Using Machine Theory of Mind to Learn Agent Social Network Structures from Observed Interactive Behaviors

with Targets Yun-Shiuan Chuang, Hsin-Yi Hung, Edwinn Gamborino, Joshua Goh\*, Tsung-Ren Huang, Yu-Ling Chang, Su-Ling

Yeh, Li-Chen Fu

-------------------------------------------------------------------------------------------------------------------------------

DECISION: Accepted as Regular paper

The final version of the paper, together with the IEEE copyright form must be submitted using PaperPlaza by 27 July

2020 (23:59 US Pacific time)

selected numbers of presented papers will be invited to submit the revised version of

their work in our special issues.

-------------------------------------------------------------------------------------------------------------------------------

REVIEWER 9 (ID: 1631)

This paper is about proposing a ML approach to infer

implicit relationships in people's social networks from

observation. They consider this from a NN and interpersonal

context (not social media network analysis). This kind of

knowledge is certainly relevant to have social agents know

in order to interact with people in a more appropriate way.

I have several concerns about the paper. I provide this

feedback in the hopes that the authors can use it to

improve and strengthen the work. It will also help to make

this work more relevant for Ro-Man that tends to be a

human-robot interaction conference.

1. There is a long tradition of modeling Theory of Mind in

social robots. The authors do not seem to be aware of this

prior work, and rather only cite their on papers.

Scassellati first explored this in the early 2000s.

Breazeal has done a fair amount of work on this topic in

her group. Yiannis Demiris as as well. The authors do not

justify their contribution in light of all this prior work.

2. Theory of Mind is a very sophisticated social-cognitive

sent of abilities. The authors use of ToM in their context

is \*not\* appropriate. ToM is also much more than inferring

false beliefs. I recommend they not use ToM to refer to

using NN to learn hidden states.

3. Other works have inferred implicit social network

connections from in-person encounters with social robots. I

recommend the authors read the work of Kanda. They do not

seem to be aware of this work. Kanda did not use NN, but

it's important that any contribution show advancement

beyond the state of the art. Without explicit discussion

with prior related work, the authors can't claim a

contribution.

4. The authors are using simulated data, or a mix of human

questionnaire data and simulated data. The actual task is

hard to understand in the paper. It's difficult to

understand the results if one can't quite make out what the

task/game/goals are in the activity. In any case, it seems

to assume the agent can see everything "birds eye view".

This isn't realistic in real-world settings. Ro-Man is a

conference that tends to prefer real-world work with robots

and people. If simulation is used, it needs to be clear how

it's results advance real-world HRI. I don't think these

assumptions are reasonable to do so.

5. The authors make a claim that using a NN model somehow

automatically lends insight into how human make these

inferences. This is significantly over-reaching. If their

goal was to model human decision making, then they should

have designed a study that was really focusing on

understanding human behavior. This paper is primarily about

a algorithmic approach rather than deep cognitive

neuroscience modeling.

For all these reasons, I think this paper is far from being

suitable for a conference like Ro\_Man. A new experiment

would need to be done. Far more discussion would need to be

added in terms of related work, too.

REVIEWER 14 (ID: 1641)

The authors applied an artificial neural network with

machine theory of mind (ToMnet+) to learn and predict

social preferences based on implicit information from the

way agents and social targets interact behaviorally. Their

findings have implications for machine applications that

seek to infer hidden information structures solely from

third-person observation of behaviors.

(1) “the evidence shows that social machines that infer

user mental states through their implicit affective

behaviors can engage actions better catered to user

preferences”, which evidence?

(2) What is the advantage of using social network model

and ToMnet+?

(3) How do you determine the parameters?

(4) Fig. 4, are there any new improvement for your

method compared to traditional ToMnet+?

(5) Pls discuss the method in “Alcoholism

identification via convolutional neural network based on

parametric ReLU, dropout, and batch normalization”

REVIEWER 19 (ID: 1651)

The authors propose a framework to represent hidden social

networks from observed interactions, which extends a work

found in the literature whose purpose is to represent the

agent's hidden false beliefs.

The authors' motivation is the need to display more

human-like behaviors which can be achieved if the social

context is understood.

The authors pose the problem as an agent interacting with

four social targets in a grid world, where each target

provides a social reward and they are separated by some

random distance and barriers.

Overall, the idea is interesting and the paper follows an

intuitive, clear argumentation.

The main concerns of the reviewer are:

- It is unclear how the authors support the validity of

considering the physical distance in the grid as the social

distance between the agent and the target.

- As discussed by the authors, the game with human

participants can only reflect the desire to maximize the

score of the game rather than the underlying social

network. The paper would be enriched if potential

countermeasures are discussed by the authors.

- If the social rewards are very different between the

targets, will the distance in the grid lack of

significance?

Minor issues:

- In the first page (lower left), authors with the same

affiliation can be grouped.

- Use the proper latex command for arg max