of the 24th International Conference on Artificial Intelligence, IJCAI'15*, pp. 3–9 (2015)
- [Dinakar 15] Dinakar, K., Picard, R., and Lieberman, H.: Common Sense Reasoning for Detection, Prevention, and Mitigation of Cyberbullying, in *Proceedings of the 24th International Conference on Artificial Intelligence, IJCAI'15*, pp. 4168–4172 (2015)
- [Oishi 16] 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'16)*, pp. 664–667 (2016)
- [Oishi 17a] Oishi, S. and Fukuta, N.: MstdnDeck: An Agent-based Protection of Cyber-bullying on Distributedly Managed Linked Microbloggings, in *Proceedings of the International Conference on Web Intelligence, WI '17*, pp. 1195–1198 (2017)
- [Oishi 17b] Oishi, S. and Fukuta, N.: Toward a Negotiation-based Cooperation Mechanism for User Assistance Agents and Humans, in *the 10th International Workshop on Agent-based Complex Automated Negotiations (ACAN'17)* (2017)
- [Sutton 98] Sutton, R. S. and Barto, A. G.: *Introduction to Reinforcement Learning*, MIT Press, Cambridge, MA, USA, 1st edition (1998)
- [大石 16a] 大石 翔, 福田 直樹: 個人支援エージェントの強化学習に基づく動的処理負荷調整機構の試作, 2016 年度電子情報通信学会総合大会 (2016)
- [大石 16b] 大石 翔, 福田 直樹: 個人支援エージェントの能力オントロジーに基づく協調的タスク遂行機構の試作, 2016 年度人工知能学会全国大会 (2016)