Impact of Robot Tutor (Nonverbal Social Behavior) on Child Learning

Master Seminar

Sudeshna Dasgupta, Shruthi Narayani Venkatesh

Fakultät für Informatik
Technische Universität München
Objective

To propose solutions or improvements to address the shortcomings of methods used in past research to evaluate impact of robot tutor non verbal social behavior on child learning.
Introduction

• Efficiency of robots as tutors in educational contexts have been demonstrated in previous research works

• Increased social behavior of robot -> increased social presence -> better interaction outcomes

• Necessity to measure impact of social behavior on outcomes (non-verbal)
Literature Review

• Non Verbal Immediacy (NVI) as a feature to measure robot social behavior
• NVI -> “the extent to which communication behaviours enhance closeness to and nonverbal interaction with another”
• Social cues such as touching, distance, forward lean, eye contact, and body orientation etc., are used as parameters in NVI measurement questionnaire
• Hypothesis: Increased social behaviour -> increased learning outcome
Experiment

- Conducted between Aldebaran NAO robot and children of age 8-9 years
- Experiment consisted of children having 4 interactions with the robot and 1 with a human
- An experimenter briefs the child and introduces the child to the tutor
- Children complete pre-tests and post-tests in prime number identification, and division by 2, 3, 5, and 7 before and after interaction with tutor

Experiment

• During interaction, the tutor provides lessons on primes and dividing by 2, 3, 5, and 7.
• Children are not provided any feedback from the experimenter or tutor during the tests.
• Interactions with the tutor would last for around 10–15 min.
• Learning is measured through the improvement of the child’s score from pretest to posttest.
• After the interaction Nonverbal Immediacy Scores are collected from children and adults through questionnaire.
Non Verbal Immediacy Questionnaire

1. The robot uses its hands and arms to gesture while talking to you
   - Never
   - Rarely
   - Sometimes
   - Often
   - Very Often

2. The robot uses a dull voice while talking to you
   - Never
   - Rarely
   - Sometimes
   - Often
   - Very Often

3. The robot looks at you while talking to you
   - Never
   - Rarely
   - Sometimes
   - Often
   - Very Often

4. The robot frowns while talking to you
   - Never
   - Rarely
   - Sometimes
   - Often
   - Very Often

**Outcome**

- Earlier hypothesis that ‘Increased social behaviour -> increased learning outcome’ holds
- However, nonverbal immediacy does not account for all of the differences in learning
- Few results were in disagreement
- Possible explanations for discrepancy:
  - Timing of social cues is not calculated as a part of NVI
  - Cue (In)congruencies were not accounted for
Problem Statement 1

How can timing of social cues be accounted for in the measurement of Non-Verbal Social Behavior?
Experiment Data

Adult and child nonverbal immediacy ratings and child learning by tutor condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Adult M NVI rating [95% CI]</th>
<th>Child M NVI rating [95% CI]</th>
<th>Child learning (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low NVI robot</td>
<td>40.2 [38.1, 42.2]</td>
<td>51.0 [47.6, 54.4]</td>
<td>0.30</td>
</tr>
<tr>
<td>High NVI robot</td>
<td>48.4 [46.9, 50.0]</td>
<td>55.1 [52.3, 57.6]</td>
<td>0.67</td>
</tr>
<tr>
<td>Social robot</td>
<td>49.0 [47.6, 50.4]</td>
<td>N/A</td>
<td>0.51</td>
</tr>
<tr>
<td>Asocial robot</td>
<td>48.5 [46.1, 50.8]</td>
<td>N/A</td>
<td>0.89</td>
</tr>
<tr>
<td>Human</td>
<td>47.7 [45.3, 50.1]</td>
<td>54.4 [52.9, 55.9]</td>
<td>0.89</td>
</tr>
</tbody>
</table>


- Each child-robot interaction was 10-15 mins long and was recorded.
- A 42 s long video clip of each interaction was shown to adults to collect their feedback through the Robot Nonverbal Immediacy Questionnaire.
Proposed Methodology

• Identify each response time of the robot following a reaction from the child.
• Robot gets score 1 for each correct response and -1 for incorrect/delayed response.
• Response score = (Total positive response score) – (Total negative response score)
• Correlate response score with child learning.
Possible Results

• Pattern that suggests correlation between response score and child learning.

Limitations

• Final analyzed pattern derived from human-robot responses might not provide a generalized response.
Future work

• Increasing the experiment dataset may help to give an improved insight.

• Identification of the appropriate statistical model to retrieve accurate insights from the dataset.

• Identification of threshold value which can work as baseline for robot-human learning process.
Recap

• Hypothesis:
  Increased robot social behavior -> Increased child learning

• Discrepancy due to not accounting for social cue (in)congruency

• Revised Hypothesis:
  Increased robot social behavior + *Increased Cue Congruency* -> Increased child learning

• To measure cue congruency: Used Guttman’s Lambda

• However, this is not the best solution to measure cue congruency
Cue Congruency

- Social cues are interpreted as a single percept by the humans
- Congruence -> Synchrony in social cues
- Incongruence -> Mismatch between cues
- Incongruency between the social cues -> adverse effect in perception -> diminish learning outcome
Measuring (In)Congruency

• Guttman’s Lambda:
  Consistency measure applied across NVI questionnaire items to measure variability

• Shortcomings:
  • Isolation of cues
  • proxy for the congruency of cues as observed by the study participants

• Need for different way to measure cue congruency:
  Involves identifying combinatorial contextual expectations for social cues
Problem Statement 2

Is there a way to measure (In)Congruency of combinatorial contextual social cues directly from the study participants?

Proposed Solution

Using Bayesian models to measure combinatorial contextual social cues with user observation ratings
Bayesian Approach

- Modelling cue integration in multimodal physical perception and social cognition
- Perceivers view multimodal perceptual and social cues as ‘Conditional Probabilities’
- Example: Context -> won Olympic gold medal

\[
P(\text{happy} | \text{crying, medal}) \propto P(\text{crying} | \text{happy}) \times P(\text{happy} | \text{medal}) \times P(\text{happy})
\]

Methodology

• Applying the previous Bayesian analogy

• Modeling behavior of robot in a certain context given a set of cues

• Example: Context -> Directing/Pointing attention to something

\[
P(C|gaze, head, finger) = \frac{P(gaze|C) \cdot P(head|C) \cdot P(finger|C) \cdot P(C)}{P(gaze) \cdot P(head) \cdot P(finger)}
\]
Extending the model

• The previous equation models only the integration of cues
• But we need ‘Contextual’ integrations of cues

\[
P(C|\text{directing, gaze, head, finger}) = \frac{P(\text{directing}) \times P(\text{gaze}|C) \times P(\text{head}|C) \times P(\text{finger}|C) \times P(C)}{P(\text{directing}) \times P(\text{gaze}) \times P(\text{head}) \times P(\text{finger})}
\]

• Training data -> combinations of cues in various contexts rated as congruent or incongruent as observed by study participants
• If \( P(\text{congruence}|\text{context, cue1, cue2, ...}) > 0.5 \) it is classified as Congruent, otherwise Incongruent
Expected Result

• After training, the model is expected to classify an integrated set of cues in a given context as either congruent or incongruent

Discussion

• The probability measure can be used as an indicator of the level of (in)congruency
• Provides a holistic measure compared to Guttman’s Lambda
• Contribute to better evaluation of impact of cue congruency on learning
Limitation

- Data collection -> curating all possible cues in various contexts and obtaining participants’ (in)congruency ratings -> can be tedious

Future work

- More sophisticated model development
- More accurate and relevant parameter/feature selection
Conclusion

- Shortcomings of the methods proposed in previous literature, to evaluate impact of robot tutor non-verbal social behavior on child learning, were addressed and solutions were proposed.

- Shortcomings:
  - Timing not a part of NVI metric
  - Measuring cue (in)congruency on isolated cues and not incorporating participants’ observations.
Conclusion

• Proposed Solutions:
  • Account robot response time and correlate with child learning score to identify a pattern and a threshold response time that makes the social behavior of the robot more helpful for a child
  • Using Bayesian models on combinatorial contextual social cues with user observation ratings to measure social cue (in)congruency

• Contribution:
  • Aid in better measurement and modeling of robot tutor behavior in order to evaluate and increase the impact on child learning
Bibliography

