The Case for Transparent Models of Non-Verbal Behaviours in Interdisciplinary Robotics Research

Extended Abstract

Carina E. I. Westling
University of Sussex
Brighton BN1 9RG
United Kingdom
c.e.i.westling@sussex.ac.uk

Nachiappan Chockalingam
University of Staffordshire
Stoke-on-Trent
United Kingdom
n.chockalingam@staffs.ac.uk

Harry J. Witchel
BSMS
Brighton BN1 9RY
United Kingdom
h.witchel@bsms.ac.uk

ABSTRACT

The lack of transparency in machine learning for the decoding of non-verbal behaviours in robot-human interaction fails to adequately support interdisciplinary research. Recent research questions both the neutrality and the collegial and collaborative underpinnings of machine learning methodologies. In disciplines such as neuroscience (or medicine more broadly) and experimental psychology, transparency is required in models derived from research, so that results can be modularly verified (i.e. replicated piecemeal) as well as expanded and built upon. The opacity of models developed with hypothesis-free machine learning can thus limit the potential for interdisciplinary research progression and development.

Our team seeks to integrate progress made in the field of robotics with extant and emergent research in the field of non-verbal behavior research. Our contributions to the field of non-verbal behavior in the context of human-computer interaction include the recognition and definition of Non-Instrumental Movement Inhibition (NIMI), a metric of cognitive engagement and alignment with stimulus objectives that describes the degree to which 'unnecessary' (non-instrumental) actions are suppressed during engagement with a stimulus. This approach allows the differentiation of dynamic engagement from rapt engagement, and the differentiation of lethargic boredom from restless boredom. The NIMI model presumes the task/activity incorporates screenic engagement (gaze) and thus requires context-awareness in interaction design. Hence, it offers a case in point for transparent models.

CCS CONCEPTS

• Human-centered computing~HCI theory, concepts and models • Human-centered computing~Interaction design theory, concepts and paradigms • *Computer systems organization~Robotics*

KEYWORDS

Machine learning, motion capture, engagement, robot-human interaction, non-verbal communication, NIMI, non-instrumental movement inhibition

ACM Reference format:

C. E. I. Westling, N. Chockalingam, and H. J. Witchel. 2017. The Case for Transparent Models of Non-Verbal Behaviours in Interdisciplinary Robotics Research. In *Proceedings of European Conference of Cognitive Ergonomics 2017 conference, Umeå, Sweden, September 2017 (ECCE'2017)*, 2 pages.

ECCE 2017, Umeå

1 INTRODUCTION

Recent research has shed light on issues with embedded bias and the difficulty of reviewing machine learning algorithms [1], both of which have potentially serious consequences as more decisions are delegated to automated systems. As the purported neutrality of machine learning comes under scrutiny, additional problems associated with this lack of transparency will enter the academic discourse, including interdisciplinary collaboration over methods that cannot be reviewed. Meaningful communication between two or more agents requires not just the parsing of interaction in the moving present, but interpretation within a framework of mutual 'understanding' [2], i.e. a shared system of meaning. Thus communication is a product not only of the need or desire to exchange symbols and acts of meaning in a given situation, but also of the larger, shared meaning-making structures of societal groups. Such groups include the broader research community, where, in order for meaningful interaction to occur, shared frames of reference are required.

An example where a shared system of meaning is essential can be found in our team's research of the decoding of human non-verbal signals of engagement. We gathered a broad range of empirical motion data during naturalistic (i.e. non-acted) human-computer interaction with sets of discrete stimuli, which we subsequently subjected to manual and semi-automatic analysis [3]. In the course of this research, we have tested several common hypotheses from the non-verbal communication field, including the assumption that engagement leads to more movement [4, 5, 6] and that engagement is a function of positive affect [5, 6], and found that neither is necessarily true when

applied across different stimuli. To accommodate the seemingly divergent behaviors that signify engagement vs. boredom during screen engagement, we formulated the NIMI (non-instrumental movement inhibition) model [6], introducing the idea that engagement manifests as alignment with the activities that a stimulus requires from the participant (ranging from stillness to high-activity actions) [4], and that boredom or disengagement manifests as 'escape attempts', and can include both/either lethargy or restlessness, often presenting cyclically, but *misaligned* with the stimulus.

2 THE CASE FOR TRANSPARENT MODELS

Without transparent models and shared frames of references, further research, and perhaps especially interdisciplinary research, is compromised. Recent findings do not support the previously assumed neutrality of machine learning methodologies [7]. In the interest of enhancing the potential for interdisciplinary collaboration with research fields in which modular verification and review of methodologies is essential and to allow for further development of extant research, we propose that hypothesis-led and supervised learning should be revisited more generally, and specifically so for motion capture and non-verbal communication research.

2.1 NIMI in robot-human interaction

NIMI (non-instrumental movement inhibition) emphasizes the need for context awareness in robot-human interaction design, as it is key to the accurate interpretation of human non-verbal signals. Increased movement can be a sign of *either* engagement or disengagement (caused by, for example, boredom), and the issue is whether the increased movement is a response to affordances, e.g. instrumental movements, or agitation associated with suppressed 'escape attempts' or frustration.

Conversely, the *inhibition* of non-instrumental movement (NIMI), i.e. the suspension of fidgeting and baseline changes in position that maintain comfort over time, has been found to be associated with engagement [3, 4, 5]. It manifests as the diminished occurrence of movements that are unnecessary and/or unrelated to the designed activity, rather than being based on measurement of instrumental movements associated with the designed activity. NIMI, as a metric of engagement, is thus independent of the type and intensity of actions that are instrumental to any particular designed context. However, in order to distinguish between aligned and unaligned movements, NIMI (as a model for emergent human-robot non-verbal interaction) relies on context-awareness, which highlights the importance of interdisciplinary collaboration and transparent models.

2.2 Proposed applications

NIMI, as defined by our team of researchers, can be applied to refining non-verbal interaction (for the dual purposes of encoding and decoding) between human and artificial humanoid agents, as well as to the interpretation of human non-verbal interaction with other applications (for the purpose of decoding, e.g. automated tutoring systems). As it can be applied as a metric of engagement that is independent of the type and intensity of instrumental movements in any given interaction, it would be of particular value in emergent communication between humans and humanoid agents.

ACKNOWLEDGMENTS

Many thanks to Chätrin Tolga, Jacob Greaves, Terri Desmonds, and the BSMS IRP programme (for funding). Also to Till Lindemann, Richard Z. Kruspe, Paul H. Landers, Oliver Riedel, Christoph Schneider and Christian Lorenz for their original ideas on interaction design.

REFERENCES

[1] A. Datta, S. Sen and Y. Zick. 2016. Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. In *Security and Privacy (SP), 2016 IEEE Symposium on*, 598-617. IEEE.

[2] Stuart Hall (Ed.). 1997. Representation: Cultural Representations and Signifying Practices (Vol. 2). SAGE.

[3] H. J. Witchel, C. E. I. Westling, J. Tee, R. Needham, A. Haely and N. Chockalingam. 2014. A time series feature of variability to detect two types of boredom from motion capture of the head and shoulders. In *Proceedings of the European Conference on Cognitive Ergonomics (ECCE '14), Vienna, Austria, 2014.* ACM.

[4] Nadia Bianchi-Berthouze. 2013. Understanding the role of body movement in player engagement. In *Human–Computer Interaction*, 28(1), 40-75.

[5] H. J. Witchel, C. P. Santos, J. K. Ackah, C. E. I. Westling and N. Chockalingam. 2016. Non-Instrumental Movement Inhibition (NIMI) differentially suppresses head and thigh movements during screenic engagement: dependence on interaction. *Frontiers in Psychology*, 7 (157). ISSN 1664-1078.

[6] H. J. Witchel, C. E. I. Westling, J. Tee, A. Healy, R. Needham, and N. Chockalingam. 2014. What does not happen: quantifying embodied engagement using NIMI and self-adaptors. *Participations*, 11 (1), 304-331. ISSN 1749-8716.

[7] S. Hajian, F. Bonchi,and C. Castillo. 2016. Algorithmic bias: from discrimination discovery to fairness-aware data mining. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2125-2126. ACM.