Understandable Robots

What, Why and How ?

Authors : Thomas Hellström and Suna Bensch
Presenter : Alen Jahic

Outline

- ❖ Introduction
- ❖ Related earlier work
- ❖ What does it mean to understand a robot
- ❖ Modeling interaction for understanding
- ❖ Designing interaction for understanding
- ❖ Conclusion & Future work
- ❖ Citations

Introduction

❖ Research efforts within HRI focus on making Human actions and thoughts understandable to robots

❖ Techniques for recognition of human activities, intentions and emotional states

❖ Highly important and not completed line of research

❖ How can we make robots understandable to Humans ?

❖ Solving this will

➢ Increase interaction quality
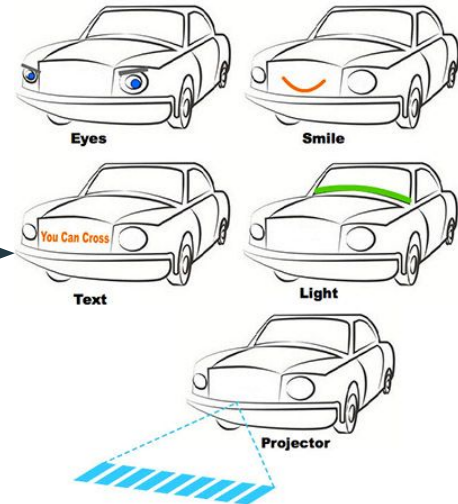
➢ Improve user experience
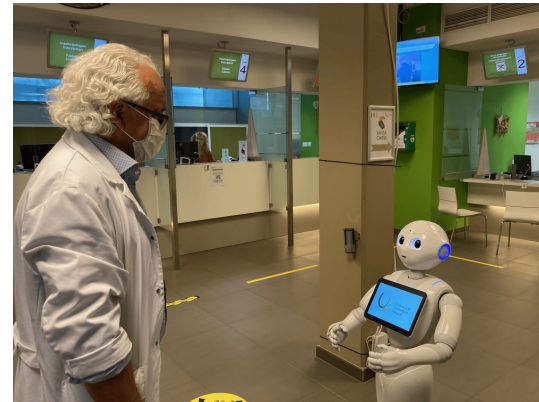
www.innovationorigins.com
Autonomous robots check patients in at Belgium hospitals

Introduction

Potential solutions →
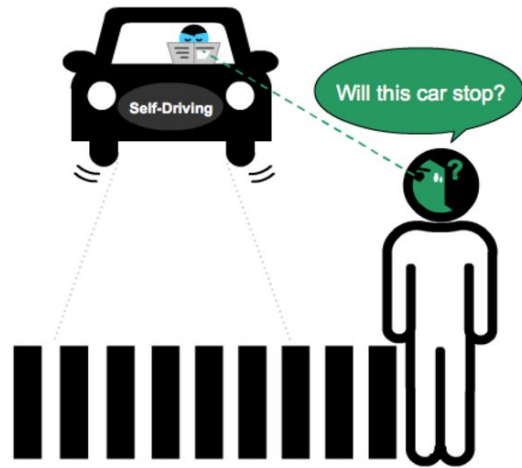
Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian

A Video-based Study Comparing Communication Modalities between an Autonomous Car and a Pedestrian

Introduction

- ❖ HRI Research Objective
  - ➢ Understandability / Understanding
- ❖ What does the concept mean ?
- ❖ How can it be formalized ?
- ❖ Objectives of the paper
  - ➢ Analyse what understanding of a robot means
  - ➢ Present a novel model of interaction for understanding
  - ➢ Formulate general guidelines for design of interaction for understanding

Related earlier work

❖ Importance of understandable robots acknowledged by the HRI community
❖ An in-depth survey of existing work has been performed
❖ Tight connection between understanding and communication
❖ Outcomes divided into areas
➢ Communication for understanding
➢ Humans understanding humans
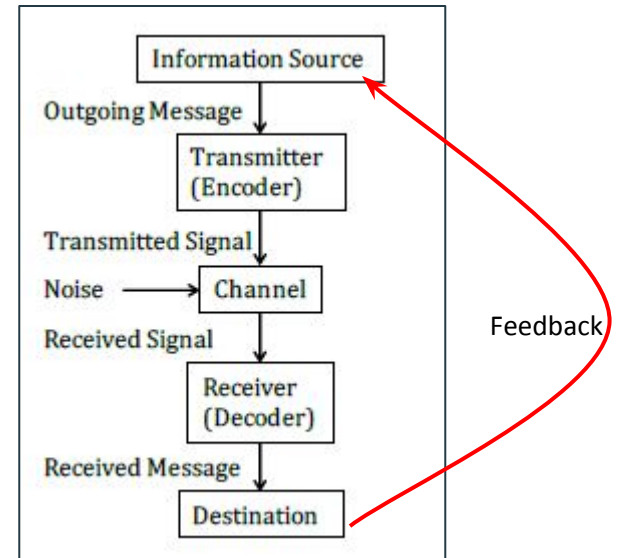➢ Humans understanding robots
➢ Robots understanding humans
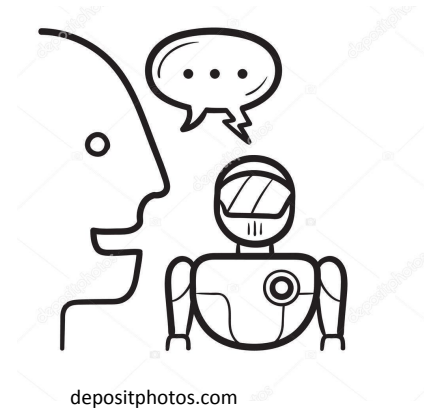
depositphotos.com

6

Related earlier work
- Communication for understanding

- ❖ Model suggested by Shannon in 1948
- ❖ Inappropriate for social sciences
- ❖ Criticism : Use of a postal metaphor
  - ➢ A physical package of information sent to a receiver
  - ➢ One-way and linear
  - ➢ Active sender
  - ➢ Passive receiver
  - ➢ No adaptation to responses
- ❖ Modified by Schramm
- ❖ Another criticism : Meaning is not taken into account
 - ➢ Decoding phase in interpersonal communication varies

Related earlier work

- Humans understanding humans

❖ Use of mindreading to estimate mental state
❖ Estimate actions by observing one's behavior
❖ Theory of Mind (ToM)
  ➢ Attribute mind with mental states
  ➢ Understand that others have believes,desires,... different from our own
  ➢ 3 important functions : Comprehend and explain, predict, manipulate
❖ 2 major views for how ToM works :
  ➢ Simulation theory
  ➢ Theory theory
❖ Principle of rationality
  ➢ All actions aim at reaching a goal
  ➢ Infer an agent's goal by observing its actions
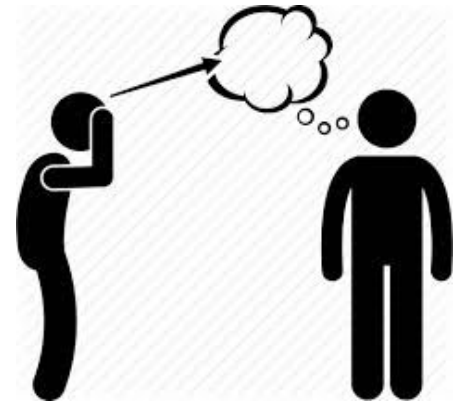
Related earlier work

- Humans understanding robots

- ❖ A Lot of research focus on equipping a robot with static functions
- ❖ Does not include ToM of interacting human
- ❖ Human's varying need of information not considered
- ❖ Exploit of Human's anthropomorphism
- ❖ Mental model of the robot's decision mechanism required
  - ➢ To predict a robot's future behaviour
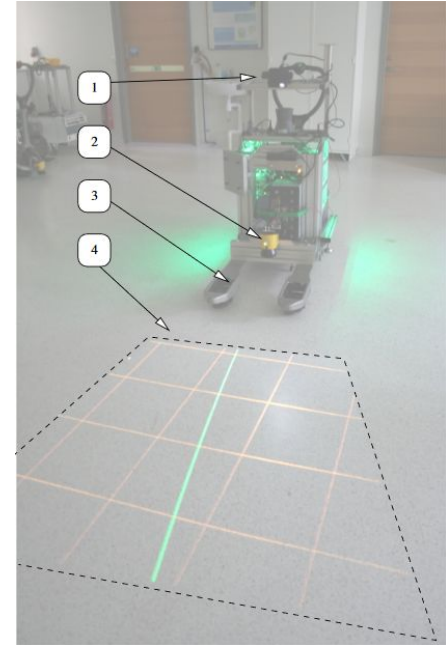
That's on my Mind! Robot to Human Intention Communication through on-board Projection on Shared Floor Space

Related earlier work

- Robots understanding humans

❖ Intention recognition
   ➢ Based on human utterances
   ➢ Inferred from body language and facial expressions
❖ Attempts to implement ToM in Robots
   ➢ By Devin and Alami
   ➢ By L. M. Hiatt, A. M. Harrison, J. G. Trafton
   ➢ ....
   ➢ Rather simplistic experiments

What does it mean to understand a robot

❖ Understanding
  ➢ "... a psychological process related to an abstract or physical object, such as a person, situation, or message whereby one is able to think about it and use concepts to deal adequately with that object" - C. Bereiter
❖ Interaction should be natural,efficient and safe
❖ Understanding of a robot
  ➢ Not limited to physical actions and intentions
  ➢ Includes non-physical entities such as knowledge, capabilities , goals, task progress, ….
  ➢ State-of-mind(SoM) = Collective of all entities

An agent's understanding of another agent is the extent to which the first agent has knowledge about the other agent's SoM in order to successfully interact with it

Modeling interaction for understanding

❖ Robots becoming more complex and autonomous hence harder to understand
❖ Support understanding by performing communicative actions
❖ Communicative actions should fit current perspectives and needs of the human
❖ Include ToM
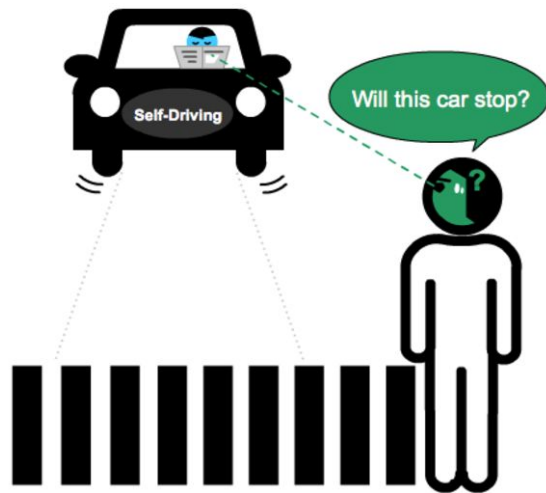 ➢ zeroth-order
  ➢ First-order
  ➢ Second-order

A communicative action is an action performed by an agent, with the intention of increasing another agent's knowledge of the first agent's SoM.

Modeling interaction for understanding
- Theory of Mind

Zeroth-order

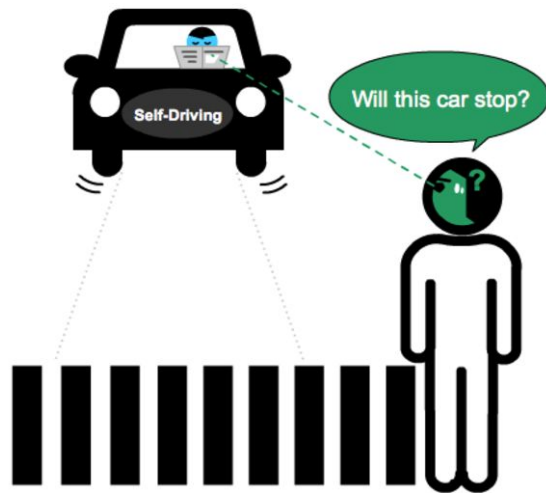❖ Car always assumes that the pedestrian will cross the road

Modeling interaction for understanding
- Theory of Mind

First-order

❖ Car uses a ToM to infer whether the pedestrian will cross the road
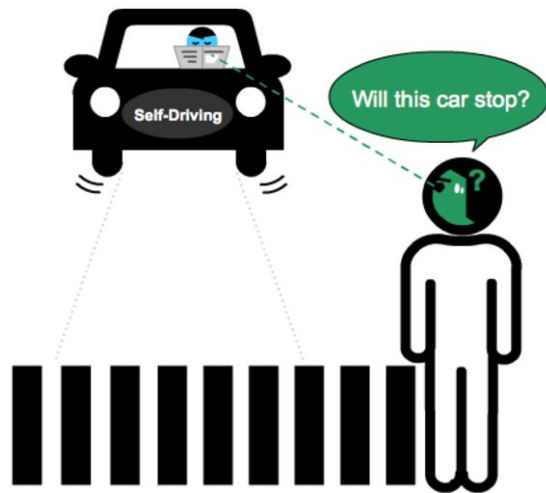

Will this car stop?

Self-Driving

Modeling interaction for understanding
-   Theory of Mind

Second-order

❖   Car uses a ToM to infer whether the pedestrian will cross the road and the pedestrian's belief regarding the car's intention to brake

Modeling interaction for understanding

- The idea

- ❖ Agent have a mind which includes a model of the interacting agent's mind
- ❖ React upon detection of mismatch between minds
  - ➢ Goal : Reduce the mismatch
- ❖ Generate, communicate, interpret communicative actions
- ❖ Build and extend on Shannon's model
- ❖ Make use of the connection between understanding and communication

Communication is the process by which we understand others and in turn endeavor to be understood by them

Modeling interaction for understanding
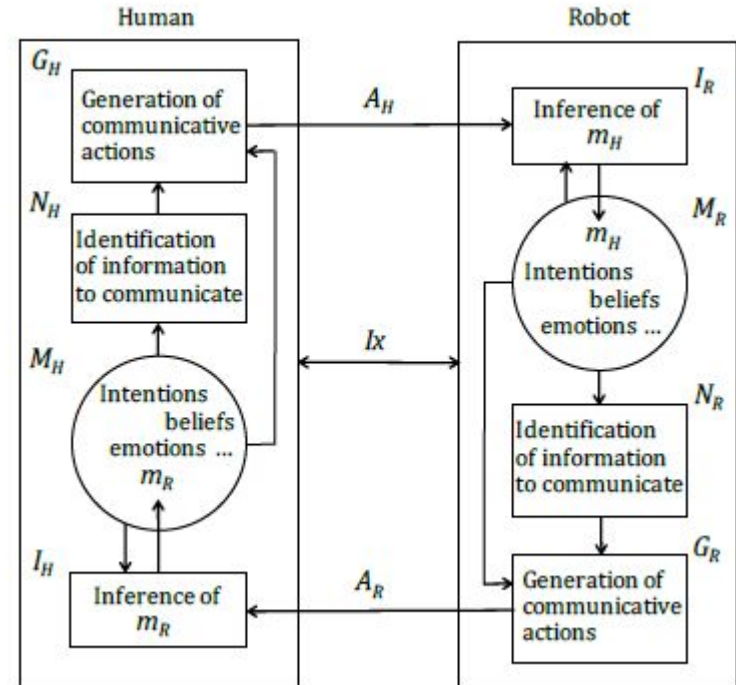- The Model

M_H = Human's SoM containing model of Robot's mind m_R

M_R = Robot's SoM containing model of Human's mind m_H

Human understanding of the Robot = $|M_R - m_R|$

Robot understanding of the Human = $|M_H - m_H|$

Modeling interaction for understanding
- Example

❖ Autonomous car detects pedestrian approaching the road

❖ Based on the traffic situation the car decides to not slow down

I_R : Car infers that the pedestrian believes that the car intends to slow down, since pedestrian is entering the road

N_R : Detects mismatch M_R and m_R. Reduce mismatch by communicating intention

G_R : Honking and flashing headlights chosen as communicative actions

I_H : Human interprets honking and flashing headlights as indicator that the car won't slow down and that it expects him to not cross the road

N_H : Estimate that there is no serious mismatch between M_H and m_H. Hence no need for communication

G_H : No communicative action generated, executed

Designing interaction for understanding

❖ Implementation requires application specific realizations of I_R , N_R and G_R
❖ Realization guided by addressing a set of questions to be answered
❖ 3 questions in the scope of the presented model
❖ 2 questions out of scope of the model

Designing interaction for understanding
- Question 1

❖ Q1 targets the realization of N_R

❖ Q1 : What information (if any) should be communicated to the human ?
  ➢ a) How should the mismatch $|M_R-m_R|$ be estimated ?
  ➢ b) How to determine if the mismatch is large enough to generate communicative actions ?
  ➢ c) Which information should be communicated to reduce the mismatch ?
  ➢ d) At which level of details should communication take place ?

Designing interaction for understanding
- Question 2 & 3

❖ Q2 targets the realization of I_R

❖ Q2 : How should the robot represent and infer the human's mind ?

➢ a) What entities (if any) of the human mind M_H should be represented in the robot's model m_H?

➢ b) How should these entities be represented?

➢ c) How should m_H be inferred from communicative actions A_H, from the robot's mind M_R, and from regular interaction I_X?

❖ Q3 targets the realization of G_R

❖ Q3 : How should communicative actions be generated to communicate the required information ?

Designing interaction for understanding
- Question 4 & 5

❖ Out of scope of the Model

❖ Q4 : To whom should the robot direct the communicative actions ?

❖ Q5 : Which mechanism should enable the model ?

➢ a) When should the information be communicated ?

➢ b) Should the robot initiate communication or should the robot respond to requests by the human ?

Designing interaction for understanding
- Example

- ❖ Example created using the work :
  - ➢ That's on my mind! robot to human intention communication through on-board projection on shared floor space by R. Chadalavada, H. Andreasson, R. Krug, A. J. Lilienthal
- ❖ Objective : Robotic forklift should interact smoothly and safely with humans moving in the same area

Designing interaction for understanding

  - Example Q1

❖ a,b) Can be answered in 2 different approaches

  ➢ TM_0 : $|M_R$-$m_R|$ constant value greater than threshold
  ➢ TM_1 : $|M_R$-$m_R|$ is assigned a greater value than threshold iff path of robot and estimated path of human lead to collision

❖ c)

  ➢ TM_0 : Communicate planned path (Robot)
  ➢ TM_1 : Communicate predicted location of Collision

❖ d) Provide large part of the planned path

Designing interaction for understanding
- Example Q2

❖ a) Depends on choice made for Q1.a and Q1.b
  ➢ TM_0 : Infers no part of M_H
  ➢ TM_1 : Infers human's intended path
❖ b) Represent path as list of floor coordinates
❖ c) Infer human's path by extrapolating perceived motion pattern

Designing interaction for understanding
- Example Q3,Q4 & Q5

❖ Q3
  ➢ Use projected light patterns on floor as communicative action
  ➢ If collision is imminent, warning signals
❖ Q4
  ➢ Direct communication to anyone listening
❖ Q5
  ➢ Generate communicative actions all the time
  ➢ Other choice, only generate when a human in motion is close

Conclusion & Future work

❖ Model of interaction for understanding proposed

❖ Applies to cases in which human and robot use ToM to understand each other

➢ Can also be applied to simpler cases

❖ Clear separation between information to be communicated and means to communicate

❖ Implementation can be guided by addressing proposed questions

❖ A solution needs to be found for the complex case when there are more Agents in a room

➢ To whom should the Robot direct the communicative actions ?

❖ No general solution for design of understandable robots

➢ Answers to the questions are mostly application specific

➢ General answers need to be found

Citations

- ❖ 38 citations
- ❖ From which 24 in 2020
- ❖ Two interesting applications
  - ➢ Robotic System for physical training of older Adults (2020)
    - ■ Omri Avioz-sarig, Samuel Olatunji, Vardit Sarne-Fleischmann and Yael edan
  - ➢ Understandable teams of Pepper robots (2020)
    - ■ A.K. Singh and Neha Baranwal
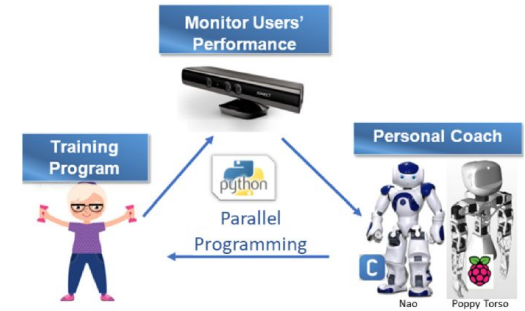
Robotic System for physical training of older Adults
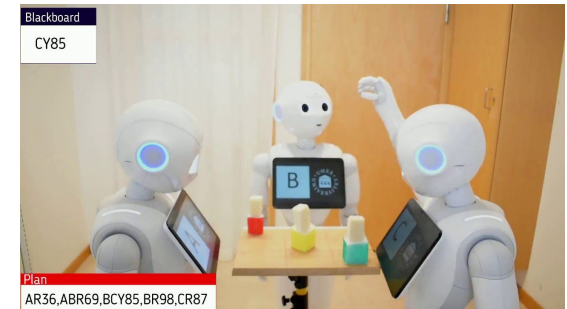
Understandable teams of Pepper Robots

Thank you for your attention !

"Even the most advanced, lifelike robots cannot reason about beliefs, desires and intentions of other agents"
- C. L. Baker, J. B. Tenenbaum (2014)