Deriving norms from actions, values, and context

Tielman, Myrthe L.; Jonker, Catholijn M.; Van Riemsdijk, M. Birna

Publication date
2019

Document Version
Final published version

Published in
18th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2019

Citation (APA)
Tielman, M. L., Jonker, C. M., & Van Riemsdijk, M. B. (2019). Deriving norms from actions, values, and context. In *18th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2019* (Vol. 4, pp. 2223-2225). International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS).

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Deriving Norms from Actions, Values, and Context

Extended Abstract

Myrthe L. Tielman
Delft University of Technology
Netherlands
m.l.tielman@tudelft.nl

Catholijn M. Jonker
Delft University of Technology
Leiden University
Netherlands
c.m.jonker@tudelft.nl

M. Birna van Riemsdijk
Delft University of Technology
University of Twente
Netherlands
m.b.vanriemsdijk@tudelft.nl

ABSTRACT

Personal technology such as electronic partners (e-partners) play an increasing role in our daily lives, and can make an important difference by supporting us in various ways. However, when they offer this support, it is important that they do so with an understanding of our choices and what is important to us. To allow an e-partner to flexibly do this, we propose a formal framework to automatically derive norms which describe how to perform a certain behavior. These norms are directly derived from the user's actions, values and the context they are in. In this way, the e-partner can take into account the user's values and offer more flexible personalized support.

KEYWORDS

Formal methods ; Values; Normative systems

ACM Reference Format:
Myrthe L. Tielman, Catholijn M. Jonker, and M. Birna van Riemsdijk. 2019. Deriving Norms from Actions, Values, and Context. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), Montreal, Canada, May 2019,* IFAAMAS, 3 pages.

1 INTRODUCTION

Digital technology plays an increasing role in our daily lives, and nowadays, many systems can be seen as personal technology, or electronic partners (e-partners). Such technology supports us in daily life activities, from simply reminding us to take our medicine or that we need to get up from our chair and take a walk [6, 15], to more complex systems for monitoring your health [7, 14] or navigation support for visually impaired people [2, 5].

As e-partners assist us on a personal level, it is important that they understand what we want, and what is important to us [1]. For instance how we wish to be reminded of taking our medicine, or what route we wish to take when we are in a hurry. Many current systems address this in the design of the technology, for instance through holding interviews with stakeholders, and incorporating the resulting design requirements [11]. However, the results are often still quite rigid in their personalization, as once designed, the system is both the same for every user, and stays the same over time. This can be a problem as individual differences exist in what we find important. Moreover, our preferences and situations also often change over time.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 2019, Montreal, Canada.

In order to be able to offer more flexible support, the system itself should not just be designed with our wishes in mind, but also able to reason about this [13]. This means that an e-partner should have an explicit internal representation of how the user wishes to perform certain behavior in a certain situation. Such an explicit internal representation allows updates, making the system more flexible [9, 12].

To represent how a user wishes to perform a behavior given their individual preferences and situation, *formal norms* can be used. Norms are at their core rules for behavior, they express what behavior is preferred in different contexts [8]. They can operate both at a societal and personal level, and within computer-science norms have, therefore, been formalized and used extensively to model the behavior of both individual agents and agent societies [3, 4]. As they represent rules from behavior, are based on what we find important as people, and have been extensively modelled formally, norms are particularly suited to be used by an e-partner to represent how a user wishes to perform their actions.

Adding formal norms to a system allows an an e-partner to be able to update these norms. There are several possible ways to achieve such updates. At their core, norms are based on values which describe what we find important, but add more detail about what this means for choices in a given situation [11]. They are, therefore, most likely to change when something changes in how values relate to these choices and situations. It makes sense, therefore, to represent this information explicitly, and directly derive norms from this. Then, if something changes to, for instance, the values, the set of norms can be automatically updated. An added advantage of such a representation is that if a user would ever wish to manually update a norm, this would be possible as well. The e-partner could then check if the new, manual norm conflicts with any of the existent norms and update its knowledge accordingly. If we would only change norms through manual input, however, this could more easily result incomplete information, and might be difficult for the user.

This work outlines how actions, values and situational context can be modelled such that norms can be automatically derived, and illustrates how such derivations can work.

2 ACTIONS, VALUES, AND CONTEXT

In this work we consider two types of norms which describe how to perform (or not perform) a certain action. Firstly, norms which describe what actions to perform as a part of another action, and secondly how to do an action more concretely. For instance, a norm can be to lock your bike as a part of stalling your bike, or a norm can be to take short route to the doctor as a more concrete way of going

to the doctor. In order to derive these norms we, therefore, need information on how an action can be performed in the first place. Firstly, information about what constitutes the different parts of an activity and whether they are necessary for completing the action. And secondly, information about the alternative more concrete ways of performing an activity, where we assume that only one alternative for performing an action can be performed at one time. This information can be structured in a small (2-level) tree, where the root is the activity, and the leafs the different parts of, or more concrete ways of performing it.

Such action trees provide the e-partner with information about the actions. However, to know which of these actions should be performed, it also requires information about the user's values. And, more specifically, how these values relate to the different leaf actions in the tree. For instance, if one can go to the doctor's via route A or route B, the e-partner needs to know if route A is safer (i.e. better for the value of safety), and route B is quicker (i.e. better for the value of efficiency), in order to decide. For this reason, we explicitly link values to leaf actions in the trees, and use a number to denote the strength of this relationship. This can be both positive and negative, as actions might also demote certain values.

Finally, the situation in which these actions are performed is also important. Route A might be safer normally due to less traffic, but late at night this might change, as route B also has less traffic. For this reason, we introduce the concept of context. Context is a situational factor, which influences the relationship between an action and a value. For instance, route A might normally promote safety by 2, but the context *late at night* could influence this relationship with -5, meaning that late at night, choosing route A would demote safety by 3.

With this information about relations between actions, how they influence values, and the role of context, we can derive norms on how to perform a certain action.

3 DERIVING NORMS

We propose several requirements for a system which derives norms from information on actions, values and context. Firstly, that system should be able to derive the norms without any further input from the user. Secondly, the resulting norm set should enable maximal support of the user's values. Thirdly, the norms set should not contain any conflicts which cannot be easily solved by the e-partner. And finally, the norm set should not contain any norms which are redundant, where redundancy indicates that following a norm would not actually change the behavior of the user given all the other norms in the norm set. For instance, if a user has an obligation to take route A, an obligation to take route A when it is raining would be redundant, as they are already doing that.

Taking these considerations into account, norms can be derived in the following way. For both part-of and concretisation norms, the first norms to be derived are for when no context factors are present or known. Then follow the norms for one context factor at a time, then two, etc. At each step, the *value sum* for the action and context is calculated. This is the sum of the value numbers related to that action and the value numbers of the context. Basically, how much the values are promoted or demoted in that context by performing that action. Then, for part of actions, the norm will be an obligation if the value sum is positive, permission if it is 0 and a prohibition if it is negative. For concretisation actions, the norm is an obligation for the action with the highest score for that context, and a prohibition for the rest. If there is a tie for the highest score, these actions become permissions and the others prohibitions, because the concretisation actions represent an exclusive choice. Then, for all norms, they are only added to the final norm set if they are not redundant, so if adding them has an effect on what the user would actually do.

In the final norm set, conflicts only occur between norms with different numbers of context factors. To solve these, the e-partner can follow *lex specialis*, which means that the most specific norm takes precedence. In this case, this is the norm with the most context factors. So for instance, if there is an obligation to take route A when it is raining and a prohibition to take route A when it is raining and you are late, the prohibition is followed. This method of solving conflict also shows why permissions are necessary, in some contexts an earlier, less specific obligations or prohibitions might need to be overruled by a permission.

4 DISCUSSION & CONCLUSION

The framework we describe in this paper operates under a number of assumptions. Firstly, relating to the *value sum*. We assume that the relationship between an action and value can be denoted by a number, and assumes ratio measurability and commensurability. Related to this, we assume mutual independence of context factors, as we sum the numbers when more than one context factor is present. Neither of these assumptions are trivial, but they are often used in order to be able to use values. Another assumption is that information on the actions, values and context is known to the system. This means that the system requires a way to receive this information, for instance through asking questions to the user about their habits.

Several directions for future work can be established. Firstly, to develop the ideas presented in this paper into a full formal framework, of which preliminary work has been done in [10]. Some further considerations for future work are to take into account the priorities between values; to develop the concept of context further, for instance also including social aspects explicitly; and to take into account societal values as well as user values. Finally, the goal of this framework is to derive norms which tell an e-partner what actions the user would wish to pursue. To actually support the user in performing these actions is, therefore, also a logical next step.

In this work we present the basic outlines of a framework for an e-partner to automatically derive norms from a user's actions, values and context. The norm set which results from this derivation represents how a user would like to perform an action given their values, and the context they are in. Using this norm set, an e-partner could ensure promoting the users values when offering support.

ACKNOWLEDGMENTS

This work is part of the research programme CoreSAEP, with project number 639.022.416, which is financed by the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

[1] 2017. *Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2.* The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. https://ethicsinaction.ieee.org/

[2] Jan Balata, Zdenek Mikovec, and Pavel Slavik. 2018. Landmark-enhanced route itineraries for navigation of blind pedestrians in urban environment. *Journal on Multimodal User Interfaces* 12, 3 (01 Sep 2018), 181–198. https://doi.org/10.1007/s12193-018-0263-5

[3] Tina Balke, Celia da Costa Pereira, Frank Dignum, Emiliano Lorini, Antonino Rotolo, Wamberto Vasconcelos, and Serena Villata. 2013. *Normative Multi-Agent Systems.* Schloss Dagstuhl.

[4] N. Criado, E. Argente, P. Noriega, and V. Botti. 2013. Human-inspired model for norm compliance decision making. *Information Sciences* 245 (2013), 218–239.

[5] L. González-Delgado, L. Serpa-Andrade, K. Calle-Urgiléz, A. Guzhñay-Lucero, V. Robles-Bykbaev, and M. Mena-Salcedo. 2016. A low-cost wearable support system for visually disabled people. In *2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC).* 1–5.

[6] Eleonora Milić, Dragan Janković, and Aleksandar Milenković. 2018. Health Care Domain Mobile Reminder for Taking Prescribed Medications. In *ICT Innovations 2016*, Georgi Stojanov and Andrea Kulakov (Eds.). Springer International Publishing, Cham, 173–181.

[7] Temiloluwa Prioleau, Elliot Moore II, and Maysam Ghovanloo. 2017. Unobtrusive and Wearable Systems for Automatic Dietary Monitoring. *IEEE Transactions on Biomedical Engineering* 64 (2017), 2075–2089.

[8] J.S. Santos, J.O. Zahn, E.A. Silvestre, V.T. Silva, and W.W. Vasconcelos. 2017. Detection and resolution of normative conflicts in multi-agent systems: a literature survey. *Journal of Autonomous Agent Multi-Agent Systems* 31 (2017), 1236–1282.

[9] John Thangarajah, Lin Padgham, and Michael Winikoff. 2003. Detecting & Exploiting Positive Goal Interaction in Intelligent Agents. In *Autonomous Agents & Multiagent Systems, AAMAS.*

[10] M.L. Tielman, C.M. Jonker, and M.B. van Riemsdijk. 2018. What should I do? Deriving norms from actions, values and context. In *International Workshop Modelling and Reasoning in Context (MRC), Held at FAIM.* Under revision at the AAMAS/IJCAI Workshop on Modeling and Reasoning in Context.

[11] Ibo van de Poel. 2013. *Translating Values into Design Requirements.* Springer, Chapter Philosophy and Engineering: Reflections on Practice, Principles and Process.

[12] M. Birna van Riemsdijk, Mehdi Dastani, Frank Dignum, and John-Jules Meyer. 2004. Dynamics of Declarative Goals in Agent Programming. In *International Workshop on Declarative Agent Languages and Technologies.*

[13] M. Birna van Riemsdijk, Catholijn M. Jonker, and Victor Lesser. 2015. Creating Socially Adaptive Electronic Partners. In *International Conference on Autonomous Agents and Multiagent Systems.*

[14] Wenxin Wang, Céline L. van Lint, Willem-Paul Brinkman, Ton J. M. Rövekamp, Sandra van Dijk, Paul J. M. van der Boog, and Mark A. Neerincx. 2017. Renal transplant patient acceptance of a self-management support system. *BMC Medical Informatics and Decision Making* 17, 1 (08 May 2017), 58. https://doi.org/10.1186/s12911-017-0456-y

[15] Yunlong Wang, Lingdan Wu, Jan-Philipp Lange, Ahmed Fadhil, and Harald Reiterer. 2018. Persuasive technology in reducing prolonged sedentary behavior at work: A systematic review. *Smart Health* 7-8 (2018), 19 – 30. https://doi.org/10.1016/j.smhl.2018.05.002