Robots as Mediators to Resolve Multi-User Preference Conflicts - Extended Abstract

Aniol Civit[1,*,†], Rebecca Stower[2,†], Iolanda Leite[2], Antonio Andriella[3] and Guillem Alenyà[1]

[1]*Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain*
[2]*KTH Royal Institute of Technology, Stockholm, Sweden*
[3]*Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Spain*

Abstract

In real-life scenarios, robots will have to make decisions that involve multiple users. The current literature does not consider scenarios where a robot interacts with users who have conflicting preferences. To address this issue, this paper proposes using the robot as a mediator. Different possible conflict resolution actions for the robot are presented, as well as the challenges and open questions arising from this proposal.

Keywords

Robot Personalisation, Multi-User Preferences, Conflict Resolution, Social Robotics

1. Introduction

A substantial body of work exists on how robots can learn human preferences and personalisation in Human-Robot Interaction (HRI) [1]. For example, preferences can be learned from observations of user behaviour in different tasks [2], from experts [3], from pairwise comparisons [4], and other approaches [5, 6]. Similarly, research on robots in groups has increased significantly in recent years [7, 8]. Different behaviours of the robot have been investigated (e.g., gaze or verbal feedback), as well as group-level outcomes such as task performance, social cohesion, turn-taking, inclusion, rational thinking, or trust [9, 10, 11, 12, 13].

What happens when multiple users are interacting with a robot with possibly competing preferences (Multi-User Multi-Objective) has yet to be explored [14]. If we consider failing to conform to a user's preference as a type of failure, then presumably this can lead to negative outcomes such as decreased trust or engagement in the interaction [15]. One strategy for potentially resolving failures is by providing explanations for the failure [16]. However, work on robot failures so far has not considered cases where robots are aware of users' preferences but are unable to conform to them, nor how the robot should cope with these scenarios.

Consequently, it is not yet clear how personalisation-based failures might be perceived by users (e.g., as a task failure, since the robot fails to achieve a goal, or as a social norm violation if the failure to comply is seen as a refusal). This could also lead to imbalances in the group dynamics if different users have different expectations and attitudes towards the robot according to whether their preference was followed or not. It could also affect interpersonal dynamics between human group members depending on whose preferences the robot follows.

Resolving differences between preferences of human users can be considered a form of conflict resolution [17]. Among HRI work which targets conflict resolution, most focus is on direct human-robot

ALTRUIST, BAILAR, SCRITA, WARN 2024: Workshop on sociAL roboTs for peRsonalized, continUous and adaptIve aSsisTance, Workshop on Behavior Adaptation and Learning for Assistive Robotics, Workshop on Trust, Acceptance and Social Cues in Human-Robot Interaction, and Workshop on Weighing the benefits of Autonomous Robot persoNalisation. August 26, 2024, Pasadena, USA

*Corresponding author.
†These authors contributed equally.
acivit@iri.upc.edu (A. Civit); stower@kth.se (R. Stower); iolanda@kth.se (I. Leite); antonio.andriella@iiia.csic.es (A. Andriella); galenya@iri.upc.edu (G. Alenyà)

conflicts in dyadic (human-robot) interactions. For example, [18, 19, 20] explore different conflict-resolution strategies when human and robot goals differ. Different outcomes were affected, such as acceptance of the robot, compliance with the request, and trust. In group settings, some work has explored conflict resolution, e.g., [21] compared a robot which intervened following a task-based or personal attack to one that did nothing. Some work has also been done looking at robots as mediators in groups of children [22, 23].

In sum, the extant literature on human-robot conflict resolution mainly focuses on dyadic conflict resolution between a human and a robot. What happens when two human users have conflicting preferences, or how the robot should resolve such conflicts, has yet to be explored. In this work, we consider the different roles of a *robot as mediator* when two human users have conflicting preferences for which action a robot should take. We contribute to the existing body of literature by *a.* identifying factors which can affect conflict resolution in human-human-robot social dynamics, and *b.* proposing initial actions that a robot mediator could use to proactively mitigate preference conflicts.

2. Factors affecting conflict resolution

Based on human-human conflict resolution theory [24] and current approaches to managing human-robot disagreements, we have identified the following factors that could potentially influence the resolution process:

- Role of Users (e.g., expert, user)
- Strength of the preferences
- Consequences of following one preference over another (e.g., psychological vs physical well-being)
- Robot appearance and behaviour (e.g., communication strategy)
- Length and/or number of interactions (short versus long-term)

Depending on these factors, different interaction outcomes could be affected, such as users' trust, acceptance, engagement or future willingness to interact with the robot.

3. Robot as a mediator

Here, we propose different decision-making actions (A) a mediator robot could employ to resolve Multi-User Multi-Objective conflicts:

(A.1) Do nothing / wait for external agreement between users

(A.2) Select a random preference

(A.3) Select the expert preference, if any

(A.4) Select an intermediate or alternative option, if any

(A.5) Weight user preferences and select the most significant one

(A.6) Weight user preferences and alternate them proportionally

(A.7) Use a Multi-Objective optimization algorithm

(A.8) Proactively discuss with one user to refine their preferences

(A.9) Proactively discuss with all users to reach an agreement

Depending on the factors identified above, these actions might have different effectiveness towards resolving preference conflicts. Some of them imply following rule-based decision-making, one uses optimization solver algorithms, and the latter proactively tries finding an agreement in the preferences (and learning them) to resolve conflicts.

Complementing the robot's decision-making in conflicting scenarios, we propose different levels of explanations (E) of those decisions to the users:

(E.1) *Act without an explanation*: The robot decides an action and executes it without explaining the reason for its selection.

(E.2) *Explain the robot's decision before/after executing the action*: The robot explains the reason for the decided action without involving the other users.

(E.3) *Explain there is a preference conflict with another user*: The robot explains there are differences in preferences from different users regarding the execution of the task.

(E.4) *Explain the reason for the other users' preference*: The robot needs to know the reason for the preferences of each user and when there are conflicts explain the reasons of the other users for those preferences.

A necessary additional property of the robot is the capability to actively learn the (potentially changing) users' preferences.

4. Challenges and open questions

Introducing robots as mediators to resolve Multi-User Multi-Objective Conflicts raises the following identified technical challenges (C) and open questions (Q):

(C.1) How can the robot continuously learn and adapt to users' preferences over time?

(C.2) How can the robot identify conflicts between user preferences?

(C.3) How can the robot find alternatives or intermediate solutions when preference conflicts arise?

(C.4) How can the robot be aware if accomplishing the preferences (or a solution) is within its capabilities?

(Q.1) How are trust and acceptance affected by different conflict resolution strategies (e.g., providing an alternative solution)?

(Q.2) How does using the robot as a mediator influence the user's engagement?

(Q.3) How are trust and acceptance affected when a robot does not conform to a user's preference (despite having the capacity to do so)?

(Q.4) How are interpersonal dynamics between users affected by differential robot preference adherence?

5. Example scenario

One of the many possible scenarios where conflicts between users may arise is when the robot interacts with an expert and a regular user for a specific task, and the robot's decisions directly affect the user.

For example, a robot is placed in the user's home and assists the user by providing them with food and drinks. The expert is the user's nutritionist and explains to the robot that the user should consume less sugar. Later, the user asks for a soda drink. In this situation, the robot receives two conflicting preferences, since the user should consume less sugar to improve their health, but they would like to drink a soda.

If the robot decides to follow strictly the expert's opinion (A.4), there exists a risk that the user loses trust or engagement with the robot and decides to fetch the soda by themselves and/or ignore the robot's suggestions in the future. However, if the robot strictly follows the user's preference (A.3), it will ignore the expert's preference for that action as well as leading to possibly detrimental health outcomes.

Interpreting the user and expert requests can give useful information for conflict resolution, in this case, the information provided by the expert concerns the user's health, and the user's input can be interpreted as wanting something to drink.

From this point, the robot can act in different ways, one of which is deciding on one of the proposed actions and acting without providing any explanation (E.1). Another solution is providing an intermediate or alternative solution such as offering water to the user (A.6). A solution involving conflict resolution could be explaining to the user that they should consume less sugar, which is the expert's preference, without involving the expert in the explanation (E.2), and additionally offer an alternative, such as offering water instead (A.6). Finally, another conflict resolution approach is explaining to the user that their nutritionist prefers that they consume less sugar and offering a cup of water (E.3 or E.4).

Additionally, the robot can give feedback to the expert, for example, commenting that the user accepted a cup of water after hearing its preference, but that they did not seem happy about it, offering the possibility to the expert to change their preference or reach a compromise (A.10), such as allowing the user to consume a limited amount of sodas during the week.

6. Conclusion and future work

In real-world scenarios, robots incorporating preference learning without expert knowledge may lead to sub-optimal or incorrect task performance. We enumerated different actions and levels of explanations to be used by a robot to mitigate conflicts in Multi-User competing preference scenarios.

Multi-Objective optimization approaches and using the robot as a mediator are promising tools for conflict resolution, but the lack of literature on that scenario raises the stated challenges and open questions, which provide a starting point for future research.

7. Acknowledgment

The S-FACTOR project from NordForsk partially funded this work (R.S and I.L). A.C has been supported by AGAUR-FI ajuts (2023 FI-1 00536) Joan Oró of the Generalitat of Catalonia and the European Social Plus Fund.

References

[1] M. Chetouani, Interactive robot learning: an overview, ECCAI Advanced Course on Artificial Intelligence (2021).

[2] B. Woodworth, F. Ferrari, T. E. Zosa, L. D. Riek, Preference learning in assistive robotics: Observational repeated inverse reinforcement learning, in: Machine learning for healthcare conference, 2018.

[3] A. Andriella, C. Torras, C. Abdelnour, G. Alenyà, Introducing caresser: A framework for in situ learning robot social assistance from expert knowledge and demonstrations, User Modeling and User-Adapted Interaction (2023).

[4] A. Bemporad, D. Piga, Global optimization based on active preference learning with radial basis functions, Machine Learning (2021).

[5] E. Senft, S. Lemaignan, P. E. Baxter, M. Bartlett, T. Belpaeme, Teaching robots social autonomy from in situ human guidance, Science Robotics (2019).

[6] A. Z. Ren, A. Dixit, A. Bodrova, S. Singh, S. Tu, N. Brown, P. Xu, L. Takayama, F. Xia, J. Varley, Z. Xu, D. Sadigh, A. Zeng, A. Majumdar, Robots that ask for help: Uncertainty alignment for large language model planners, in: Proceedings of The 7th Conference on Robot Learning, 2023.

[7] R. Oliveira, P. Arriaga, A. Paiva, Human-robot interaction in groups: Methodological and research practices, Multimodal Technologies and Interaction (2021).

[8] S. Sebo, B. Stoll, B. Scassellati, M. F. Jung, Robots in groups and teams: a literature review, Proceedings of the ACM on Human-Computer Interaction (2020).

[9] S. Gillet, M. T. Parreira, M. Vázquez, I. Leite, Learning gaze behaviors for balancing participation in group human-robot interactions, in: 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2022.

[10] S. Sebo, L. L. Dong, N. Chang, M. Lewkowicz, M. Schutzman, B. Scassellati, The influence of robot verbal support on human team members: Encouraging outgroup contributions and suppressing ingroup supportive behavior, Frontiers in Psychology (2020).

[11] C. Birmingham, Z. Hu, K. Mahajan, E. Reber, M. J. Matarić, Can i trust you? a user study of robot mediation of a support group, in: IEEE International Conference on Robotics and Automation (ICRA), 2020.

[12] S. Forgas, R. Huertas, A. Andriella, G. Alenyà, How do consumers' gender and rational thinking affect the acceptance of entertainment social robots?, International Journal of Social Robotics (2022).

[13] M. R. Fraune, S. Šabanović, T. Kanda, Human group presence, group characteristics, and group norms affect human-robot interaction in naturalistic settings, Frontiers in Robotics and AI (2019).

[14] S. Gillet, M. Vázquez, S. Andrist, I. Leite, S. Sebo, Interaction-shaping robotics: Robots that influence interactions between other agents, ACM Transactions on Human-Robot Interaction (2024).

[15] D. Ullman, B. F. Malle, Measuring gains and losses in human-robot trust: Evidence for differentiable components of trust, in: 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019.

[16] S. S. Sebo, P. Krishnamurthi, B. Scassellati, "i don't believe you": Investigating the effects of robot trust violation and repair, in: 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019.

[17] T. H. Weisswange, H. Javed, M. Dietrich, T. V. Pham, M. T. Parreira, M. Sack, N. Jamali, What could a social mediator robot do? lessons from real-world mediation scenarios, arXiv preprint arXiv:2306.17379 (2023).

[18] F. Babel, J. M. Kraus, M. Baumann, Development and testing of psychological conflict resolution strategies for assertive robots to resolve human-robot goal conflict, Frontiers in Robotics and AI (2021).

[19] F. Babel, P. Hock, J. Kraus, M. Baumann, Human-robot conflict resolution at an elevator-the effect of robot type, request politeness and modality, in: 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2022.

[20] F. Babel, P. Hock, J. Kraus, M. Baumann, It will not take long! longitudinal effects of robot conflict resolution strategies on compliance, acceptance and trust, in: 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2022.

[21] M. F. Jung, N. Martelaro, P. J. Hinds, Using robots to moderate team conflict: the case of repairing violations, in: Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction, 2015.

[22] S. Shen, P. Slovak, M. F. Jung, "stop. i see a conflict happening." a robot mediator for young children's interpersonal conflict resolution, in: Proceedings of the ACM/IEEE international conference on human-robot interaction, 2018.

[23] S. Gillet, W. van den Bos, I. Leite, A social robot mediator to foster collaboration and inclusion among children., in: Robotics: Science and Systems, 2020.

[24] J. Bercovitch, R. Jackson, Negotiation or mediation?: An exploration of factors affecting the choice of conflict management in international conflict, Negotiation Journal (2001).