Capturing Contextual Morality: Applying Game Theory on Smartphones

Niels van Berkel
n.vanberkel@ucl.ac.uk
University College London
London, United Kingdom

Simo Hosio
simo.hosio@oulu.fi
University of Oulu
Oulu, Finland

Benjamin Tag
benjamin.tag@unimelb.edu.au
The University of Melbourne
Melbourne, Australia

Jorge Goncalves
jorge.goncalves@unimelb.edu.au
The University of Melbourne
Melbourne, Australia

ABSTRACT

In order to build more fair Artificial Intelligence applications, a thorough understanding of human morality is required. Given the variable nature of human moral values, AI algorithms will have to adjust their behaviour based on the moral values of its users in order to align with end user expectations. Quantifying human moral values is, however, a challenging task which cannot easily be completed using *e.g.* surveys. In order to address this problem, we propose the use of game theory in longitudinal mobile sensing deployments. Game theory has long been used in disciplines such as Economics to quantify human preferences by asking participants to choose between a set of hypothetical options and outcomes. The behaviour observed in these games, combined with the use of mobile sensors, enables researchers to obtain unique insights into the effect of context on participant convictions.

KEYWORDS

Mobile sensing; context; moral; ethics; artificial intelligence; game theory.

ACM Reference Format:
Niels van Berkel, Simo Hosio, Benjamin Tag, and Jorge Goncalves. 2019. Capturing Contextual Morality: Applying Game Theory on Smartphones. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct), September 9–13, 2019, London, United Kingdom.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3341162.3344846

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6869-8/19/09…$15.00
https://doi.org/10.1145/3341162.3344846

1 INTRODUCTION

The increasingly widespread deployment of Artificial Intelligence (AI) in our daily lives has raised concerns regarding the fairness, accountability, and transparency of these high-stake algorithms. Identified negative consequences of automated decision making (*e.g.*, racism, sexism [24]) have uncovered a gap between AI algorithms and the surrounding society's moral values. Indeed, initial work within HCI and other disciplines have started to direct focus on improving the fairness as well as transparency of these systems, and an interesting consideration of these algorithms is their connection, or lack thereof, to the actual moral values and beliefs which drive human decision making [8]. These values are contextual: previous work clearly indicates differences in moral values and decisions between people, cultures, and contexts [9]. Therefore, it is critical that future AI systems are trained on a contextually rich dataset and/or tailored to a specific context [12]. This knowledge, which we call *contextual morality*, is critical in ensuring that AI systems meet set design goals as well as the requirements of people in an increasingly complex interplay of technology and society.

Obtaining insightful measurements on the (contextual) morality of people is, however, not straightforward. Previous work investigating human morality has focused on the completion of surveys in the lab or online, often using stimuli to trigger a specific context [11, 14]. In this work, we explore the potential of game theory in collecting longitudinal insights about human morality. Game theory models interactive scenarios in which participants are offered a choice between two (or more) options and their respective reward. An essential element of game theory scenarios is the fact that the reward (or punishment) of the game is dependent on

Table 1: Outcomes for Player A and Player B in the Prisoner's Dilemma

	Player B cooperates (stays silent)	Player B defects (betrays Player A)
Player A cooperates (stays silent)	Player A serves 1 year in prison. Player B serves 1 year in prison.	Player A serves 3 years in prison. Player B goes free.
Player A defects (betrays Player B)	Player A goes free. Player B serves 3 years in prison.	Player A serves 2 years in prison. Player B serves 2 years in prison.

the strategy of both players. As such, combining the application of smartphone-based game theory with mobile sensing allows researchers to obtain rich insights into the effects of context on the participant's strategies and convictions.

In this work we explore the possibilities of employing game theory to collect longitudinal insights into contextual morality. To assist in the quantification of morality, we follow an Utilitarian perspective (as opposed to *e.g.* a Deontological perspective). Utilitarianism states that ethical decision making maximises value (*i.e.*, 'utility' – sometimes labelled as 'happiness') during decision making [20]. Following this perspective, a contextually moral decision can thus be defined as the decision which maximises utility in a given context. Subsequently, the study of contextual morality aims to systematically explore the effect of context on our moral values. As the use of game theory forces participants to make a choice (*i.e.* to 'pick a side'), appropriate use of game theory can identify how context affects participant decision making while maintaining identical game scenarios. Furthermore, we also identify challenges that researchers must overcome when running studies applying game theory. Our long-term goal is to apply game theory to identify user's expectations of the behaviour of AI applications in various day-to-day scenarios. Following the discussion of opportunities and challenges facing the deployment of in-the-wild game theory studies, we also outline opportunities for future work in this area.

2 RELATED WORK

Contextual morality

Research in the field of Psychology has revealed considerable differences in human morality following changes in their context. For example, Leavitt et al. asked participants to assume a professional identity (*e.g.*, manager versus engineer) and observed substantial influences on the participants' moral judgement [17]. Similarly, the (social) presence of others has been extensively studied and shown to affect our moral preferences [9].

In an ongoing attempt to better understand the expected behaviour of AI-systems, researchers have explored human morality in decision making. An example of this is provided by Awad et al., who study the moral implications of decisions made by autonomous vehicles [1]. In an online crowdsourcing study, participants were asked to indicate their preference for the car's collision avoidance, and subsequent potential casualty of the passengers, as based on the characteristics of pedestrians in front of the car (*e.g.*, elderly, children). Results from these human evaluations can offer insights and guidance for the implementation of both algorithms and policy. Although the authors identified differences in moral decision-making between cultures (*e.g.*, Southern cultures showed a strong preference to save younger people over older people, whereas this difference is much less pronounced in Eastern cultures) [1], no additional elements of context were considered in their analysis.

Game theory

Game theory models the strategic decision making between two or more players by providing a pre-set number of options and outcomes between which players are asked to make a choice. Given the ability of game theory to quantify human preferences while simultaneously offering versatility to adjust game scenarios, it has become a popular research instrument in Economics [7]. Arguably the most famous game theory example is the 'Prisoner's Dilemma'. Shown in Table 1, the Prisoner's Dilemma highlights the fact that players can act either as rivals or partners, and that both players operate under uncertainty (*i.e.*, what will the other player choose). An important concept in game theory is the Nash Equilibrium – a situation in which all of a game's players pursue the best possible strategy *given the strategies of all other players*. A commonly used interpretation of this is the 'no regret' interpretation; "*there is no player who, after observing the opponent's choice, regrets his own choice*" [7].

Game theory in HCI. Smartphones are highly suitable devices to be used for data collection activities due to the variety of onboard sensors as well as their interactive capabilities. With smartphones, different types of *auctions* [19] have been popular to collect data. For instance, in a 2-week long field trial, Hosio et al. leveraged a reverse second price auction to collect contextual data using mobile sensing as well as subjective insights on the monetary value of remaining battery life in a real life context [13].

3 OPPORTUNITIES

The related work typically relies on the use of surveys to obtain insights in human morality [9, 17], and has presented a number of validated scales to do so. A recent example is Black et al.'s Moral Identity Questionnaire [6] – a 20-item questionnaire which assesses both the salience of moral integrity and moral self. Although this approach can increase our understanding of human behaviour, its implications for *e.g.* ethical algorithmic behaviour are not directly obvious. Given the complexity of our motivations and convictions in different situations, richer and more eloquent ways of capturing our moral choices are increasingly being called for, in particular those that reach beyond simple articulations of words or Likert-scales [18]. As such, recent work by *inter alia* Awad et al. [1] has offered participants concrete alternative outcomes – allowing participants to identify their preferred outcome. The use of versatile mobile applications allows researchers to obtain both types of information (*i.e.*, short surveys and more concrete application-based questionnaires), potentially allowing for a correlation of these different measures. We outline some of the opportunities of employing game theory in longitudinal mobile studies.

Quantify convictions

Game theory has been extensively applied to study human behaviour under situations of incomplete information. This is a critical distinction with situations or experiments where 'all cards are on the table'. As such, game theory provides a more realistic representation of reality. Furthermore, participants are forced to weigh up their options – offering an insight into their decision-making process. For these reasons, game theory has been employed in a variety of research domains. For example, game theory has had a major influence on the development of Economics [7], allowing researchers to quantify decision making from microeconomics to macroeconomics. Rather than asking participants to rate their own or a peer's level of morality, we consider game theory an inventive solution which can be used to determine participant convictions.

Human contributions

As aforementioned, smartphones have become excellent research tools due to their connectedness, personal nature, and interactive capabilities. While traditional applications can naturally be used to elicit contributions, the Experience Sampling Method (ESM), pioneered in the late 1970's [16], has been very successfully transitioned to the mobile domain [2, 3]. The ESM allows for repeated *in situ* data collection through participant self-reports, enabling researchers to obtain insights into shifts in participant answers. The combination of the mobile ESM and a reverse second price Vickrey Auction is a promising approach due to the always-on nature of smartphones and the capacity of auctions to elicit truthful information [19] – something that can arguably be difficult to obtain when it comes to surveying issues about morality. Furthermore, participant completion of game theory scenarios can be complemented with additional information not easily obtained from sensor information (*e.g.*, company of the participants).

Contextual information

Smartphones have established themselves as a widely used research instrument, specifically for the collection of participant data *in situ* [15, 21]. Not only can researchers develop and present interactive applications (such as those used for game theory), smartphones also allow for the continuous and automatic collection of contextual information about the participant, as shown by Tag et al. [23]. The longitudinal collection of this data allows researchers to "*run perceptually and behaviorally rich experiments*" [21]. As participants naturally change their context throughout the day (*e.g.*, going to work, spending time with friends), capturing participant answers to game theory scenarios over an extended period of time enables the collection of responses across a rich contextual diversity.

4 CHALLENGES

Previous work on *in situ* mobile studies has highlighted a number of challenges faced by researchers when deploying their studies, *e.g.*, connectivity, recruitment. The use of game theory in a non-controlled and mobile study setup introduces additional challenges, both technical and methodological in nature. Here, we discuss the main challenges we identify and highlight potential solutions to address these challenges.

Game partner

Most scenarios found in the game theory literature require a partner to interact with. Allowing participants to complete a scenario on their own is not viable, as this does not satisfy the criteria of incomplete information. It is critical that both players are unaware of the other player's choice until after they have made their decision. Furthermore, game theory requires that the actions of one player affect the end-result of the other player(s). Although it is straightforward to conduct experiments in the lab in which two participants are present, it is significantly more challenging when running studies 'in the wild'. As participants will not be constantly available, the system has to find alternative solutions. First, the system can match participants with whomever is available at the time rather than consistently matching up participants in identical pairs. Second, researchers can consider alternative game theory scenarios in which a large number of actors can simultaneously participate. Third, as the aforementioned can

be technically challenging and does not necessarily resolve all issues (*i.e.*, uneven numbers of participants in the case of two-player scenarios), researchers can consider deception of participants as a viable method. The use of automated bots, or simply a random answer selector, instead of real players will allow participants to always find a game partner.

Game synchrony

Although game theory does not require instantaneous and simultaneous decisions from participants, a typical element of these games is that participants receive the outcome of the game shortly after submitting their decision. This can be challenging to achieve if a participant is matched with someone who fails to respond to an incoming notification. Sampling strategies found in the Experience Sampling literature frequently allow for a time-window within which participants should respond to a notification before the notification expires [2]. A similar timeout system can be employed in the context of mobile game theory, ensuring that participants are matched with participants as based on their availability while allowing for a time window of response time.

Longitudinal motivation

Study fatigue has been shown to effect participants in longitudinal studies, resulting in a drop of both the quantity and quality of participant responses to mobile questionnaires [4, 5]. It is therefore critical for researchers to identify mechanisms which motivate participants over extended periods of time. An extrinsic method frequently used in the literature is that of financial incentives [13]. Although the use of (extensive) monetary rewards can introduce additional complications with participant motivation [22], the use of monetary rewards – in particular the use of micro-payments – is complimentary to game theory applications. By linking micro-payments to the potential outcomes of a game theory scenario, participant motivation may be sustained for longer periods, while simultaneously ensuring a vested interested among participants to obtain a favourable outcome.

Technological challenges

While the increased prevalence of smartphone sensing capabilities via onboard sensors has enabled various new applications as well as research prototypes, the current fragmentation of device families makes development challenging. To this end, different mobile sensing frameworks have emerged. For example, the AWARE framework [10] and the CARP framework [1] both enable building mobile applications with powerful sensing capabilities for capturing the user context. The open-source nature of these endeavours allows researchers to relatively quickly deploy an application.

[1] https://github.com/cph-cachet/carp.sensing-flutter

5 FUTURE WORK

Even though this work offers only a preliminary perspective on the study of contextual morality using game theory, a number of avenues for future works emerge.

- Explore suitable game theory scenarios. The literature distinguishes a number of mechanisms used in the design of game theory scenarios, such as dynamic games (games that are sequential rather than simultaneous in nature), second-price auction ('sealed-bid' auction in which the highest bidder wins but ends up paying the monetary value of the second-highest bid), and the pivotal mechanism (used to obtain honest reporting of participant preferences for *e.g.* a public cause), to name a few [7]. Which mechanisms are most suitable for the collection of contextual morality requires further investigation.
- Develop data collection software. The deployment of mobile applications for the purpose of experimenting with game theory will require substantial technical development. In particular, synchronising game states between participants who may or may not respond to incoming notifications raises methodological and technical challenges.
- Collect human responses in the wild. The critical difference between the presented concept and previous work on game theory is the collection of contextual data. This will allow researchers to identify the effect of context on participants' actions, and thereby inform the development of future AI algorithms. This paper is a call to action, and we believe that the HCI community has an important role to play in the development of AI technology which is adaptive to people's context.

Potential study outline

We present an outline of a potential future study in which game theory is applied to measure contextual morality in a longitudinal 'in the wild' setting.

This study explores the effect of both location and peer-presence on participants' morality. Hypotheses are that participants behave more moral in the presence of others and while in public locations (as compared to while being alone and in private locations, *e.g.* at home). To obtain an insight into (contextual changes in) the participants' morality, we collect mobile sensor data and human input data using the Experience Sampling Method at multiple times throughout the day [2]. Human input data consists of both 'traditional' self-report data as well as the repeated completion of a straightforward game theory scenario. Participant self-report data consists of the Moral Identity Questionnaire [6], as well as the participants' answer to questions on their current company

and physical location. In the presented game scenario, participants play a modified version of the Prisoner's Dilemma in collaboration with another participant. Participants can earn a small payment, depending on the outcome of the game. By collecting these data over an extensive period of time, there is a high propensity to obtain contextually rich readings – allowing us to obtain insights into the anticipated effect of location and peers on our contextual morality.

6 CONCLUSION

In this work, we outline the concept of applying game theory to capture insights on human contextual morality. Given the increasing ubiquity of AI applications, there is a growing need to understand the effect of context on human morality and subsequently adjust AI applications accordingly. Given the inherit difficult in obtaining insights in an individual's morality, we consider the use of game theory a viable approach worth exploring. Although we outline a number of open challenges, the potential payoffs of an increased understanding of contextual morality are substantial, both addressing an intellectual curiosity in human morality and a practical application in matching AI behaviour and user requirements.

REFERENCES

[1] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. 2018. The Moral Machine experiment. *Nature* 563, 7729 (2018), 59.

[2] Niels van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. *ACM Comput. Surv.* 50, 6, Article 93 (Dec. 2017), 40 pages. https://doi.org/10.1145/3123988

[3] Niels van Berkel, Jorge Goncalves, Simo Hosio, and Vassilis Kostakos. 2017. Gamification of Mobile Experience Sampling Improves Data Quality and Quantity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 107 (Sept. 2017), 21 pages. https://doi.org/10.1145/3130972

[4] Niels van Berkel, Jorge Goncalves, Peter Koval, Simo Hosio, Tilman Dingler, Denzil Ferreira, and Vassilis Kostakos. 2019. Context-Informed Scheduling and Analysis: Improving Accuracy of Mobile Self-Reports. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 51, 12 pages. https://doi.org/10.1145/3290605.3300281

[5] Niels van Berkel, Jorge Goncalves, Lauri Loven, Denzil Ferreira, Simo Hosio, and Vassilis Kostakos. 2019. Effect of experience sampling schedules on response rate and recall accuracy of objective self-reports. *International Journal of Human-Computer Studies* 125 (2019), 118 – 128. https://doi.org/10.1016/j.ijhcs.2018.12.002

[6] Jessica E. Black and William M. Reynolds. 2016. Development, reliability, and validity of the Moral Identity Questionnaire. *Personality and Individual Differences* 97 (2016), 120 – 129. https://doi.org/10.1016/j.paid.2016.03.041

[7] Giacomo Bonanno. 2018. *Game Theory*. North Charleston, SC, USA: CreateSpace Independent Publishing Platform. 592 pages.

[8] Kate Crawford and Ryan Calo. 2016. There is a blind spot in AI research. *Nature News* 538, 7625 (2016), 311.

[9] Jelle De Boer. 2017. Social Preferences and Context Sensitivity. *Games* 8, 4 (2017). https://doi.org/10.3390/g8040043

[10] Denzil Ferreira, Vassilis Kostakos, and Anind K Dey. 2015. AWARE: mobile context instrumentation framework. *Frontiers in ICT* 2 (2015), 6.

[11] Matthew Grizzard, Ron Tamborini, Robert J. Lewis, Lu Wang, and Sujay Prabhu. 2014. Being Bad in a Video Game Can Make Us Morally Sensitive. *Cyberpsychology, Behavior, and Social Networking* 17, 8 (2014), 499–504. https://doi.org/10.1089/cyber.2013.0658 PMID: 24950172.

[12] Simo Hosio, Andy Alorwu, Niels van Berkel, Miguel Bordallo López, Mahalakshmy Seetharaman, Jonas Oppenlaender, and Jorge Goncalves. 2019. Fueling AI with Public Displays?: A Feasibility Study of Collecting Biometrically Tagged Consensual Data on a University Campus. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays (PerDis '19)*. ACM, New York, NY, USA, Article 14, 7 pages. https://doi.org/10.1145/3321335.3324943

[13] Simo Hosio, Denzil Ferreira, Jorge Goncalves, Niels van Berkel, Chu Luo, Muzamil Ahmed, Huber Flores, and Vassilis Kostakos. 2016. Monetary Assessment of Battery Life on Smartphones. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1869–1880. https://doi.org/10.1145/2858036.2858285

[14] Dan R. Johnson, Brandie L. Huffman, and Danny M. Jasper. 2014. Changing Race Boundary Perception by Reading Narrative Fiction. *Basic and Applied Social Psychology* 36, 1 (2014), 83–90. https://doi.org/10.1080/01973533.2013.856791

[15] Vassilis Kostakos, Denzil Ferreira, Jorge Goncalves, and Simo Hosio. 2016. Modelling Smartphone Usage: A Markov State Transition Model. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 486–497. https://doi.org/10.1145/2971648.2971669

[16] Reed Larson and Mihaly Csikszentmihalyi. 2014. *The Experience Sampling Method*. Springer Netherlands, Dordrecht, 21–34. https://doi.org/10.1007/978-94-017-9088-8_2

[17] Keith Leavitt, Scott J. Reynolds, Christopher M. Barnes, Pauline Schilpzand, and Sean T. Hannah. 2012. Different Hats, Different Obligations: Plural Occupational Identities and Situated Moral Judgments. *Academy of Management Journal* 55, 6 (2012), 1316–1333. https://doi.org/10.5465/amj.2010.1023 arXiv:https://doi.org/10.5465/amj.2010.1023

[18] Christoph Lütge, Hannes Rusch, Matthias Uhl, and Christoph Luetge. 2014. *Experimental ethics: Toward an empirical moral philosophy*. Springer.

[19] R. Preston McAfee and John McMillan. 1987. Auctions and Bidding. *Journal of Economic Literature* 25, 2 (1987), 699–738. http://www.jstor.org/stable/2726107

[20] J.S. Mill. 1863. *Utilitarianism*. Parker. https://books.google.co.uk/books?id=lyUCAAAAQAAJ

[21] Geoffrey Miller. 2012. The Smartphone Psychology Manifesto. *Perspectives on Psychological Science* 7, 3 (2012), 221–237. https://doi.org/10.1177/1745691612441215 PMID: 26168460.

[22] Arthur A. Stone, Ronald C. Kessler, and Jennifer A. Haythomthwatte. 1991. Measuring Daily Events and Experiences: Decisions for the Researcher. *Journal of Personality* 59, 3 (1991), 575–607. https://doi.org/10.1111/j.1467-6494.1991.tb00260.x

[23] Benjamin Tag, Andrew W Vargo, Aman Gupta, George Chernyshov, Kai Kunze, and Tilman Dingler. 2019. Continuous Alertness Assessments: Using EOG Glasses to Unobtrusively Monitor Fatigue Levels In-The-Wild. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 464:1–464:12. https://doi.org/10.1145/3290605.3300694

[24] James Zou and Londa Schiebinger. 2018. AI can be sexist and racist - it's time to make it fair. 559 (2018), 324–326. https://doi.org/10.1038/d41586-018-05707-8