Social Robotics

PDEEC PhD course on Social Robotics

João Silva Sequeira[1]

[1]Instituto Superior Técnico
Institute for Systems and Robotics

Spring 2021

TÉCNICO LISBOA

Social norms I

- A social norm is a customary rule/habit/behavioral pattern that, if not followed, results in rejection by observers (see [Andrighetto et al, 1998])

- A social norm is a collection of degrees of acceptability/rejection of a number of factors

  "'Norm' is a term used to refer to a variety of behaviors, and accompanying expectations." (see [Bicchieri, 2006])

Social norms II

- Game theory and Utility theory often used as framework to model social norms (see [Bicchieri, 2006])

- Descriptive norms and conventions as solutions to coordination games, ([Bicchieri, 2006] argues that these are not social norms)

- Some social norms, on the contrary, often go against narrow self-interest, [Bicchieri, 2006]

  Social norms are situations in which there is conflict of interest, but also a potential for joint gain, [Bicchieri, 2006]

Social norms III

The pragmatic aspects surrounding a social norm include the following,
[Savarimuthu al, 2010]:

- "Normative expectation of a behavioural regularity: There is a general agreement
  within the society that a behaviour is expected on the part of an agent (or actor)
  by others in a society, in a given circumstance."

- "Norm enforcement mechanism: When an agent does not follow a norm, it could
  be subjected to a sanction. The sanction could include monetary or physical
  punishment in the real world which can trigger emotions (embarrassment, guilt,
  etc.) or direct loss of utility. Other kind of sanctions could include agents not
  being willing to interact with an agent that violated the norm or the decrease of
  its reputation score."

- "Norm spreading mechanism: Having obtained a norm, an agent may spread the
  norm to other agents in the society. Examples of norm spreading mechanisms
  include the notion of advice from powerful leaders, imitation and learning on the
  part of an agent."

TÉCNICO LISBOA

Indicators of social norm compliance I

Example of indicators for socially accepted behavior in a (undergrad) classroom

- Frequency of (any) words by students much smaller than that of teacher
  - Which may not be of clear interest for the student(s) that must keep up with FB

- High correlation between the words by the students and those by the teacher (suggesting that students are interested in the topic lectured)

Indicators of social norm compliance II
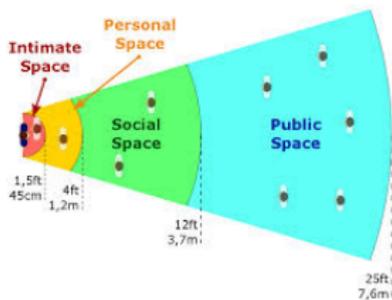
(a)
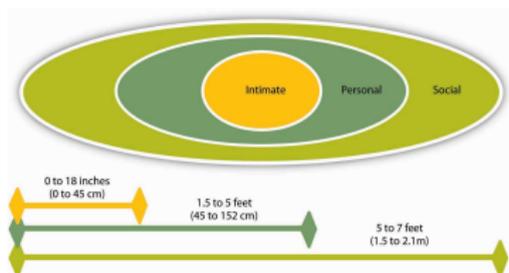
(b)

- Proxemics is a social norm

- A robot can use the trajectories of other agents to estimate their compliance with the norm

  Similarly, it can generate its own trajectories to comply with proxemics social norms of others

TÉCNICO LISBOA

Adjusting to social norms

- Members of society gradually learn to follow a system of conventions via "a slow progression, and by our repeated experience of the inconveniences of transgressing it" (David Hume quoted in [Vanderschraaf, 2014])

- *Fictitious play*; "human-like" technique used to foresee the evolution of systems

  Markov fictitious play; variant of the original technique (see [Skyrms, 2014])

  The idea is to use a model to estimate the evolution and check if it complies with the norms

The social contract in social robotics I

- Social contracts (SC) in human societies amount to maintain an equilibrium in the respective social dynamics

- A SC expresses a "certain" form of behavior
    - Society members cooperate

    - Society members do not cooperate

- Human SCs are often implicit and depend on trust

The social contract in social robotics II

- "Stag Hunt (SH) game constitutes the prototypical example of the social contract ...", [Skyrms, 2014] – The values in the table are payoffs (reward) obtained by each player following their decisions

	Stag	Hare
Stag	*a, a*	*c, b*
Hare	*b, c*	*d, d*

$a > b \geq d > c$

- Two hunters must decide separately, and without the other knowing, whether to hunt a stag or a hare

- Both hunters know the only way to successfully hunt a stag is collaborating with each others

- If both opt for non-collaboration it's certain that both get a hare, which is not a worst case situation

The social contract in social robotics III

Prisoner's Dilemma		Player 2	
		Confess	Stay Silent
Player 1	Confess	(2,2)	(4,1)
	Stay Silent	(1,4)	(3,3)

- The payoffs in the table can be interpreted as the prison years given to each player

- A crime was committed and 2 people are being charged

- Each player does not know what action the opponent is going to choose

- Staying silent is the best way to ensure that none of them gets the maximum penality

The social contract in social robotics IV

- Game theory is often used to model SCs

  "Whenever collective action of groups of individuals is at stake N-person games are appropriate", [Skyrms, 2014]

- A SC is a social relationship, i.e., it represents a behavioral model

Additional remarks on social models

- Understand how any of the relevant state variables evolves

  State variables are those variables such that their evolution in time describes completely the system

  Often very difficult to have an algebraic representation

- Is it always possible to enumerate all state variables in a realistic social robotics problem?

  Some variables may not be observable; others may lie hidden

Good properties in social models I

- Disturbances are bounded

- Observable variables are known; unobservable may/may not be known

  In some situations relevant (state) variables may not be known to exist, i.e., the dimension of the system is underestimated

- Markov property

- Short/Medium term stability

- Equilibria and long term stability – meaning that interesting social contracts may exist

An (incomplete) attempt to define a social robot I

- A social robot has personality

  Behaviors are admissible for execution according to the personality of the robot

- It is capable of displaying and recognizing emotions

  Generated emotions act as configurators for behaviors

- It is capable of estimating social norms

  These norms are estimated from the robot's perspective, i.e., they may not be coherent with estimates from other agents (in which case the robot may end up being rejected)

- It can adjust its selection of admissible behaviors and their configuration to a set of social norms

  This amounts to establishing a social contract – in a sense, this is a control goal

An (incomplete) attempt to define a social robot II

- It does not need to be able to solve complex problems

  It can be a low IQ robot, no need for intelligence of types 5 (intrapersonal) and 7 (logical-mathematical)– see [Gardner, 1983].

  It may need intelligence of type 4 (interpersonal)

Local bibliography

📄 Giulia Andrighetto, Cristiano Castelfranchi, Eunate Mayor, John McBreen, Maite Lopez-Sanchez, Simon Parsons
"(Social) Norm Dynamics"
In "Normative Multi-Agent Systems". Dagstuhl Follow-Ups, Vol. 4. ISBN 978-3-939897-51-4. Editors: Giulia Andrighetto, Guido Governatori, Pablo Noriega, and Leendert W.N. van der Torre; pp. 135–170, Dagstuhl Publishing

📄 Brian Skyrms
"Social Dynamics"
Oxford University Press, 2014

📄 Peter Vanderschraaf
"Learning to Take Turns: The Basics"
In [Skyrms, 2014]

📄 Cristina Bicchieri
"The Grammar of Society: The nature and Dynamics of Social Norms"
Cambridge University Press 2006

📄 Tony Savarimuthu, Stephen Cranefield, Maryam A. Purvis, Martin K. Purvis
"Obligation Norm Identification in Agent Societies"
Journal of Artificial Societies and Social Simulation 13 (4) 3

📄 H. Gardner
"Frames of Mind: The Theory of Multiple Intelligences"
1983

TÉCNICO LISBOA