Towards Multi-Agent Interactive Reinforcement Learning for Opportunistic Software Composition in Ambient Environments

Doctoral Consortium

Kevin Delcourt
IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3
Toulouse, France
Kevin.Delcourt@irit.fr

ABSTRACT

In order to manage the ever-growing number of devices present in modern and future ambient environments, as well as their dynamics and openness, we aim to propose a distributed multi-agent system that learns, in interaction with a human user, what would be their preferred applications given the services available.

The goal of this Ph.D. thesis is to focus on the interaction between a reinforcement learning system and the human user, to improve the system's learning capabilities as well as the user's ease with the system, and ultimately build a working prototype, usable by end-users.

KEYWORDS

Ambient Intelligence; Machine Learning; Human-AI Interaction; Multi-Agent System; Human-in-the-loop; Emergence

ACM Reference Format:

Kevin Delcourt. 2022. Towards Multi-Agent Interactive Reinforcement Learning for Opportunistic Software Composition in Ambient Environments: Doctoral Consortium. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 2 pages.

1 CONTEXT

1.1 Ambient Intelligence

Ambient Intelligence [11] aims to provide a human user with a personalized cyber-physical environment and associated software, that are tailored to their specific situation and needs. Openness, dynamism, unpredictability of ambient systems are all difficulties to overcome. Furthermore, the ever-changing needs of each user are another source of complexity.

As a result, it is extremely onerous for a designer to anticipate every possible state of the ambient environment and the user needs, then design user-adapted applications in the traditional top-down way.

1.2 Opportunistic Software Composition

My Ph. D. thesis is part of a research project that addresses those challenges by designing and developing a solution that learns from its user, to bring out tailored applications composed in a bottom-up manner from software services that are present at the time in their ubiquitous environment.

This opportunistic composition model [16] is based on continuous reinforcement learning [13] from the user's feedback. Learning and decision are distributed in a multi-agent system (MAS) [15] to meet requirements such as scalability in number of devices as well as distribution in the cyber-physical space.

In this model, every agent manages a basic service of the user's environment, and collaborates to decide on the composition of its service with another. Thus, due to agents' distributed interactions, applications emerge from the environment in the form of a set of service compositions. Agents' decision is built from their local view of the environment, and from the feedback provided by the user during their everyday routine, *i.e.* their preferred applications for any given environment.

1.3 Project Status

This project is at the crossroads of several research areas: ambient intelligence, machine learning, human-computer interaction, and MAS. We considered several research questions in past works:

- A prior work examined the problem of building composite applications with the user-in-the-loop [16]. However, the solution takes into account the human user in a limited way, over-stressing them.
- As preliminary work, I published a demonstration of our solution's prototype in a controlled ambient environment [5], also available as a video[1].
- In order to present composite applications to an end-user in an understandable way, we developed a domain-specific language, and a tool allowing for a user to have access to several views of the application [14].

2 PH.D. THESIS SCOPE

2.1 Research Questions

Human users have complex and dynamic goals that depend on their preferences and the services available in the ambient cyber-physical environment. Moreover, they are prone to make mistakes when interacting with learning systems. And so, since our past works only sporadically treated human factors, we focus on the interactions between the human user and the learning system and how best to learn from them:

- How to best build and adapt the life-long model of the human user?

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online. © 2022 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

[1] https://www.irit.fr/%7eSylvie.Trouilhet/demo/outletSeeking.mp4

- How to decentralize this user-oriented intelligence in a MAS, in order to allow, among other things, the learning system to scale up?
- How can autonomous learning agents collaborate to build useful collective knowledge?
- How to do so and not burden the user with excessive solicitation? [3]
- Is knowledge about one user transferable to another, and how? Indeed, for now, we only studied the case of a single user in an ambient environment.
- Finally, how to assess any proposed solution and validate that the above issues are taken into account?

Concerning the fourth research question, I have explored in a prior work [6] the possibility of leveraging software product lines [10] in order to decrease the amount of feedback required from the user.

Working on addressing these issues led me to study the field of Interactive Machine Learning.

2.2 Interactive Machine Learning

Interactive machine learning (IML) systems regroup all systems that allow end-users, devoid of ML knowledge, to produce an ML-based solution satisfying their needs [1]. Thus, recommendation systems [12] and end-user programming systems of intelligent software [7] are related to this domain. IML being at the intersection of human-computer interaction (HCI) and machine learning (ML), most studies in the field tend to lean towards one or the other.
On the HCI side, work on IML tends to focus on the impact of human-AI interaction on the user experience, improving global performance by focusing on usability issues:

- Amershi *et al.* extracted 18 guidelines from the literature that could be used by designers to produce more user-friendly intelligent systems [2].
- In [9], Honeycutt *et al.* study the impact of the interaction with an automated learning system on the user's trust. It is an important factor to consider: if a user does not trust a system enough, its usage will decline. On the contrary a higher trust than necessary means the user won't be watchful for system poor decisions with potentially dire consequences.

On the ML side, studies focus on finding and tuning ML algorithms to best learn from human interaction, therefore increasing global performance:

- In [8], Holzinger *et al.* showed that an ant colony optimization system produced better results when receiving human guidance.
- Christiano *et al.* present an interactive reinforcement learning system leveraging deep neural networks to minimize the interaction needed to learn a sequence of actions from human feedback [4].

2.3 Research Proposal

The main objective of my thesis is to propose a non-intrusive solution taking into account human factors, based on the insights gained from IML and related domains. And later to produce a software prototype out of this model. In this contribution, the interactions between the learning system and the user would benefit the system which gets feedback on its predictions, as well as the user who gets a pertinent service offer.

Finally, we would then need to study and validate this solution in an open dynamic environment, in relation to end-users, in order to affirm the quality of our proposal regarding the user experience of the system. That is a transverse issue as every contribution in this project would benefit from being thoroughly validated.

3 CONCLUSION

We presented in this extended abstract the research domain and the specific questions to be addressed during my Ph.D., namely how to learn from the human user and distribute knowledge through a MAS, to bring out in an ambient environment the applications that most suit their need. To the extent of our knowledge, this is a relatively emergent field of study, and we hope to get valuable feedback and insights on our past, present, and future works during the doctoral consortium.

REFERENCES

[1] S. Amershi, M. Cakmak, W. Knox, and T. Kulesza. 2014. Power to the people: The role of humans in interactive machine learning. *Ai Magazine* 35, 4 (2014), 105–120.
[2] S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, and K. Inkpen. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.
[3] C. Bach and D. Scapin. 2003. Adaptation of ergonomic criteria to human-virtual environments interactions. In *Proceedings of Interact*, Vol. 3. Citeseer, 880–883.
[4] P. Christiano, J. Leike, T. Brown, M. Martic, S. Legg, and D. Amodei. 2017. Deep reinforcement learning from human preferences. *arXiv preprint arXiv:1706.03741* (2017).
[5] K. Delcourt, F. Adreit, J.-P. Arcangeli, K. Hacid, S. Trouilhet, and W. Younes. 2021. Automatic and Intelligent Composition of Pervasive Applications - Demonstration. In *19th IEEE International Conference on Pervasive Computing and Communications (PerCom 2021)*. Kassel (virtual), Germany. https://hal.archives-ouvertes.fr/hal-03200989
[6] K. Delcourt, F. Adreit, J.-P. Arcangeli, and S. Trouilhet. 2021. Interactive Reinforcement Learning for Software Composition via Software Product Lines - Approach and Research Questions. In *Unpublished paper*.
[7] J. Fails and D. Olsen Jr. 2003. Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces*. 39–45.
[8] A. Holzinger, M. Plass, M. Kickmeier-Rust, K. Holzinger, G. Crişan, C. Pintea, and V. Palade. 2019. Interactive machine learning: experimental evidence for the human in the algorithmic loop. *Applied Intelligence* 49, 7 (2019), 2401–2414.
[9] D. Honeycutt, M. Nourani, and E. Ragan. 2020. Soliciting human-in-the-loop user feedback for interactive machine learning reduces user trust and impressions of model accuracy. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 8. 63–72.
[10] A. Metzger and K. Pohl. 2014. Software product line engineering and variability management: achievements and challenges. In *Future of Software Engineering Proceedings*. 70–84.
[11] F. Sadri. 2011. Ambient intelligence: A survey. *ACM Computing Surveys (CSUR)* 43, 4 (2011), 1–66.
[12] T. Schnabel, S. Amershi, P. Bennett, P. Bailey, and T. Joachims. 2020. The Impact of More Transparent Interfaces on Behavior in Personalized Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20)*. Association for Computing Machinery, New York, NY, USA, 991–1000. https://doi.org/10.1145/3397271.3401117
[13] R. Sutton and A. Barto. 2018. *Reinforcement learning: An introduction*. MIT press.
[14] S. Trouilhet, J.-P. Arcangeli, Bruel J.-M., and M. Koussaifi. 2021. Model-Driven Engineering for End-Users in the Loop in Smart Ambient Systems. *Journal of Universal Computer Science* 27, 7 (July 2021).
[15] M. Wooldridge. 2009. *An Introduction to MultiAgent Systems* (2nd ed.). Wiley Publishing.
[16] W. Younes, S. Trouilhet, F. Adreit, and J.-P. Arcangeli. 2020. Agent-mediated application emergence through reinforcement learning from user feedback. In *29th IEEE Int. Conf. on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*. IEEE Press. https://hal.archives-ouvertes.fr/hal-02895011.