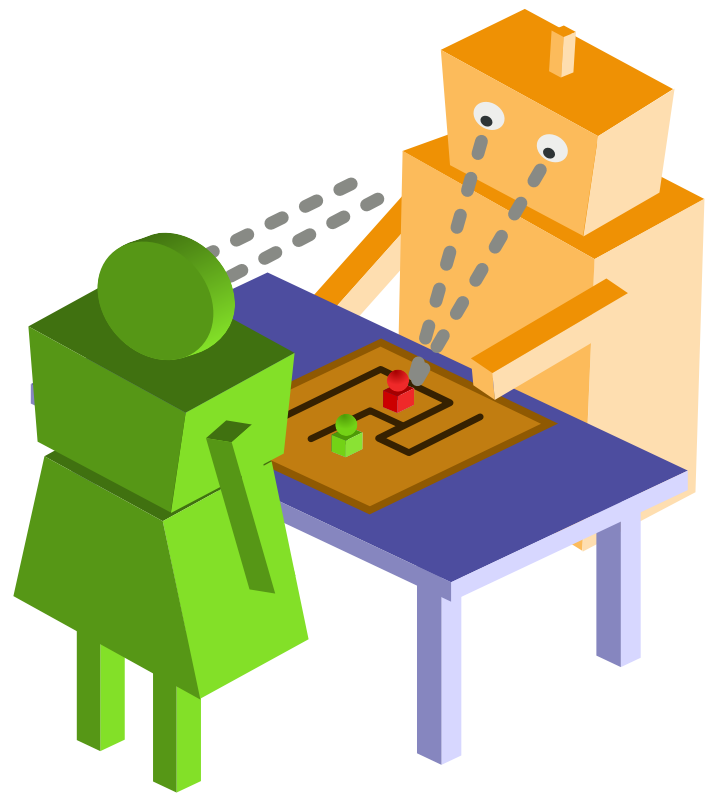
WITH PLYMOUTH UNIVERSITY

MARIE CURIE ACTIONS

Of Cognition and Social Robots

on the way to artificial social cognition in HRI

University of Edinburgh – **15 May 2017**



Séverin Lemaignan

Centre for Robotics and Neural Systems
Plymouth University

MSc in **AI for Learning technologies**
(Paris 5)

top student of the year


PhD with High Distinction @TUM

SINCE MY PHD

2013–2015 **Post-doc at EPFL**

- built up a leading group in child-robot interaction

SINCE MY PHD

2013–2015 **Post-doc at EPFL**

- built up a leading group in child-robot interaction
- went up to 3 PhD + 3 MSc students

SINCE MY PHD

2013–2015 **Post-doc at EPFL**

- built up a leading group in child-robot interaction
- went up to 3 PhD + 3 MSc students
- joint supervisions with Portugal, France

SINCE MY PHD

2013–2015 **Post-doc at EPFL**

- built up a leading group in child-robot interaction
- went up to 3 PhD + 3 MSc students
- joint supervisions with Portugal, France
- 2 high-visibility projects with large media coverage

SINCE MY PHD

2013–2015 **Post-doc at EPFL**

- built up a leading group in child-robot interaction
- went up to 3 PhD + 3 MSc students
- joint supervisions with Portugal, France
- 2 high-visibility projects with large media coverage
- 8 papers in 5 years at high-profile HRI conference, 3 awards

SINCE MY PHD

2013–2015 **Post-doc at EPFL**

- built up a leading group in child-robot interaction
- went up to 3 PhD + 3 MSc students
- joint supervisions with Portugal, France
- 2 high-visibility projects with large media coverage
- 8 papers in 5 years at high-profile HRI conference, 3 awards

2015–2017 **EU Marie Skłodowska-Curie fellow** at Plymouth University on **Theory of Mind** & **Social cognition for robots**

SINCE MY PHD

2013–2015 **Post-doc at EPFL**

- built up a leading group in child-robot interaction
- went up to 3 PhD + 3 MSc students
- joint supervisions with Portugal, France
- 2 high-visibility projects with large media coverage
- 8 papers in 5 years at high-profile HRI conference, 3 awards

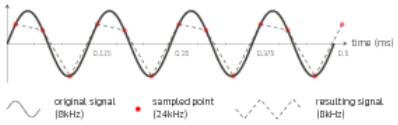
2015–2017 **EU Marie Skłodowska-Curie fellow** at Plymouth University on **Theory of Mind** & **Social cognition for robots**

60+ publications; **1000+ citations**; **H-index=15**

SAMPLING RATE

Nyquist theorem

- The sampling rate has to be at least **twice as high** as the fastest changes. If not, you are going to miss relevant information.



- e.g. if sound signal changes at 3kHz, you have to sample at at least 6kHz to not miss anything of the signal.

CORRECT STEP EQUATIONS (MEASUREMENT UPDATE)

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \cdot \mathbf{H}^\mathsf{T} \cdot (\mathbf{H} \cdot \mathbf{P}_{k|k-1} \cdot \mathbf{H}^\mathsf{T} + \mathbf{R})^{-1}$$

The **Kalman gain**, this needs to calculated first

Sensor noise

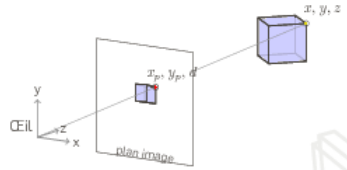
$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \cdot (\mathbf{z}_k - \mathbf{H} \cdot \hat{\mathbf{x}}_{k|k-1})$$

The *a posteriori* estimated state

$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k \cdot \mathbf{H}) \cdot \mathbf{P}_{k|k-1}$$

The *a posteriori* estimated covariance of our state

EN 3D

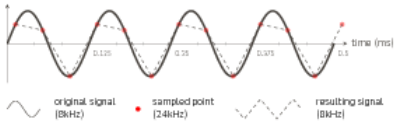


$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{d} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ \frac{z}{d} \end{bmatrix} \simeq \begin{bmatrix} \frac{dx}{z} \\ \frac{dy}{z} \\ d \\ 1 \end{bmatrix}$$

SAMPLING RATE

Nyquist theorem

- The sampling rate has to be at least **twice as high** as the fastest changes. If not, you are going to miss relevant information.



- e.g. if sound signal changes at 3kHz, you have to sample at at least 6kHz to not miss anything of the signal.

CORRECT STEP EQUATIONS (MEASUREMENT UPDATE)

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \cdot \mathbf{H}^\mathsf{T} \cdot (\mathbf{H} \cdot \mathbf{P}_{k|k-1} \cdot \mathbf{H}^\mathsf{T} + \mathbf{R})^{-1}$$

The **Kalman gain**, this needs to calculated first

Sensor noise

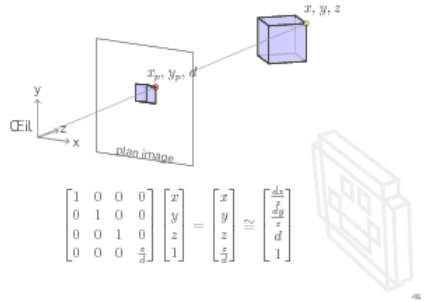
$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \cdot (\mathbf{z}_k - \mathbf{H} \cdot \hat{\mathbf{x}}_{k|k-1})$$

The *a posteriori* estimated state

$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k \cdot \mathbf{H}) \cdot \mathbf{P}_{k|k-1}$$

The *a posteriori* estimated covariance of our state

EN 3D

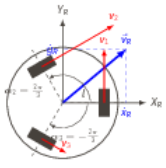


$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ \frac{z}{d} \end{bmatrix} \simeq \begin{bmatrix} \frac{dx}{z} \\ \frac{dy}{z} \\ d \\ 1 \end{bmatrix}$$

KINEMATICS OF AN OMNIDIRECTIONAL ROBOT: ROBOT FRAME

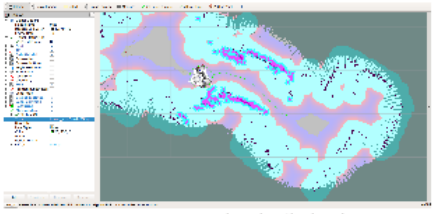
$$\begin{aligned} v_i &= \dot{\varphi}_i \cdot r \\ &= -\dot{x}_R \cdot \sin(\alpha_i) + \dot{y}_R \cdot \cos(\alpha_i) + l \cdot \dot{\theta} \end{aligned}$$

$$\begin{aligned} \dot{\varphi} \cdot r &= \begin{bmatrix} \dot{\varphi}_1 \\ \dot{\varphi}_2 \\ \dot{\varphi}_3 \end{bmatrix} \cdot r \\ &= \begin{bmatrix} -\sin\alpha_1 & \cos\alpha_1 & l \\ -\sin\alpha_2 & \cos\alpha_2 & l \\ -\sin\alpha_3 & \cos\alpha_3 & l \end{bmatrix} \cdot \begin{bmatrix} \dot{x}_R \\ \dot{y}_R \\ \dot{\theta} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & l \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & l \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & l \end{bmatrix} \cdot \begin{bmatrix} \dot{x}_R \\ \dot{y}_R \\ \dot{\theta} \end{bmatrix} \end{aligned}$$



⇒ we can determine the wheels angular velocities to obtain a specific robot velocity.

GLOBAL VS LOCAL PATH PLANNING



Global path planning

- long distances (i.e. large map, slower calculations)
- static environment

Local path planning

- short horizon
- dynamic environment (use of sensors)

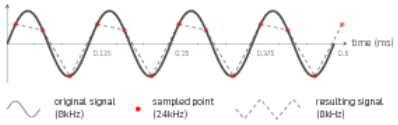
GRASSFIRE ALGORITHM

	A	B	C	D	E	F	G	H	I	J	K	L
1	13	12	11	10	9	8	7	8	9	10	11	10
2	12	11	10	9	8	7	6					9
3	11	10	9	8	7	6	5	4	5	6	7	8
4	10	9	8	7	6	5	4	3	4	5	6	7
5	11	S	9			4	3	2	3	4	5	6
6	12	11	10	11		3	2	1	2		6	7
7	11	10	11	12		2	1	G	1		7	8
8	10	9				3	2	1	2		8	9
9	9	8	7	6	5	4	3	2	3		9	10
10	10	9	8	7	6	5	4	3	4		10	11

SAMPLING RATE

Nyquist theorem

- The sampling rate has to be at least **twice as high** as the fastest changes. If not, you are going to miss relevant information.



- e.g. if sound signal changes at 3kHz, you have to sample at at least 6kHz to not miss anything of the signal.

CORRECT STEP EQUATIONS (MEASUREMENT UPDATE)

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \cdot \mathbf{H}^\mathsf{T} \cdot (\mathbf{H} \cdot \mathbf{P}_{k|k-1} \cdot \mathbf{H}^\mathsf{T} + \mathbf{R})^{-1}$$

The **Kalman gain**, this needs to calculated first

Sensor noise

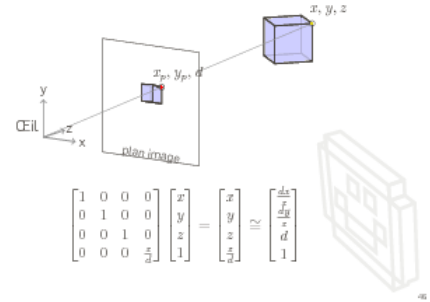
$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \cdot (\mathbf{z}_k - \mathbf{H} \cdot \hat{\mathbf{x}}_{k|k-1})$$

The *a posteriori* estimated state

$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k \cdot \mathbf{H}) \cdot \mathbf{P}_{k|k-1}$$

The *a posteriori* estimated covariance of our state

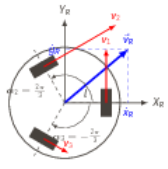
EN 3D



$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ \frac{z}{d} \end{bmatrix} \simeq \begin{bmatrix} \frac{dx}{z} \\ \frac{dy}{z} \\ d \\ 1 \end{bmatrix}$$

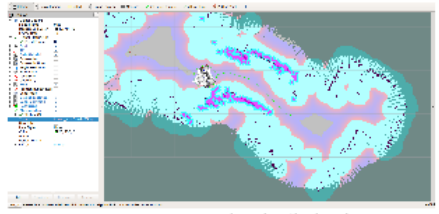
KINEMATICS OF AN OMNIDIRECTIONAL ROBOT: ROBOT FRAME

$$\begin{aligned} v_i &= \dot{\varphi}_i \cdot r \\ &= -\dot{x}_R \cdot \sin(\alpha_i) + \dot{y}_R \cdot \cos(\alpha_i) + l \cdot \dot{\theta} \\ \dot{\varphi} \cdot r &= \begin{bmatrix} \dot{\varphi}_1 \\ \dot{\varphi}_2 \\ \dot{\varphi}_3 \end{bmatrix} \cdot r \\ &= \begin{bmatrix} -\sin\alpha_1 & \cos\alpha_1 & l \\ -\sin\alpha_2 & \cos\alpha_2 & l \\ -\sin\alpha_3 & \cos\alpha_3 & l \end{bmatrix} \cdot \begin{bmatrix} \dot{x}_R \\ \dot{y}_R \\ \dot{\theta} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & l \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & l \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & l \end{bmatrix} \cdot \begin{bmatrix} \dot{x}_R \\ \dot{y}_R \\ \dot{\theta} \end{bmatrix} \end{aligned}$$



⇒ we can determine the wheels angular velocities to obtain a specific robot velocity.

GLOBAL VS LOCAL PATH PLANNING



Global path planning

- long distances (i.e. large map, slower calculations)
- static environment

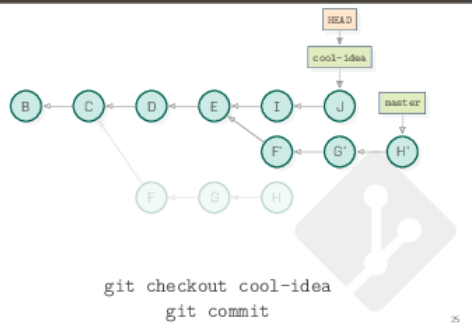
Local path planning

- short horizon
- dynamic environment (use of sensors)

GRASSFIRE ALGORITHM

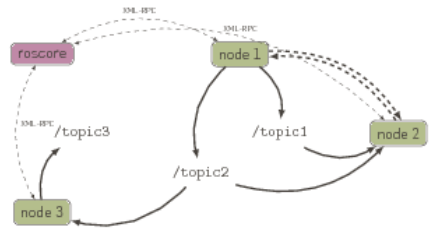
	A	B	C	D	E	F	G	H	I	J	K	L
1	13	12	11	10	9	8	7	8	9	10	11	10
2	12	11	10	9	8	7	6					9
3	11	10	9	8	7	6	5	4	5	6	7	8
4	10	9	8	7	6	5	4	3	4	5	6	7
5	11	S	9			4	3	2	3	4	5	6
6	12	11	10	11		3	2	1	2		6	7
7	11	10	11	12		2	1	G	1		7	8
8	10	9				3	2	1	2		8	9
9	9	8	7	6	5	4	3	2	3		9	10
10	10	9	8	7	6	5	4	3	4		10	11

REBASING BRANCHES




```
git checkout cool-idea
git commit
```


MIDDLEWARES' CORE PRINCIPLE: TALKING NODES



`ROS_MASTER_URI=http://<host>:<port>`

EVENT-ORIENTED PROGRAMMING

Event-oriented programming is a possible way of implementing a behavioural control paradigm:

```
def on_dummy(robot):
    robot.look_at_dummy()
    robot.blink()
    sleep = robot.fall_asleep()
    robot.lightbar(RAINBOW).wait()
    sleep.wait()

def on_dummy_removed(robot):
    robot.light_bar(colors.rand())
    robot.wakeup().wait()
    robot.move(0.4, v = 0.8).wait()
    robot.idle().wait()
```

```
def on_bumper(robot):
    pulse = robot.pulse_row(0)
    while abs(robot.state.v) > 0.01:
        robot.sleep(0.2)
    pulse.cancel()

def on_toy(robot):
    robot.playsound(SOUNDS["toy_in"])
    robot.lightbar(RAINBOW).wait()
```

CS EXPERTISE

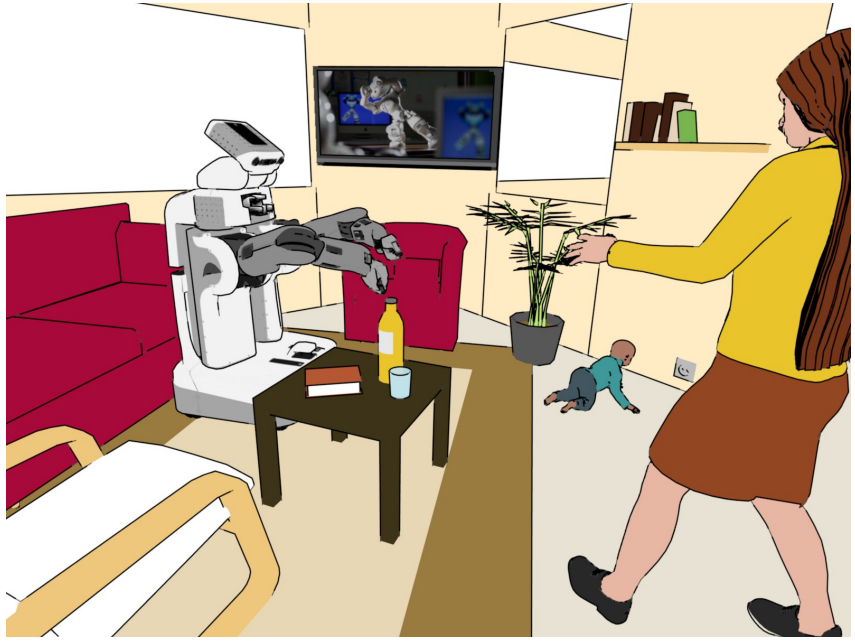
- **advanced C++** (C++11/14, STL, meta-programming)
- **advanced Python** (meta-programming)
- **software engineering** (GIT expertise, coding best practises)
- logic programming (Prolog, ontologies)
- distributed systems (middlewares)
- computer vision & 3D rendering (OpenGL)
- algorithms & data structures

120+ open-source repositories on GitHub; **contributor to major open-source projects** like **ROS**, **OpenCV**

Situated dialogue effectively evidences the challenges

How can the robot make sense of and act upon a command like:

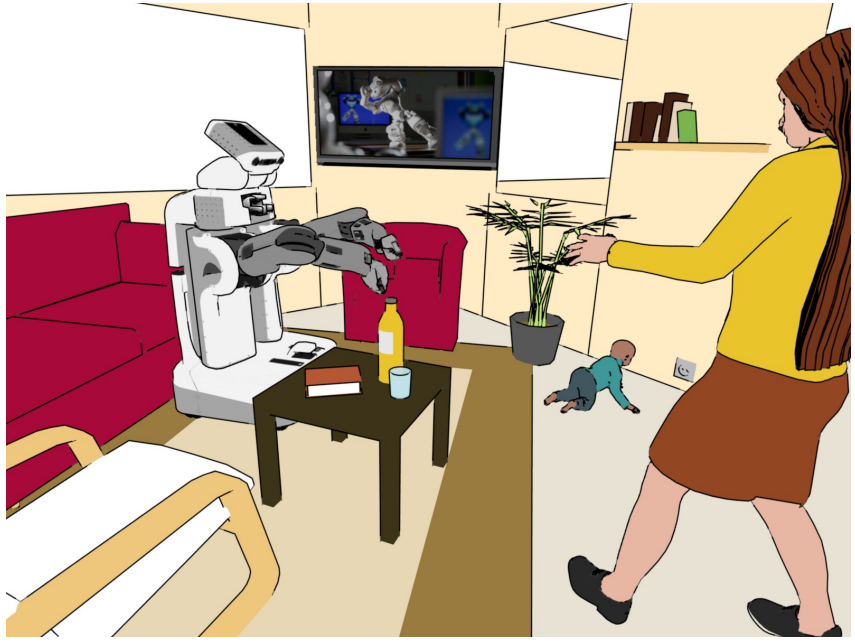
“Can you give me that book?”



Situated dialogue effectively evidences the challenges

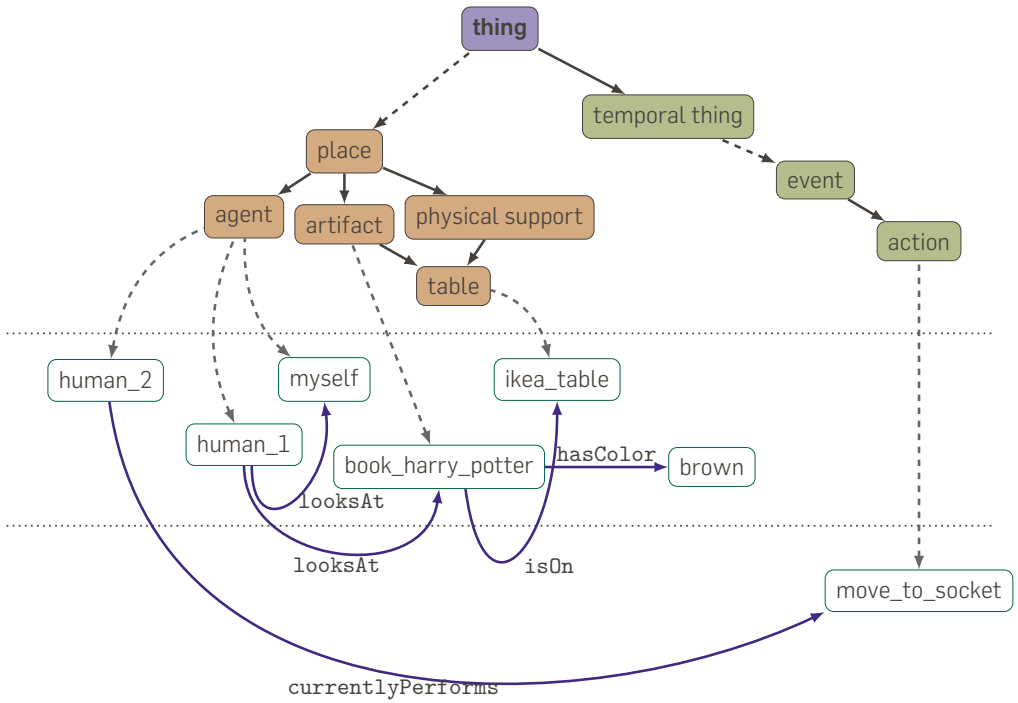
How can the robot make sense of and act upon a command like:

“Can you give me that book?”



My PhD: a symbolic approach to this problem

MULTI-MODAL SYMBOLIC SITUATION ASSESSMENT



[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

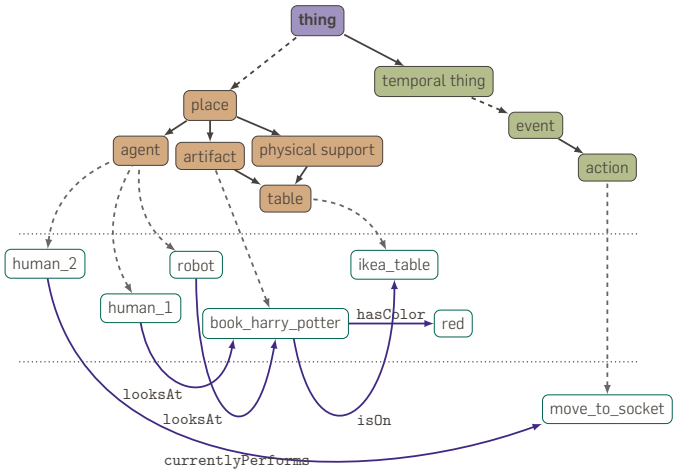

LAAS-CNRS

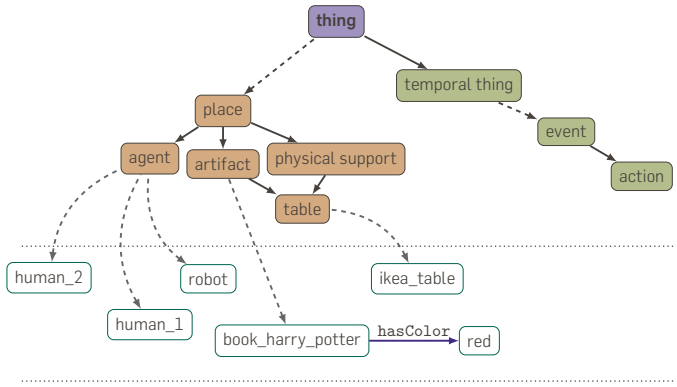
What if I ask for the object in the box, but the robot has moved it somewhere else?

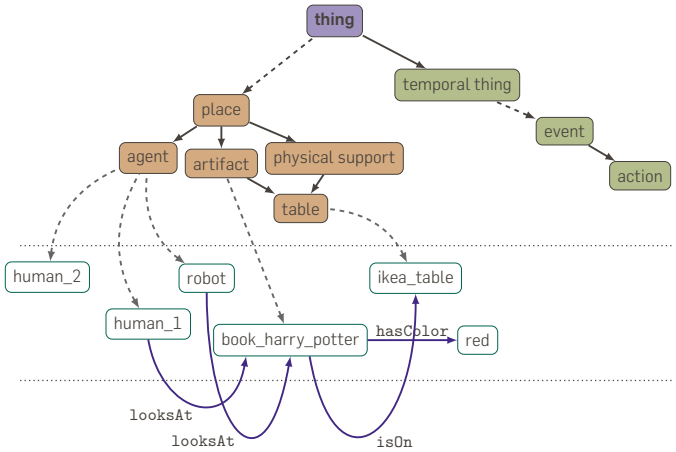
What if I ask for the object in the box, but the robot has moved it somewhere else?

False-belief situation

PARALLEL MODELS: TOWARDS A THEORY OF MIND



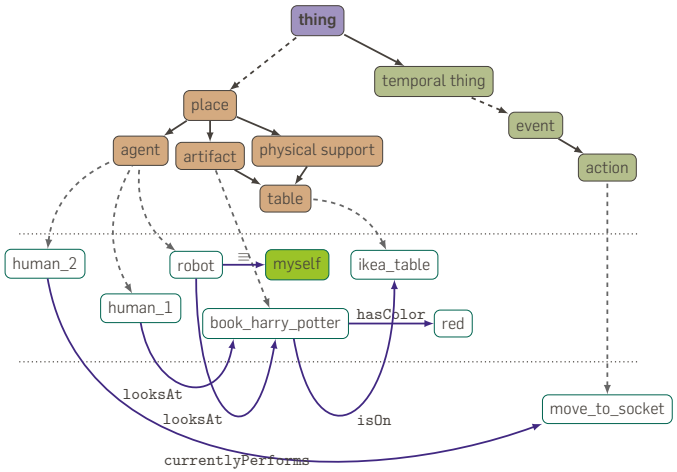


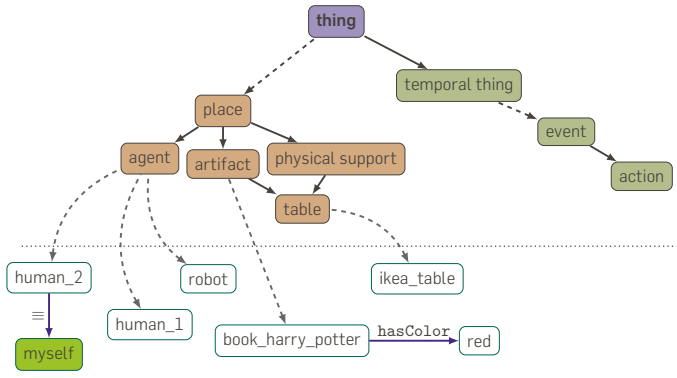


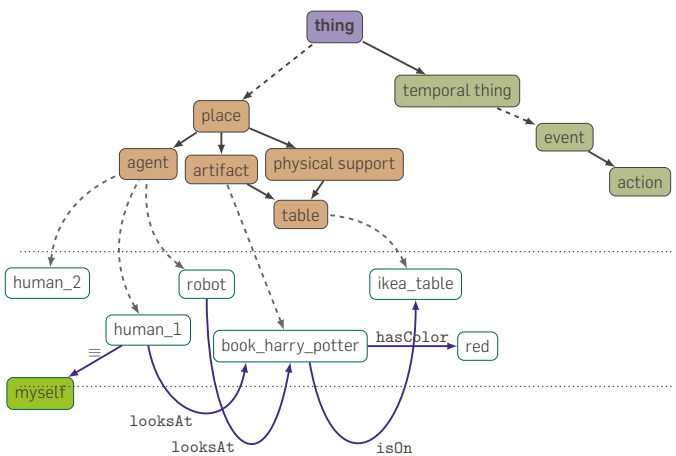
...

[Ros, Lemaignan et al., **Which One? Grounding the Referent Based on Efficient HRI**, ROMAN 2010]

PARALLEL MODELS: TOWARDS A THEORY OF MIND





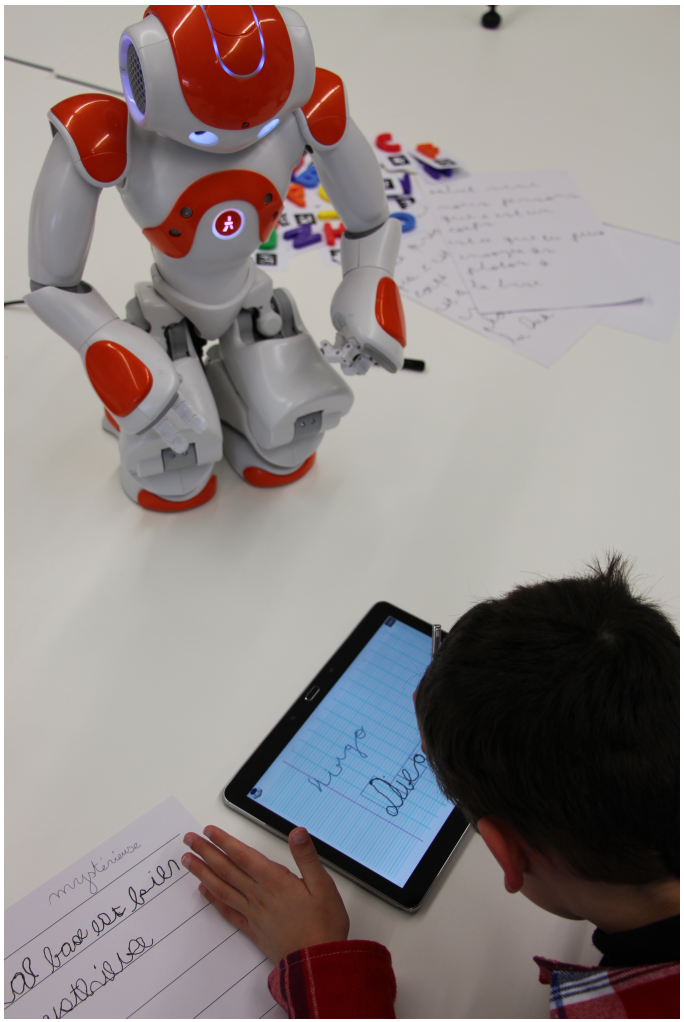


...

[Ros, Lemaignan et al., **Which One? Grounding the Referent Based on Efficient HRI**, ROMAN 2010]


LAAS-CNRS

THE ROBOT AS A SOCIAL AGENT



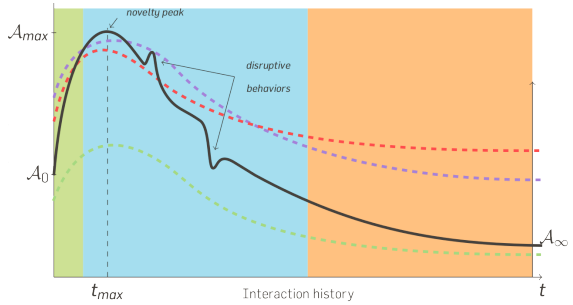
The robot as a cognitive agent is key here

- Protégé effect
- metacognition

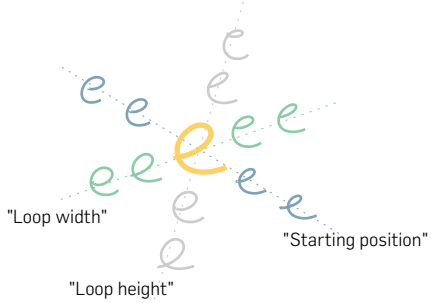
[Lemaignan et al. **Learning by Teaching a Robot: The Case of Handwriting** – Robotics and Automation Magazine 2016]

OTHER EXPERIMENTAL WORK

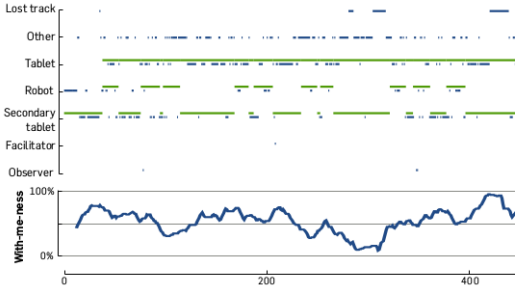
Understanding social interaction

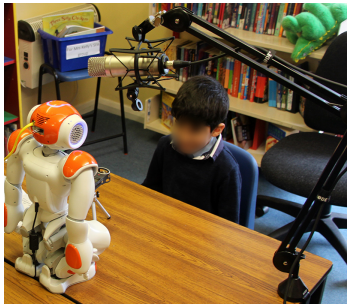


Robots for Learning



Methodology





Expectations
How do you imagine a robot? *What could it look like?* *Have you ever seen a robot before?*
Impression
When you first saw R, what did you think? *Is R a robot? How do you know?* *Did you expect R would come over to you when you call it?* *What happened when you put the domino in the box?*
Ascribe intention
Do you think R could go out the door all by itself? *Does R always obey / come over to you?* *Could R do something silly?* *Why did R not come over to you when you called it?*

Ascribe perceptual capabilities
Here is a domino. Do you think R can see it? *When I say "Hello R!", do you think R can hear it?*
Ascribe emotional state
Does R have feelings? Can R be happy or sad sometimes?
Social acceptance
Do you like R? Why (not)? *What do you (not) like about it?* *Would you like to have R at home?*
Companionship
Could R be your friend? Why (not)?
Ascribe moral standing
Assume you go on a holiday for two weeks. Is it alright to leave R alone at home? Why (not)?

How to push back the boundaries of social robotics?

open, underspecified situations
complex social dynamics
rich semantics
interplay of socio-cognitive functions

SURFACE FUNCTIONS FOR SOCIAL COGNITION

[Baxter, Lemaignan, Trafton **Workshop on Cognitive Architectures for Social HRI** – HRI 2016]

yet...

THEORETICAL FRAMEWORK: STAGES OF PLAY

In developmental psychology, Parten's **stages of play**:

1.  **Solitary (independent) play**
2.  **Onlooker play**
3.  **Parallel play**
4.  **Associative play**
5.  **Cooperative play**

[Parten, **Social participation among preschool children** Journal of Abnormal and Social Psychology 1932]

Free-play sandbox

A FRAMEWORK FOR SOCIO-COGNITIVE INVESTIGATION

- **Task-independent social dynamics**

  interaction flow, adaptation, social patterns...

- **Situation awareness**

  what happens? who does what? why?

  → mind modelling

KEY OUTCOMES FOR HRI

- **online segmentation and understanding of the interaction flow**
- **surface and deep behavioural alignment**
- **enable the robot to merge into natural social dynamics**
  - 'natural' (i.e. emergent) turn-taking
  - 'natural' protodeclarative pointing

KEY OUTCOMES FOR HRI

- **online segmentation and understanding of the interaction flow**
- **surface and deep behavioural alignment**
- **enable the robot to merge into natural social dynamics**
  - 'natural' (i.e. emergent) turn-taking
  - 'natural' protodeclarative pointing
- artifical **Theory of Mind beyond purely perceptual models**
- **emergence of Parten's stages of play?**

MY CURRENT VISION

(Deep) learning of socio-cognitive human-robot interactions

[Graziano **Consciousness and the Social Brain** – 2013]
[Jaeger **Controlling recurrent neural networks by conceptors** – 2014]

MY CURRENT VISION

(Deep) learning of socio-cognitive human-robot interactions

Working hypothesis: **Sociality emerges from interaction**

[Graziano **Consciousness and the Social Brain** – 2013]
[Jaeger **Controlling recurrent neural networks by conceptors** – 2014]

MY CURRENT VISION

(Deep) learning of socio-cognitive human-robot interactions

Working hypothesis: **Sociality emerges from interaction**

- Instrumental role of **attention**
- Unsupervised **recurrent neural networks** to model others' minds → a **connectionist theory of mind**

[Graziano **Consciousness and the Social Brain** – 2013]
[Jaeger **Controlling recurrent neural networks by conceptors** – 2014]

MY CURRENT VISION

(Deep) learning of socio-cognitive human-robot interactions

Working hypothesis: **Sociality emerges from interaction**

- Instrumental role of **attention**
- Unsupervised **recurrent neural networks** to model others' minds → a **connectionist theory of mind**

...towards a principled model of social cognition?

[Graziano **Consciousness and the Social Brain** – 2013]
[Jaeger **Controlling recurrent neural networks by conceptors** – 2014]

Thank you!

SUPPLEMENTARY MATERIAL

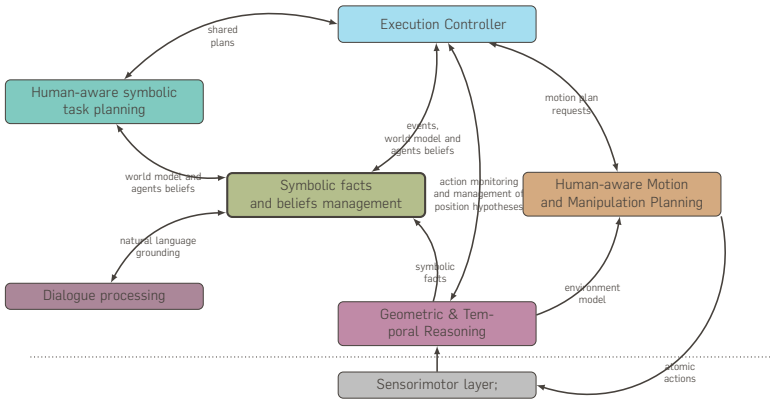
RESEARCH APPROACH

RESEARCH APPROACH

In parallel

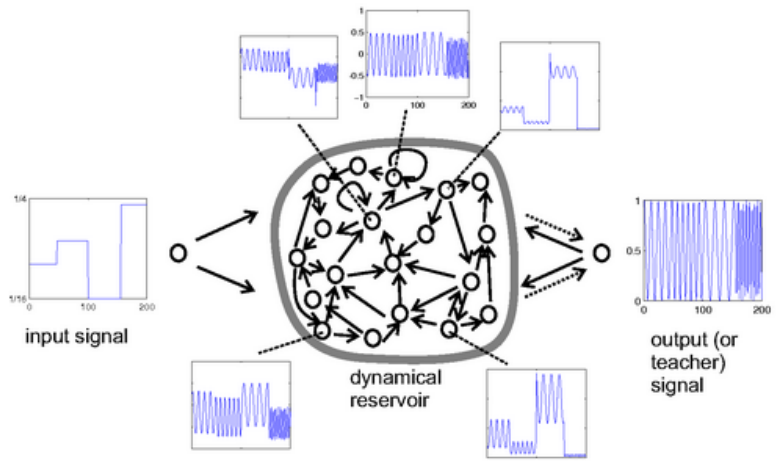
Practical

The “Edinburgh Architecture for Socio-Cognitive Robots”



Basic research

Emergent Social Cognition



→ RNNs; deep learning

SCAFFOLDING THE RESEARCH

Create a **KidsLab**, within the Robotarium

- experimental space suited for 4 to 13 years old
- fully equipped (recording, tracking, **tactile floor**)
- in connection with Wee Science lab and the RabLab

SCAFFOLDING THE RESEARCH

Create a **KidsLab**, within the Robotarium

- experimental space suited for 4 to 13 years old
- fully equipped (recording, tracking, **tactile floor**)
- in connection with Wee Science lab and the RabLab

Establish collaboration with the (new) Royal Hospital for Sick Children

- field experiments
- research with real-world impact on children

FUNDING

Short term

- **EPSRC First Grant**

  up to £125,000
  funding for the **KidsLab** project
  equipment + one post-doc

- **EPSRC Early Career Fellowship**

  max 5 years
  focus on **Emergence of artificial social cognition**
  possible partnerships: developmental robotics (e.g. INRIA Flowers), cognitive neurosciences (e.g. Graziano lab), learning technologies (e.g. EPFL CHILI Lab)

FUNDING

Within 2-3 years

- **Co-Investigator on a H2020 project**

  close contacts with IST Lisbon, LAAS-CNRS, EPFL, INRIA Bordeaux, Bristol BRL

- Submit a **EPSRC Standard Research** proposal

  $\rightarrow$ reaching independence
  focus on socio-cognitive architecture + long-term field deployment
  multi-proposal submission, possibly with Lincoln U. (expertise on Cognitive Architectures)

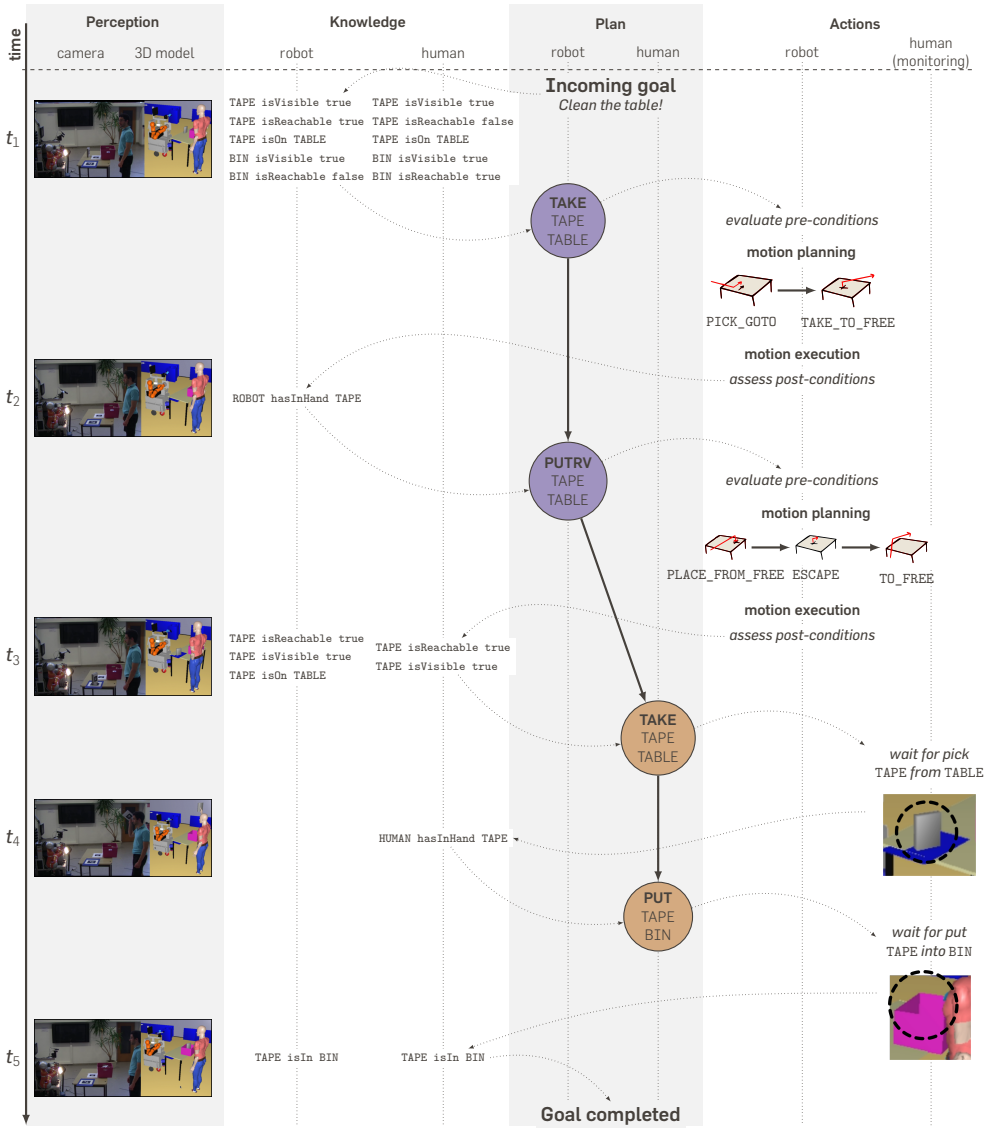
- **ERC Starting Grant**

  up to €1.5m
  within 7 years of PhD completion, i.e. before 2019
  project on **Emergence of artificial social cognition**

PERFORMING IN HUMAN ENVIRONMENTS

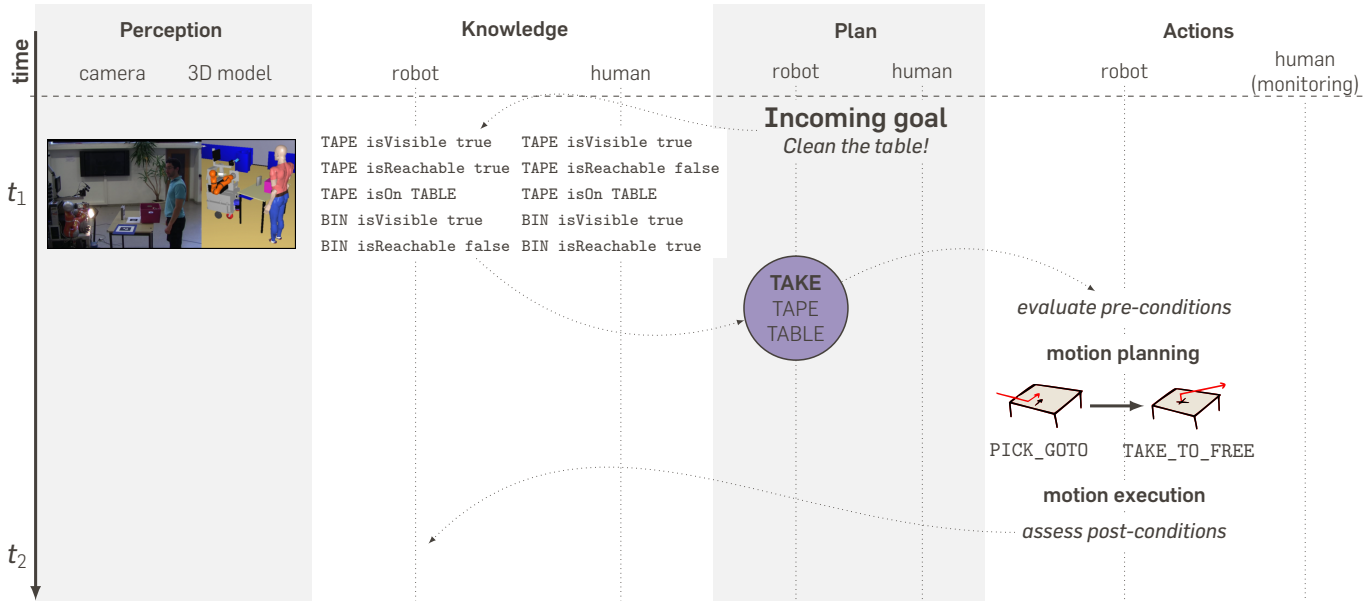

LAAS-CNRS

“CLEANING THE TABLE”…



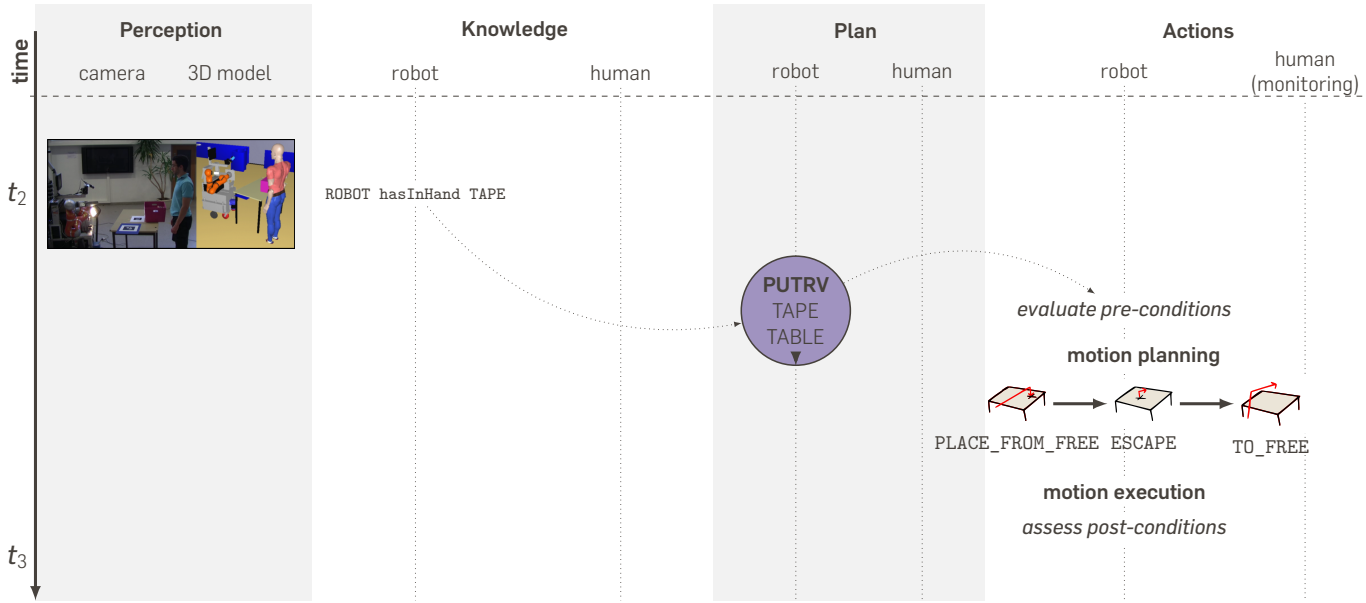
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

“CLEANING THE TABLE”...



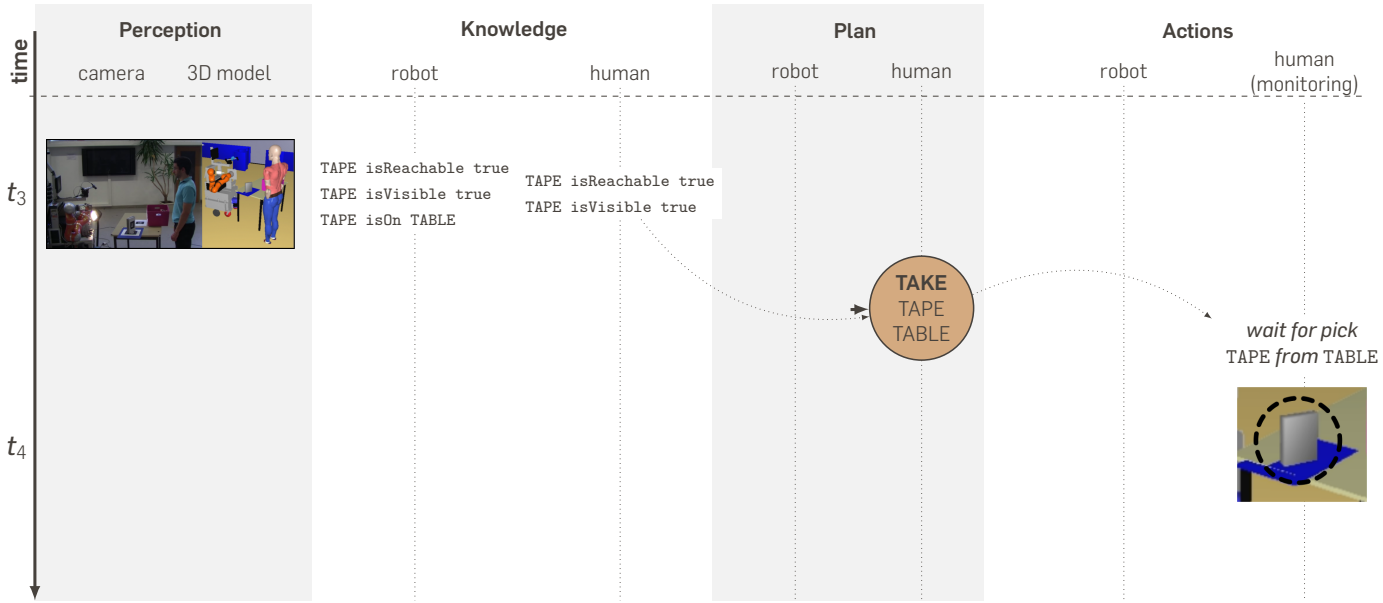
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

“CLEANING THE TABLE”...



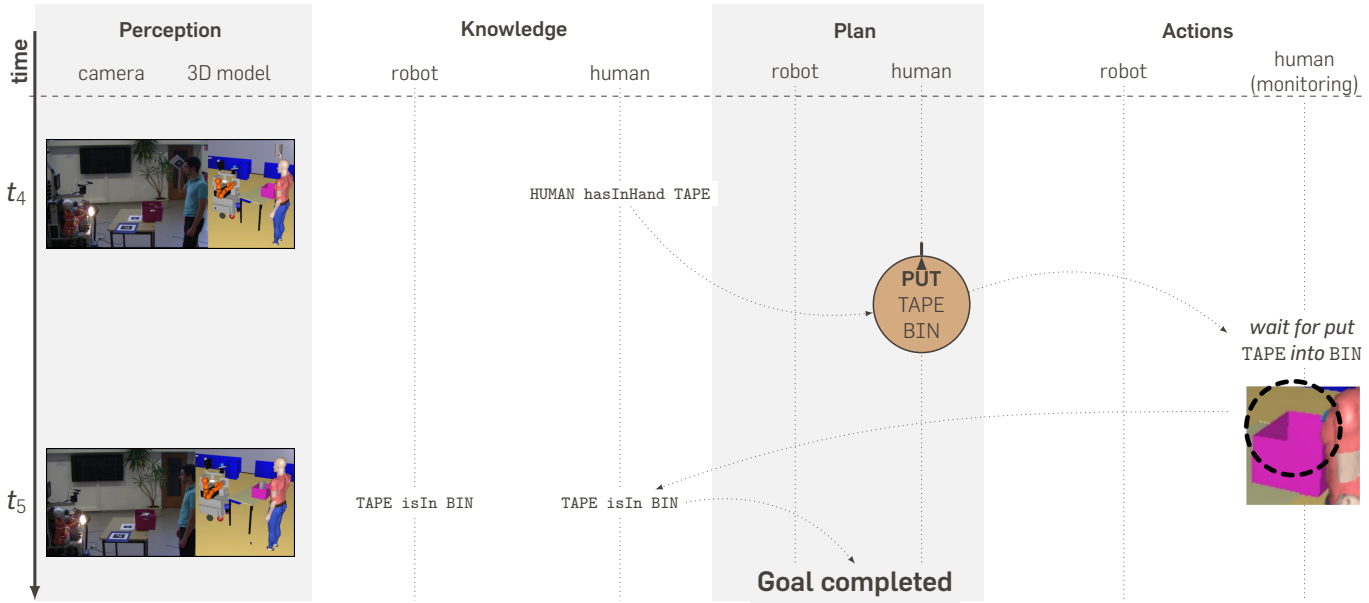
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

“CLEANING THE TABLE”...



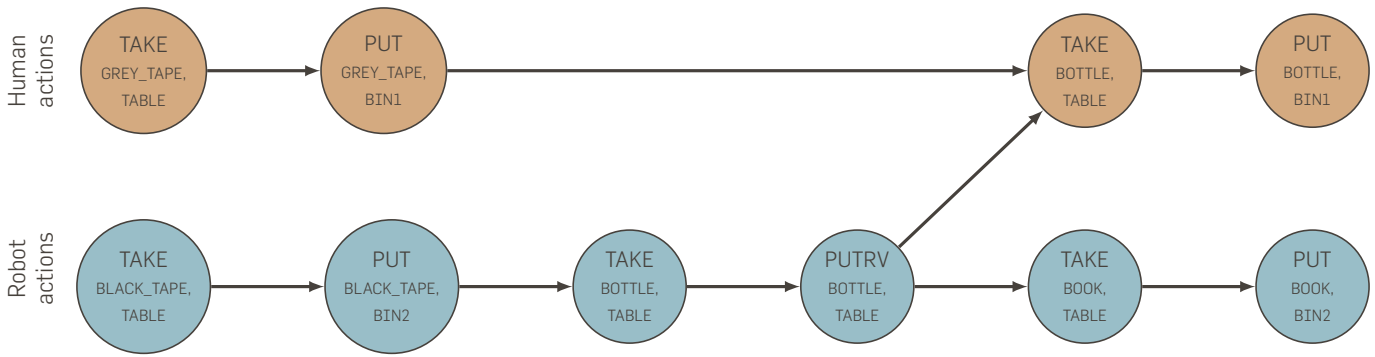
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

“CLEANING THE TABLE”...

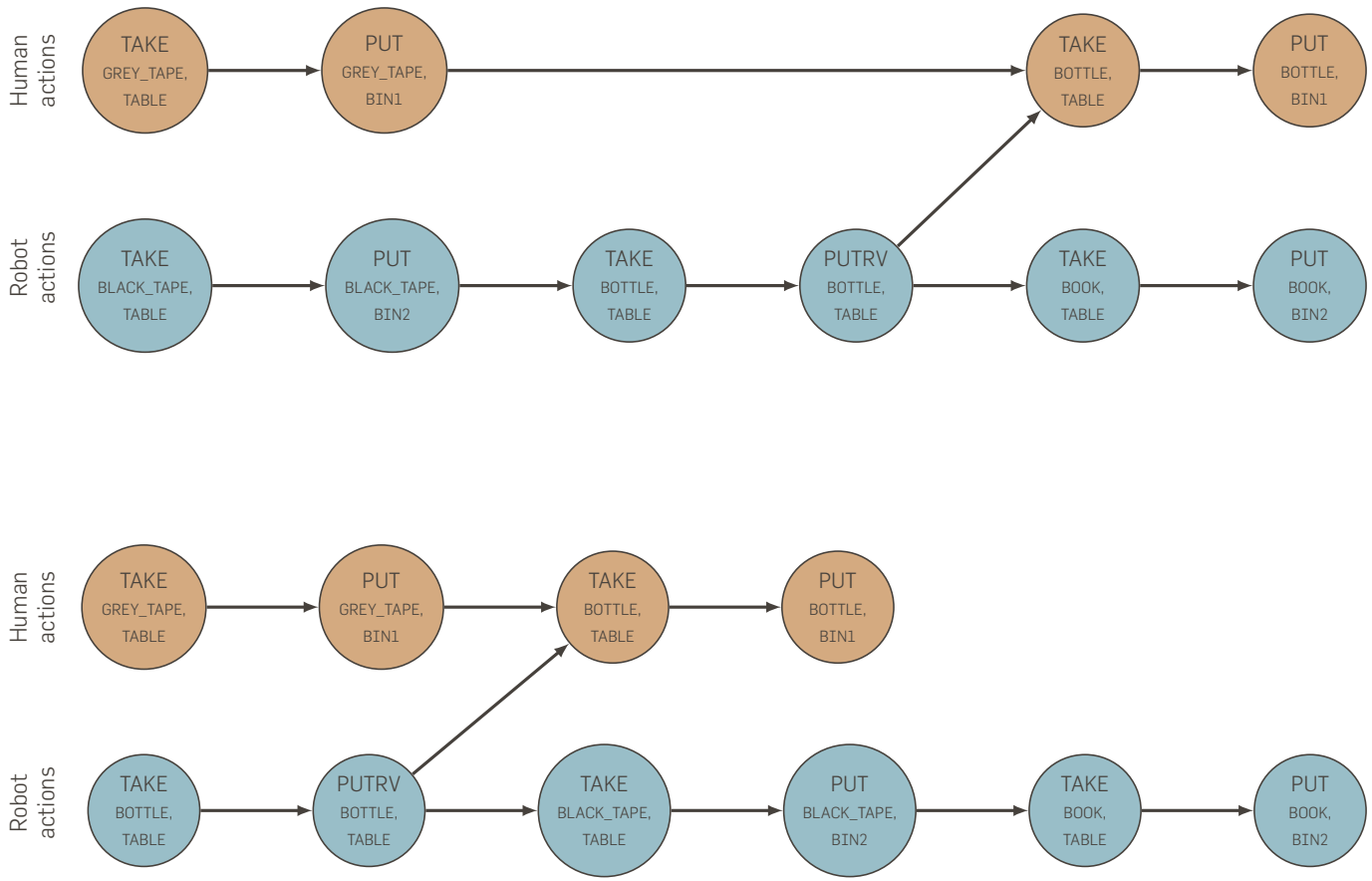


[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

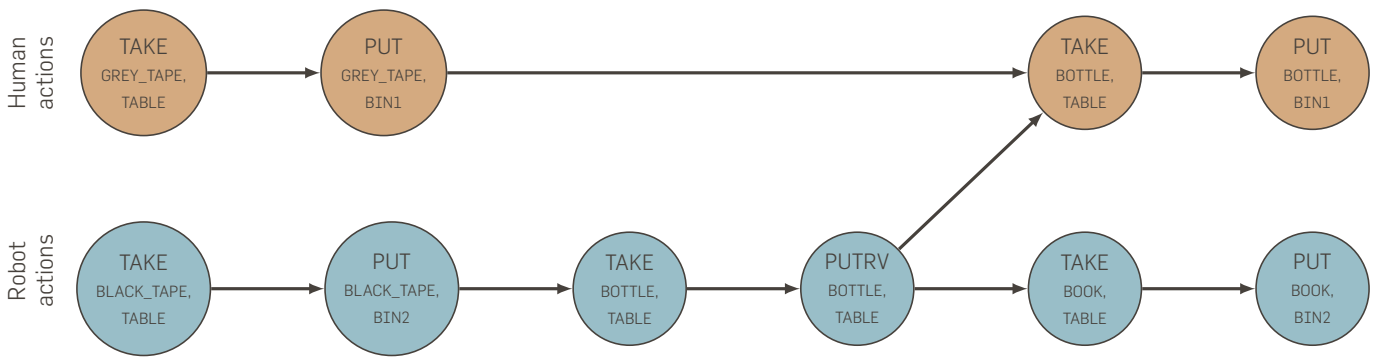
PLANNING FOR THE HUMAN



PLANNING FOR THE HUMAN



PLANNING FOR THE HUMAN




PR2
PR2

roboscopie

A Theatre Performance for a Robot and a Human

HRI 2012

```
amit_give
arms_against_torso
attachobject
basicgive
basicgrab
basictake
basket
cancel_follow
cancel_track
cancel
carry
close_gripper
configure_grippers
detect_and_grab
detect
disabledevileye
display
dock
enabledevileye
extractpose
follow
glance_to
goto
grab_gripper
grab
gym
handover
handsup_folded
handsup_folded2
handsup_folded3
handsup
hide
idle
init
larm_swinging
lock_object
look_at_ros
look_at_xyz
look_at
looksat
manipose
move_head
movearm
moveclose
open_gripper
pick
place_agent
place_object
pointsat
put_accessible
put
rarm_swinging
release_gripper
release
restpose
rotate
satisfied
say
setpose
settorso
setup_scenario
show
slow_arms_swinging
sorry
speed_arms_swinging
sweep_look
sweep
switch_cameras
take
track_human
track
translate
tuckedpose
unlock_object
wait
waypoints
```

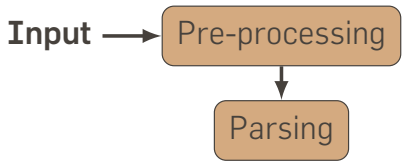
PYROBOTS

```python
from robots import GenericRobot
from robots.concurrency import action, ActionCancelled
from robots.resources import Resource, lock

class MyRobot(GenericRobot):
    # ... state + lowlevel action

WHEELS = Resource("wheels")

@lock(WHEELS)
@action
def move_forward(robot):
    target = [1.0, 0., 0., "base_link"]

    try:
        robot.goto(target)

        while(robot.dist_to(target) > 0.1):
            robot.sleep(0.5)

    except ActionCancelled:
        robot.stop()
```

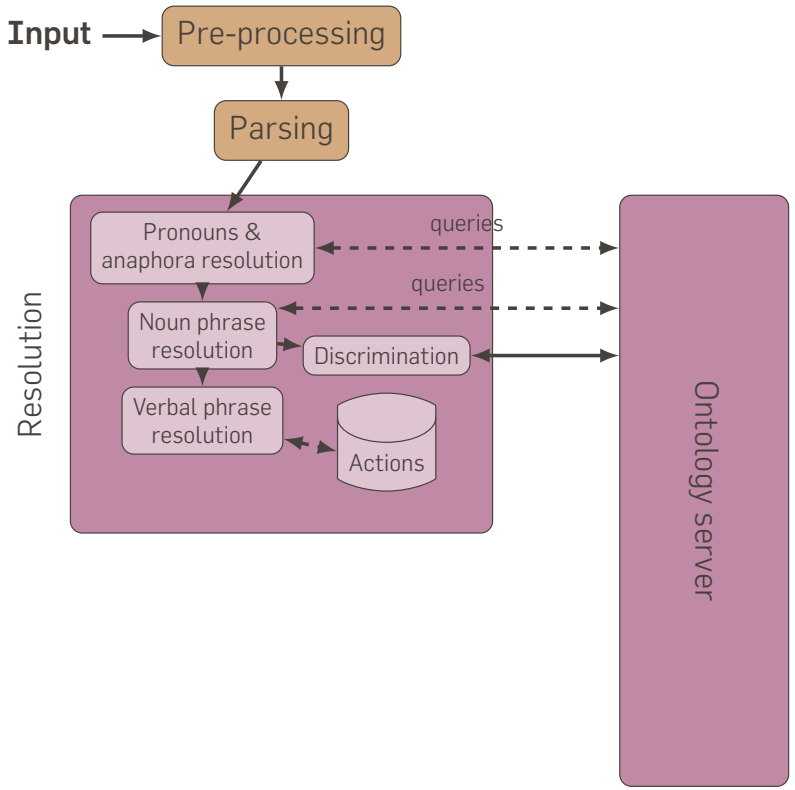
```
with MyRobot() as robot:

    robot.whenever("my_bumper", True).do(move_forward)

    try:
        while True:
            time.sleep(0.5)
    except KeyboardInterrupt:
        pass
```

DIALOGUE GROUNDING

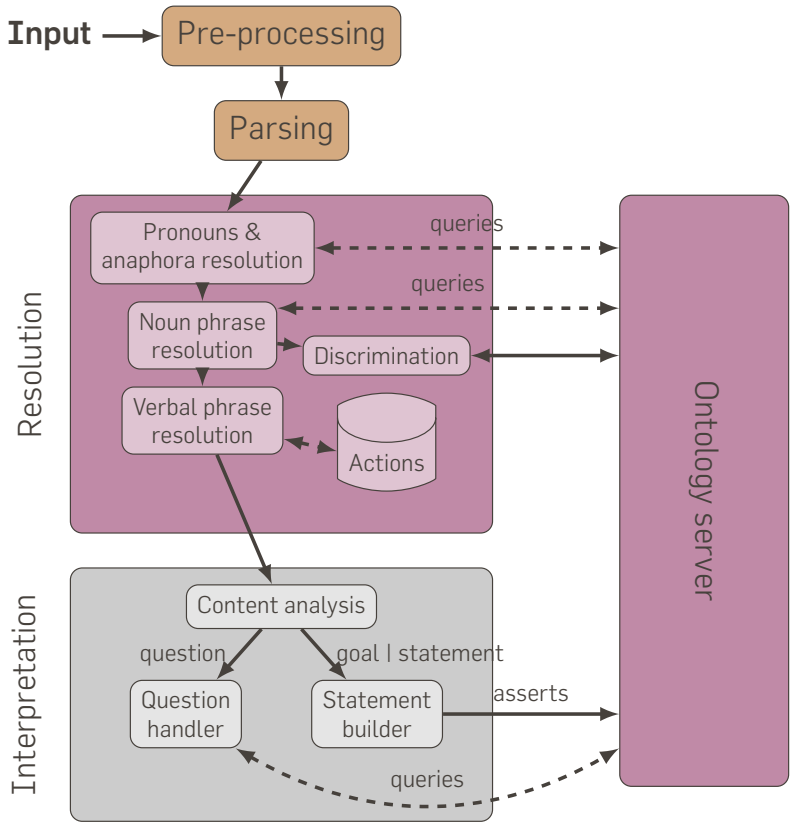
DIALOGUE GROUNDING



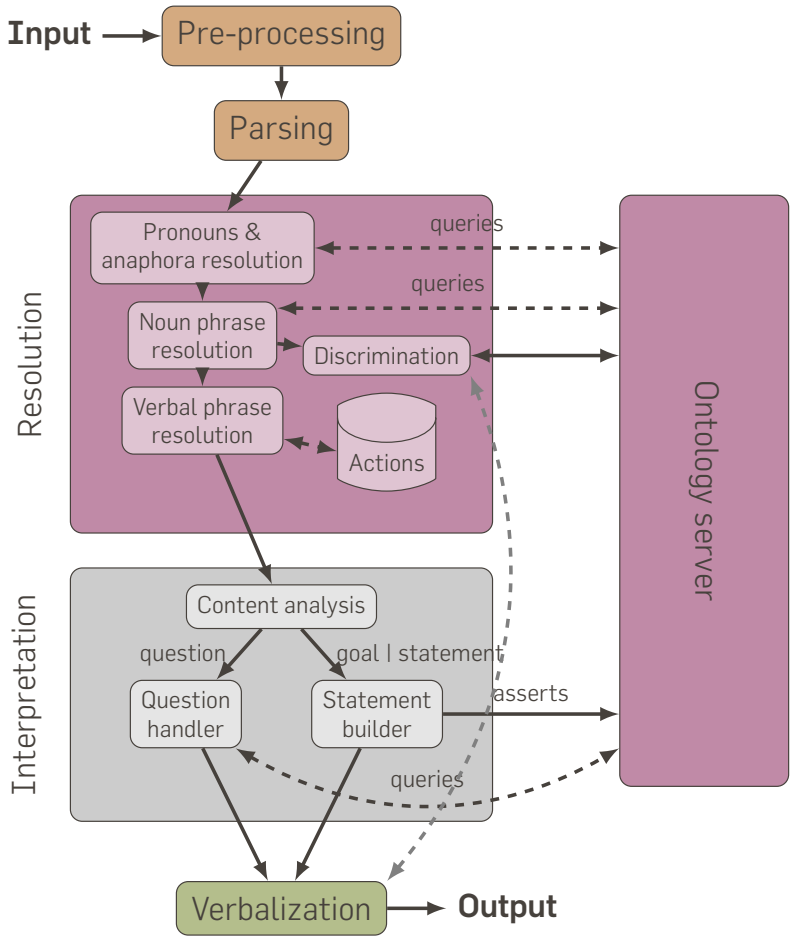
DIALOGUE GROUNDING



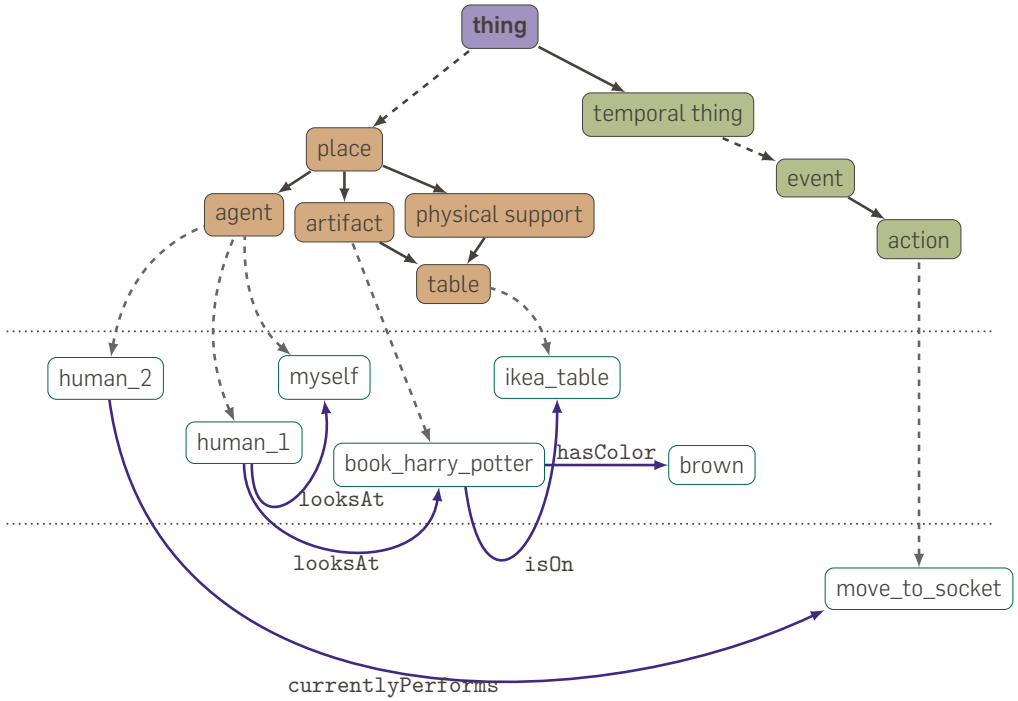
DIALOGUE GROUNDING



DIALOGUE GROUNDING



MULTI-MODAL SYMBOLIC SITUATION ASSESSMENT



DIALOGUE GROUNDING

"Give me the book on the table"

◂ Supplementary material

DIALOGUE GROUNDING

"Give me the book on the table"

$\downarrow$

$\texttt{me} \rightarrow \texttt{human\_1}$

$\text{find}(\texttt{?obj type table}) \rightarrow \texttt{ikea\_table}$

$\text{find}(\texttt{?obj type book, ?obj isOn ikea\_table}) \rightarrow \texttt{harry\_potter}$

◂ Supplementary material

DIALOGUE GROUNDING

"Give me the book on the table"

↓

`me` → `human_1`

find(`?obj type table`) → `ikea_table`

find(`?obj type book, ?obj isOn ikea_table`) → `harry_potter`

↓

```
human_1 desires action1,
action1 type Give,
action1 performedBy myself,
action1 actsOnObject harry_potter,
action1 receivedBy human_1
```

◂ Supplementary material

CRI FOR LEARNING

SOCIAL OR NOT SOCIAL?

Non-social

Social

SOCIAL OR NOT SOCIAL?

Cellulo

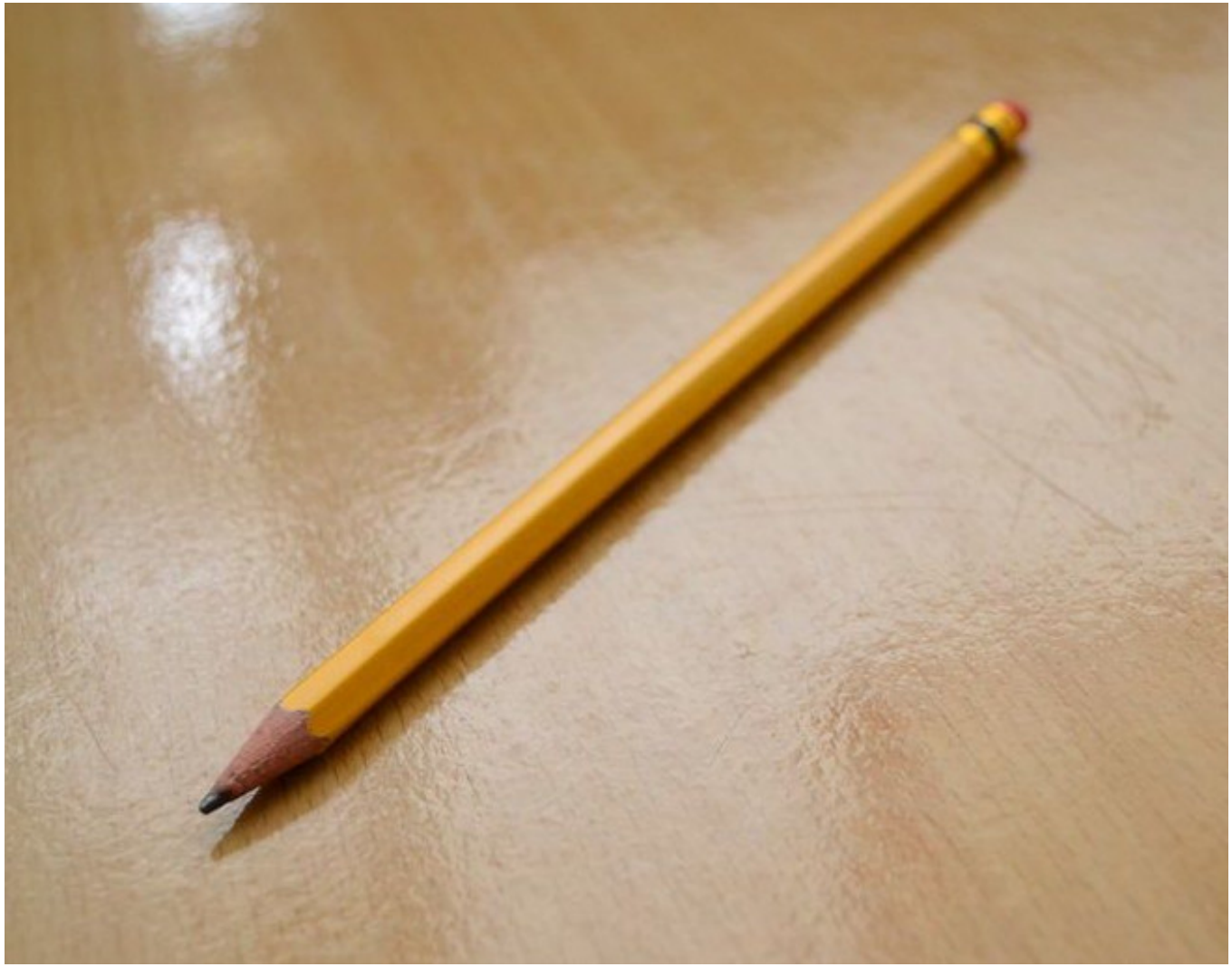
CoWriter

NON-SOCIAL INTERACTION

What is the most effective learning tool in a classroom?

NON-SOCIAL INTERACTION

What is the most effective learning tool in a classroom?



Pens and paper are pervasive...

...so should the robots

CELLULO: DESIGN PRINCIPLES

- **ubiquitous**: a pervasive yet unremarkable tool that blend into the daily learning routine; has to be trustworthy (i.e. reliable), readily replaceable (i.e. cheap, no affective bonding), intuitive (i.e. few simple affordances)

[Özgür, Lemaignan et al. **Cellulo: Versatile Handheld Robots for Education** – HRI 2017]

CELLULO: DESIGN PRINCIPLES

- **ubiquitous**: a pervasive yet unremarkable tool that blend into the daily learning routine; has to be trustworthy (i.e. reliable), readily replaceable (i.e. cheap, no affective bonding), intuitive (i.e. few simple affordances)
- **versatile**: applicable to a broad range of learning scenarios; the robots' hardware, appearance and interaction modalities must not imply or be constrained to specific use cases

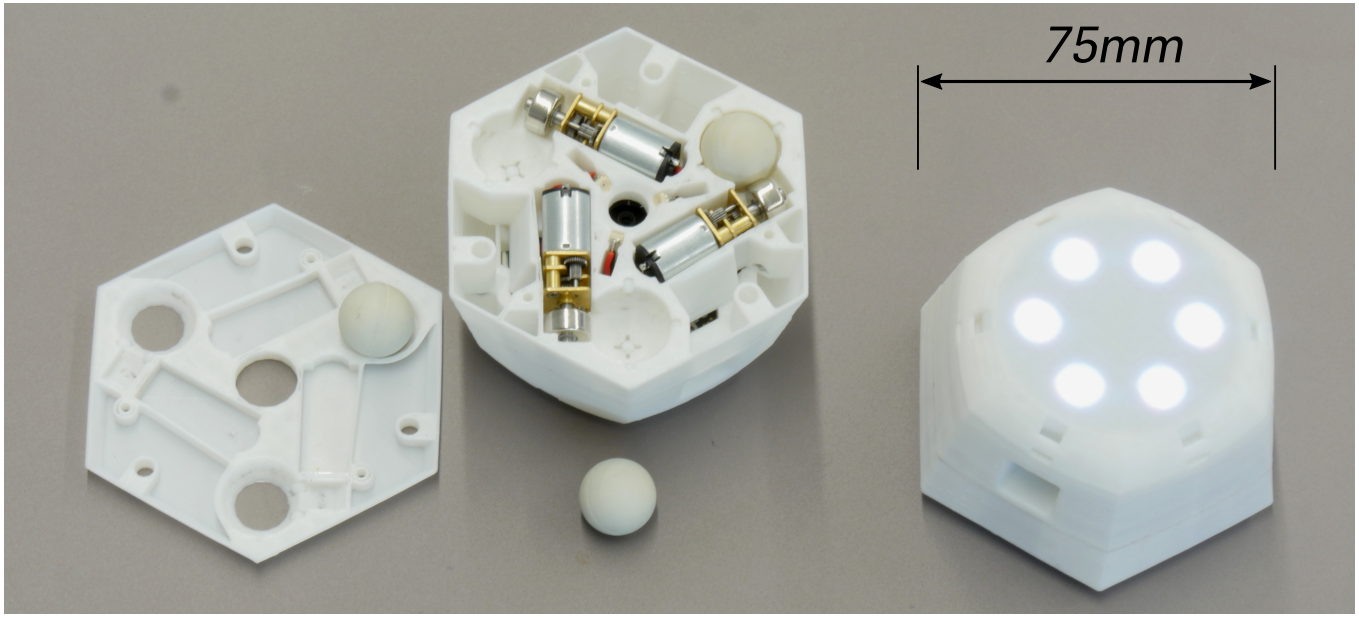
[Özgür, Lemaignan et al. **Cellulo: Versatile Handheld Robots for Education** – HRI 2017]

CELLULO: DESIGN PRINCIPLES

- **ubiquitous**: a pervasive yet unremarkable tool that blend into the daily learning routine; has to be trustworthy (i.e. reliable), readily replaceable (i.e. cheap, no affective bonding), intuitive (i.e. few simple affordances)
- **versatile**: applicable to a broad range of learning scenarios; the robots' hardware, appearance and interaction modalities must not imply or be constrained to specific use cases
- **practical**: to gain field acceptance in the classrooms, educative robots must critically represent a net educative gain and must not incur higher workload for the teachers

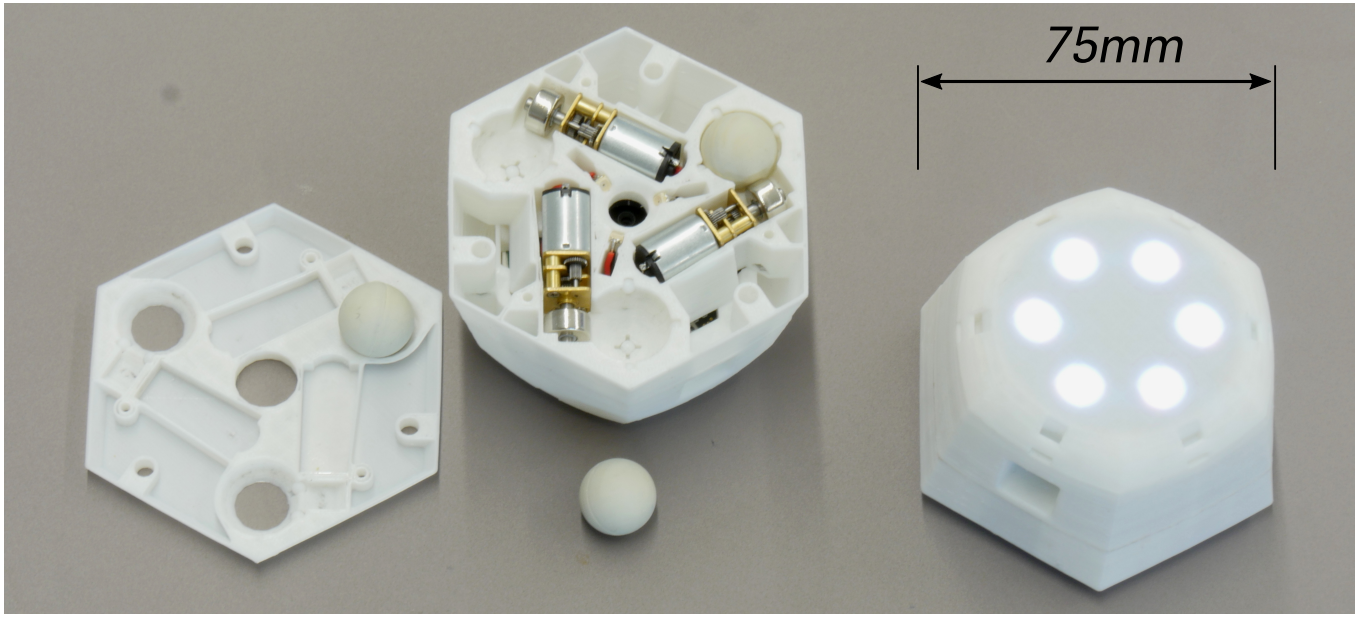
[Özgür, Lemaignan et al. **Cellulo: Versatile Handheld Robots for Education** – HRI 2017]

CELLULO: HARDWARE

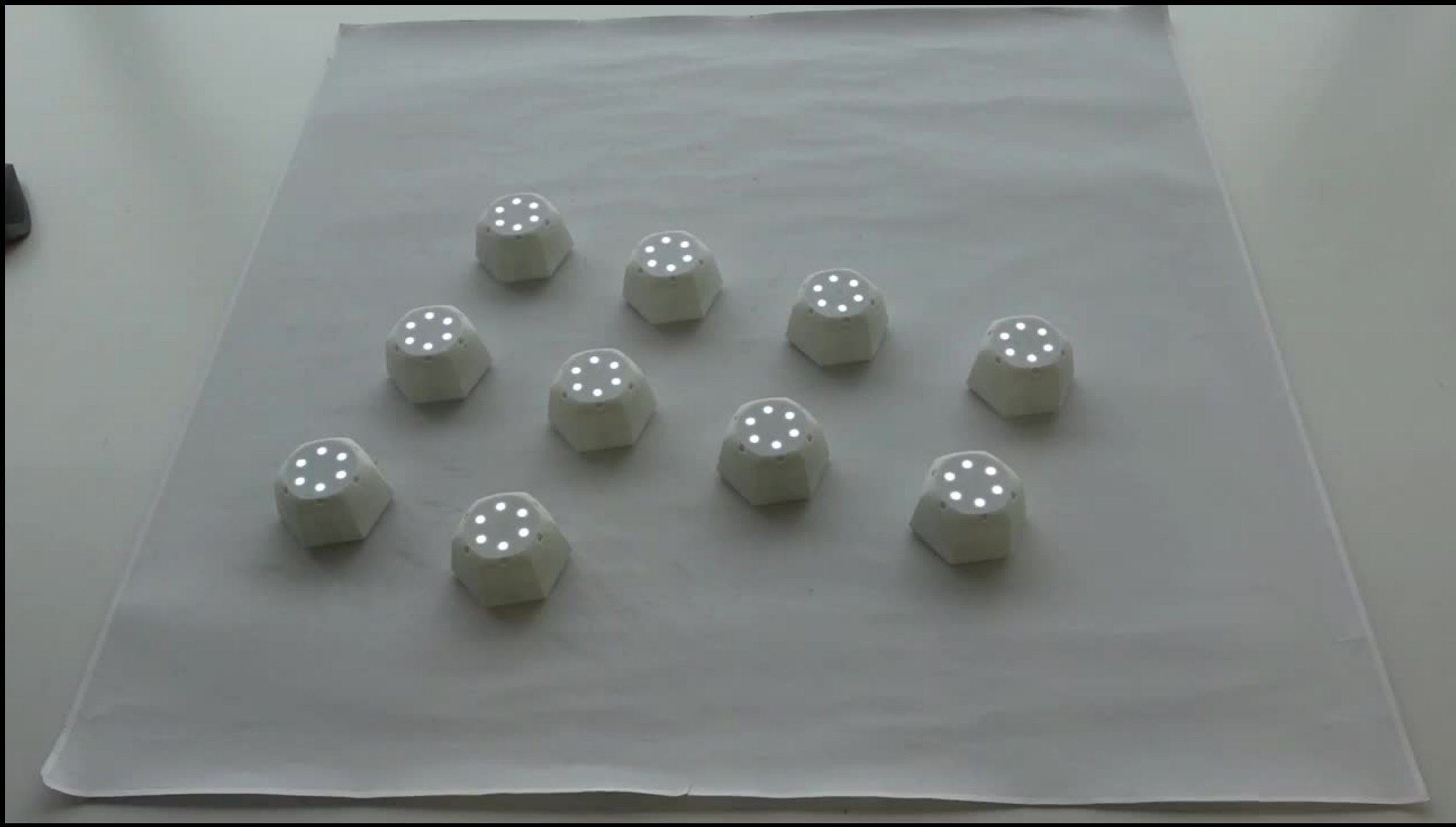


- Holomonic motion
- Sub-mm absolute localisation (no external hardware)
- Haptic feedback + tactile RGB LED buttons
- Bluetooth

CELLULO: HARDWARE



- Holomonic motion
- Sub-mm absolute localisation (no external hardware)
- Haptic feedback + tactile RGB LED buttons
- Bluetooth
- Affordable (prototype: €125)



INTERACTION WITH THE PAPER

Critically, Cellulo is meant as an **interaction between (classroom-friendly) paper and the robots**.

INTERACTION WITH THE PAPER

Critically, Cellulo is meant as an **interaction between (classroom-friendly) paper and the robots**.

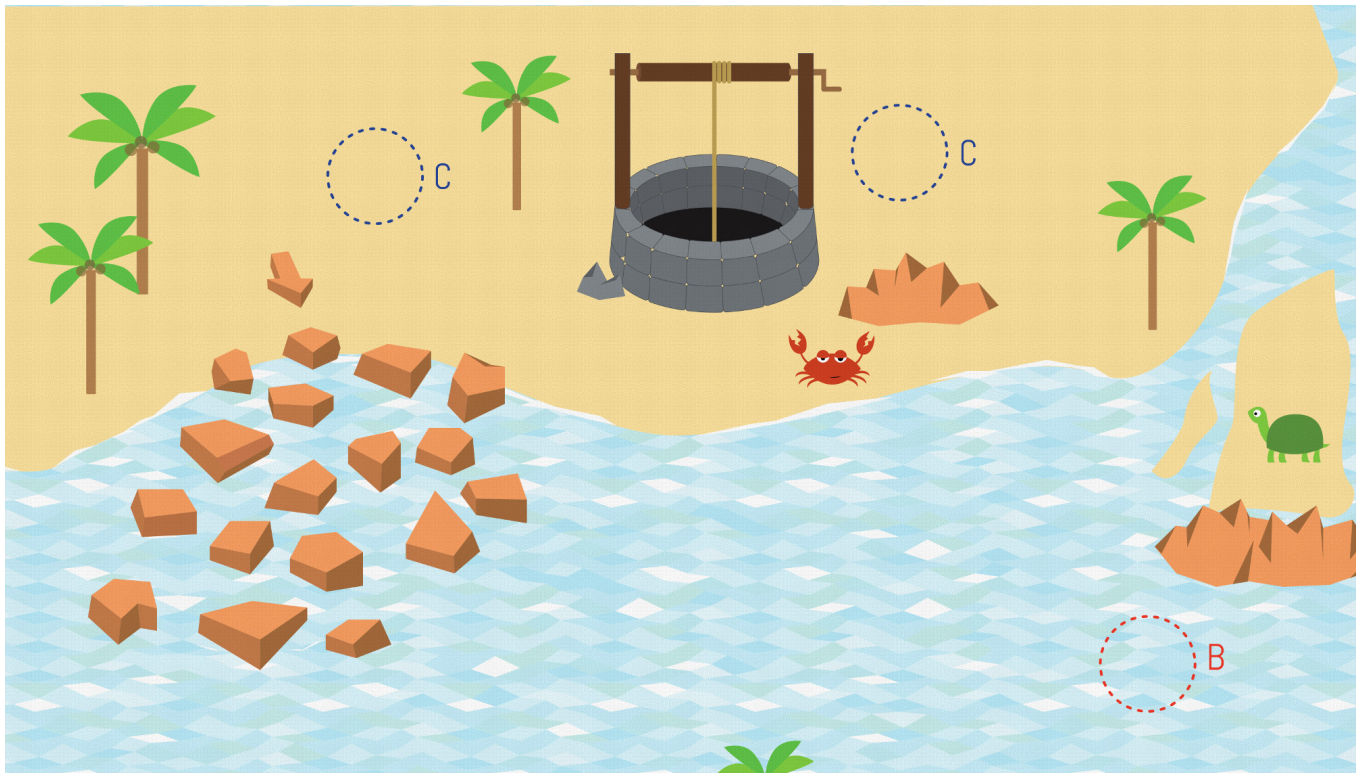
Achieved through a **paper-based absolute localisation system**

[Hostettler, Özgür, Lemaignan et al. **Real-Time High-Accuracy 2D Localization with Structured Patterns** – ICRA 2016]

INTERACTION WITH THE PAPER

Critically, Cellulo is meant as an **interaction between (classroom-friendly) paper and the robots**.

Achieved through a **paper-based absolute localisation system**

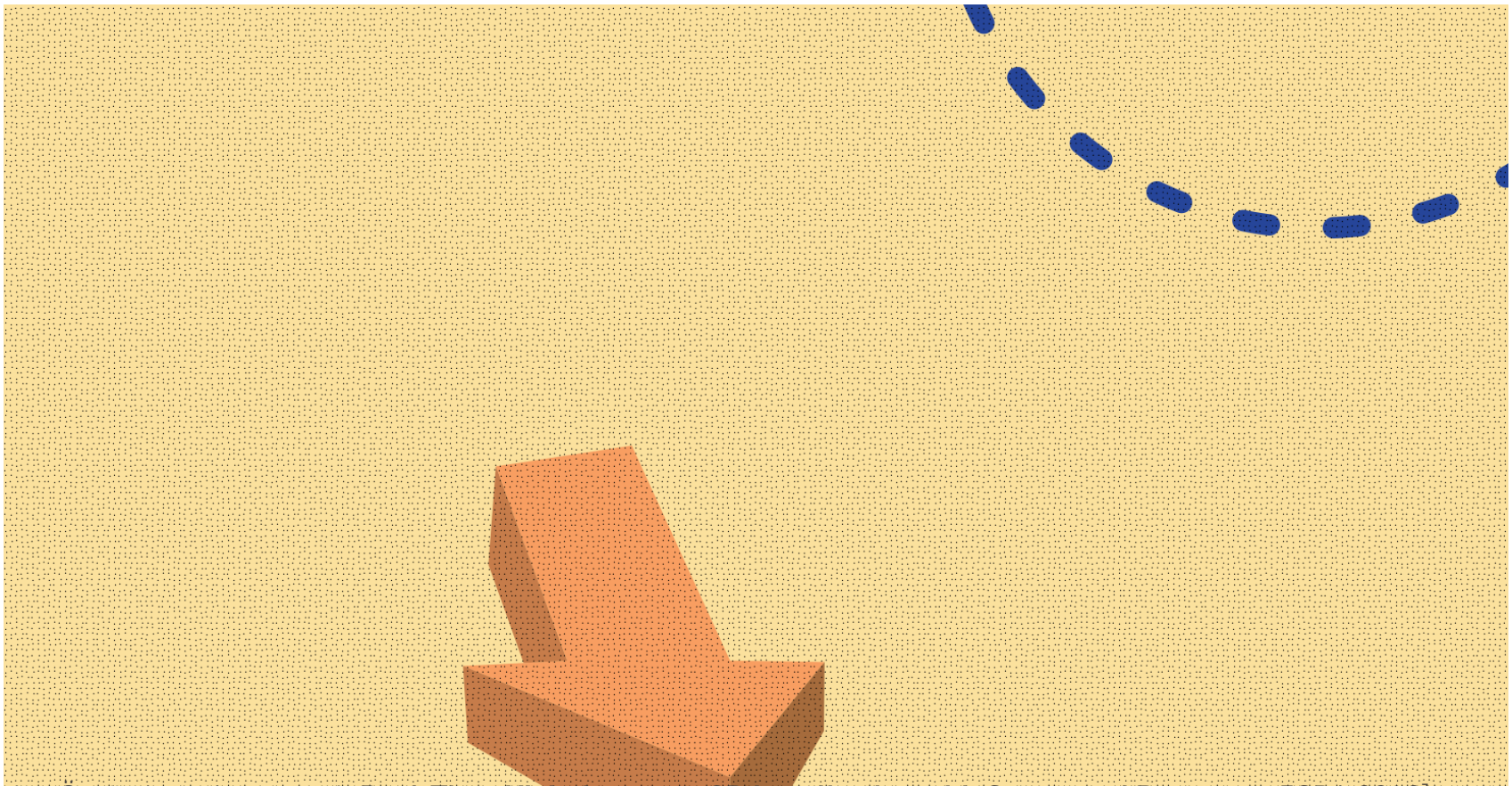


[Hostettler, Özgür, Lemaignan et al. **Real-Time High-Accuracy 2D Localization with Structured Patterns** – ICRA 2016]

INTERACTION WITH THE PAPER

Critically, Cellulo is meant as an **interaction between (classroom-friendly) paper and the robots**.

Achieved through a **paper-based absolute localisation system**



[Hostettler, Özgür, Lemaignan et al. **Real-Time High-Accuracy 2D Localization with Structured Patterns** – ICRA 2016]

INTERACTION WITH THE PAPER

Critically, Cellulo is meant as an **interaction between (classroom-friendly) paper and the robots**.

Achieved through a **paper-based absolute localisation system**

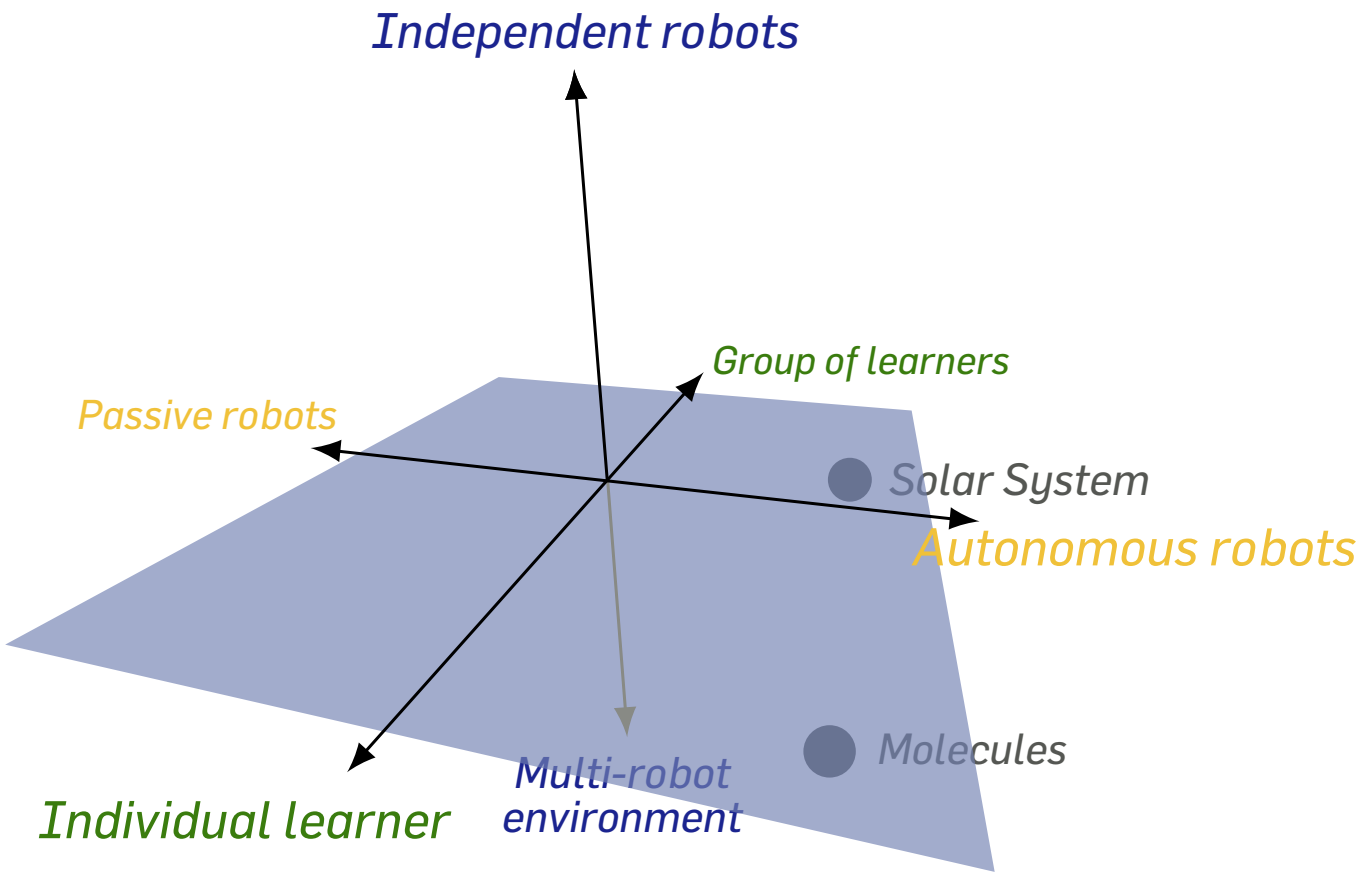
- even more than 'classroom-friendly', paper is 'teacher-friendly'
- easy to manipulate, copy, print, cutout, dispose...
- unique activity IDs: drop the robots onto the sheet, it recognizes the activity

[Hostettler, Özgür, Lemaignan et al. **Real-Time High-Accuracy 2D Localization with Structured Patterns** – ICRA 2016]

Concept 1: chemistry

Concept 2: solar system

INTERACTION DESIGN SPACE



[Özgür, Lemaignan et al. **Cellulo: Versatile Handheld Robots for Education** – HRI 2017]

...at the other end of the spectrum...

TECHNICAL CHALLENGES

- Get a child-proof robot to write...

TECHNICAL CHALLENGES

- Get a child-proof robot to write...
- ...badly...

TECHNICAL CHALLENGES

- Get a child-proof robot to write...
- ...badly...
- Make it able to learn...

TECHNICAL CHALLENGES

- Get a child-proof robot to write…
- …badly…
- Make it able to learn…
- …with the help of children

THE COWRITER PROJECT

Can we address children' hand-writing impairments with robots?

[Lemaignan et al. **Learning by Teaching a Robot: The Case of Handwriting** – Robotics and Automation Magazine 2016]

THE COWRITER PROJECT

Can we address children' hand-writing impairments with robots?

- Robots do not know how to write!

[Lemaignan et al. **Learning by Teaching a Robot: The Case of Handwriting** – Robotics and Automation Magazine 2016]

THE COWRITER PROJECT

Can we address children' hand-writing impairments with robots?

- Robots do not know how to write!
- Learning by Teaching

[Lemaignan et al. **Learning by Teaching a Robot: The Case of Handwriting** – Robotics and Automation Magazine 2016]

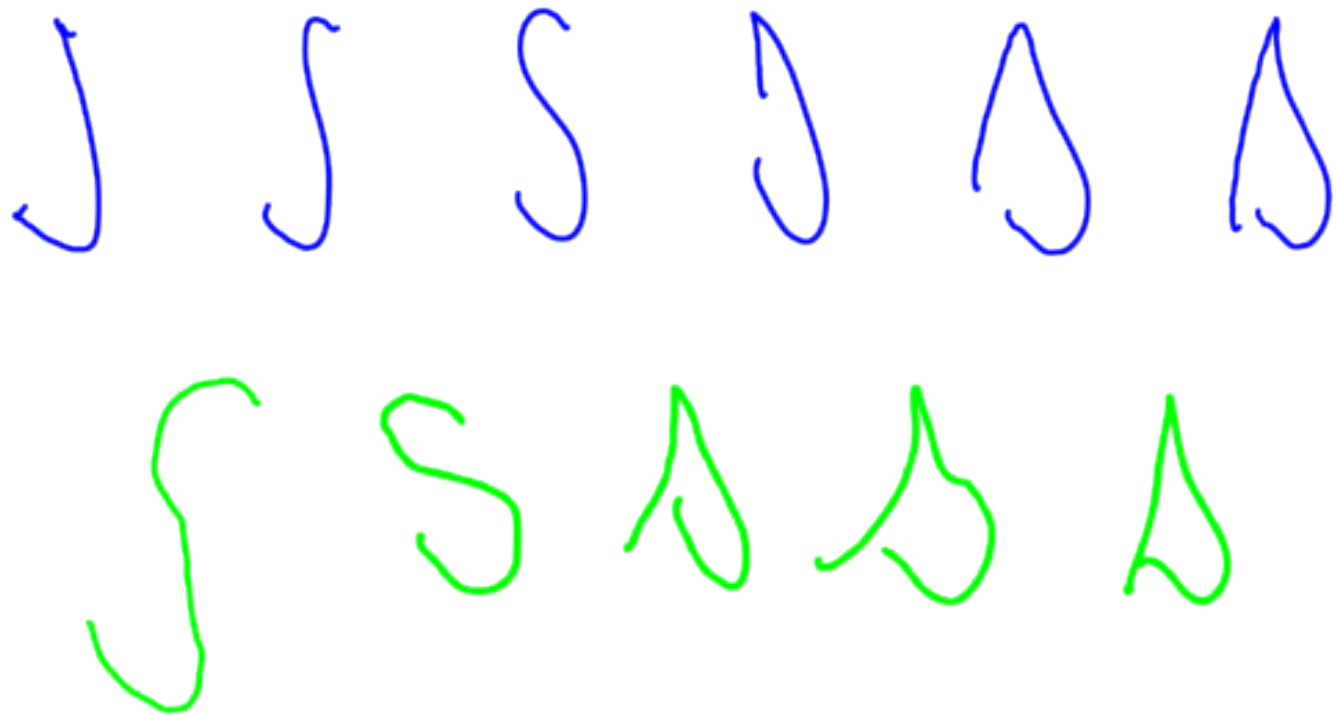
THE COWRITER PROJECT

Can we address children' hand-writing impairments with robots?

- Robots do not know how to write!
- Learning by Teaching
- (nice side-effect: we can adapt to each child and each disabilities)

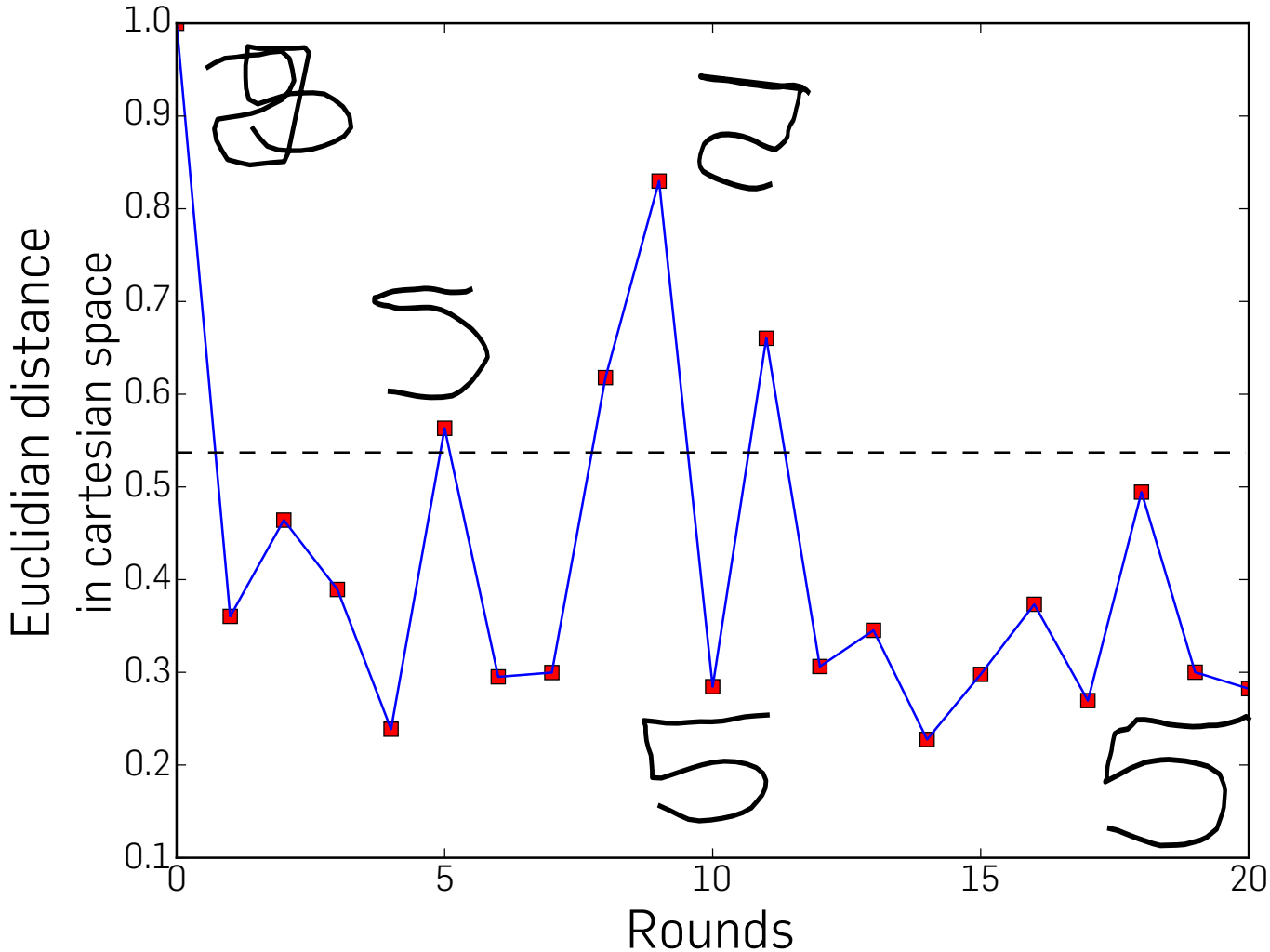
[Lemaignan et al. **Learning by Teaching a Robot: The Case of Handwriting** – Robotics and Automation Magazine 2016]

LEARNING FROM DEMONSTRATION



[Hood, Lemaignan et al. **When Children Teach a Robot to Write: An Autonomous Teachable Humanoid [...]** – HRI 2015]

LEARNING TO DRAW A 5

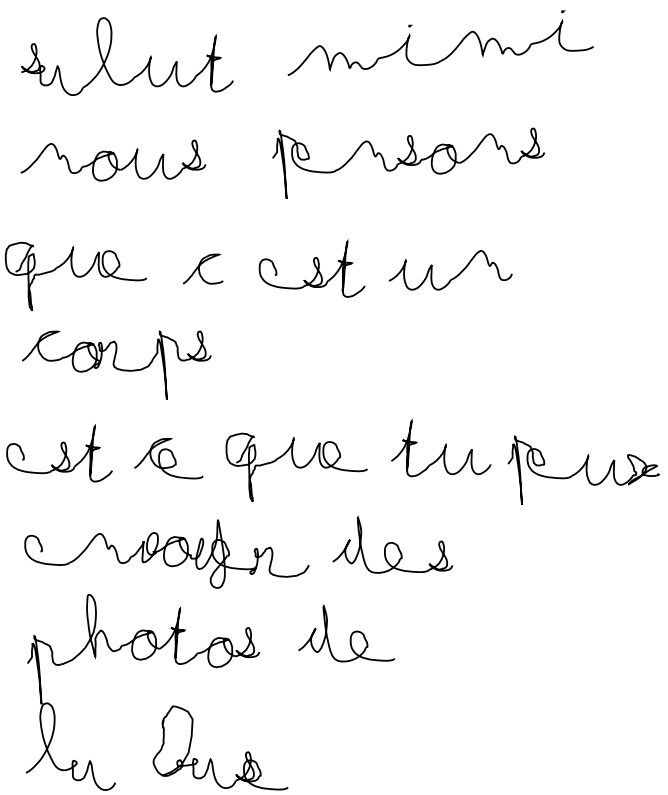


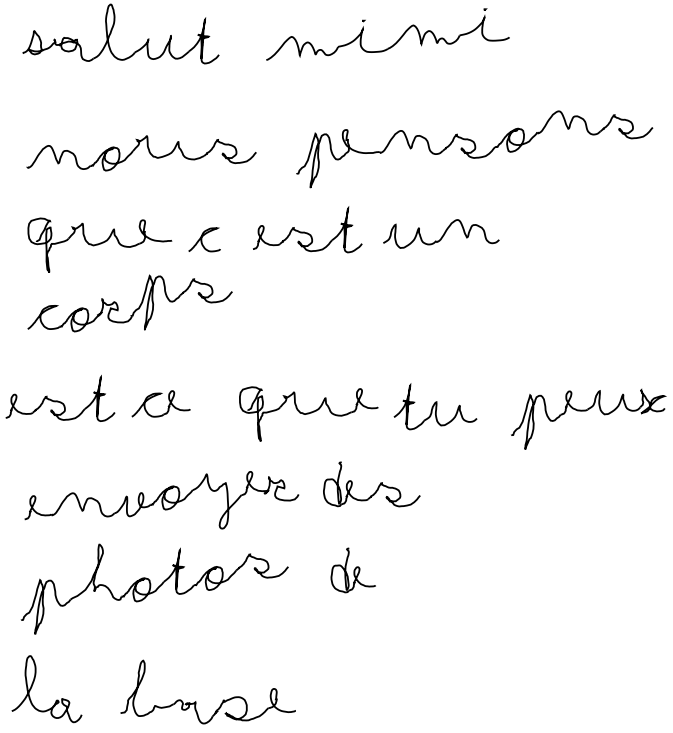
[Lemaignan et al. **Learning by Teaching a Robot: The Case of Handwriting** – Robotics and Automation Magazine 2016]

BEFORE – AFTER

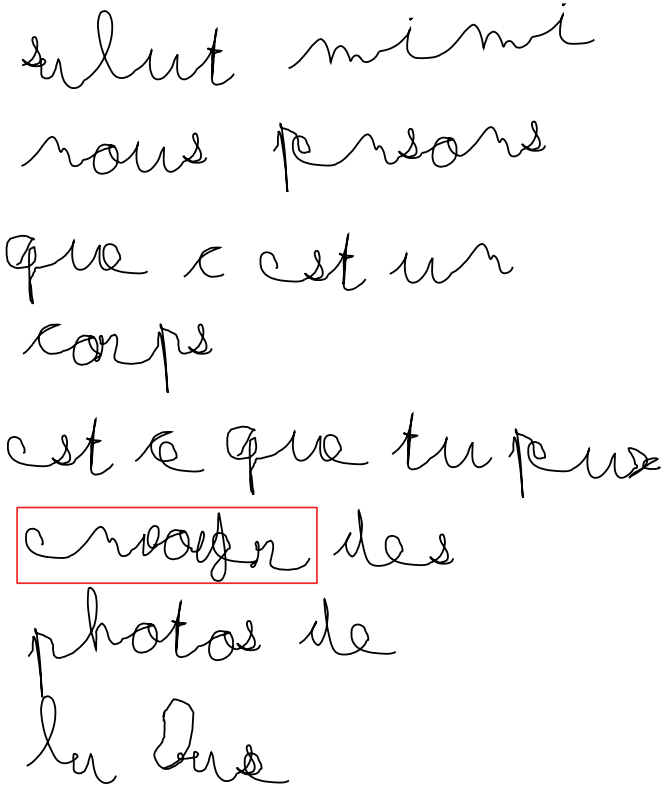
salut mimi
nous pensons
que c est un
corps
est ce que tu peux
croiser des
photos de
la Duse

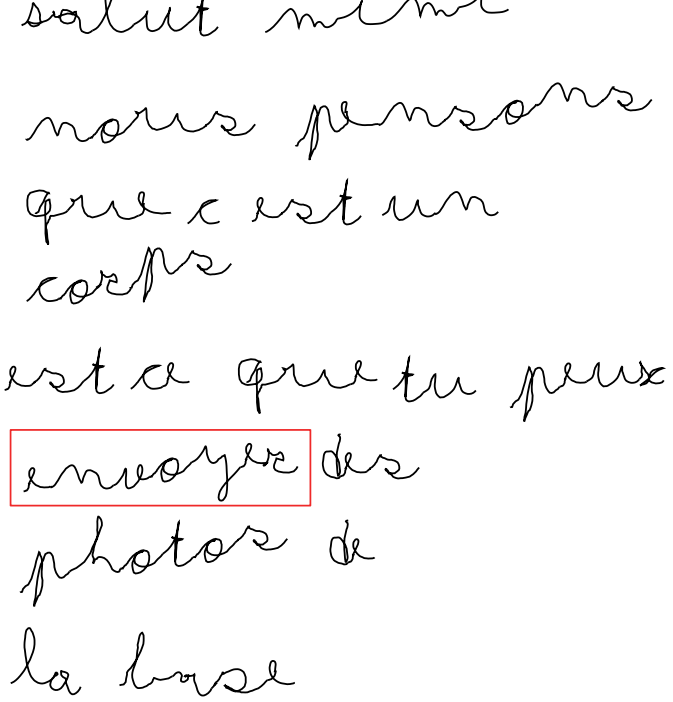
BEFORE – AFTER



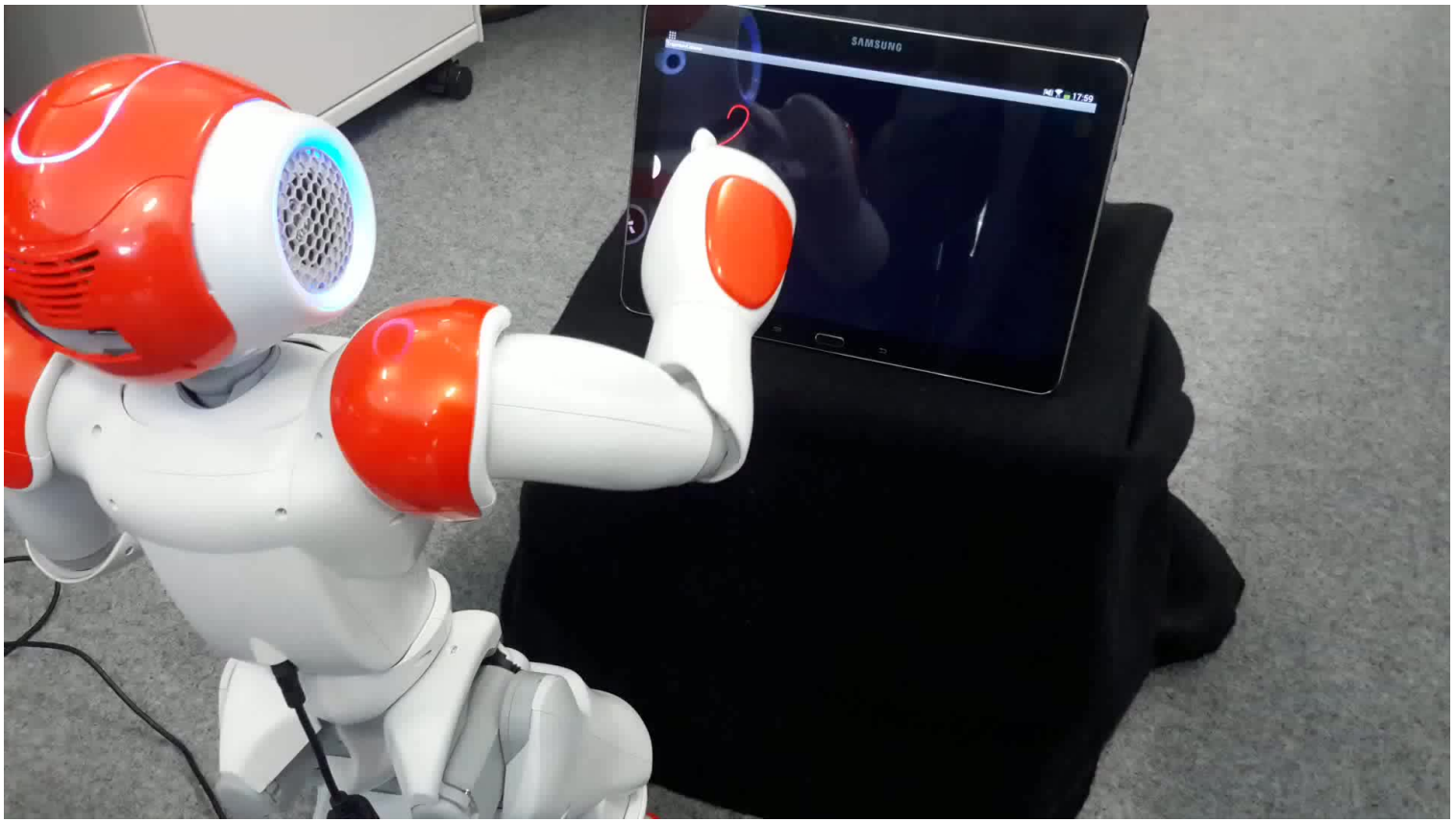


BEFORE – AFTER



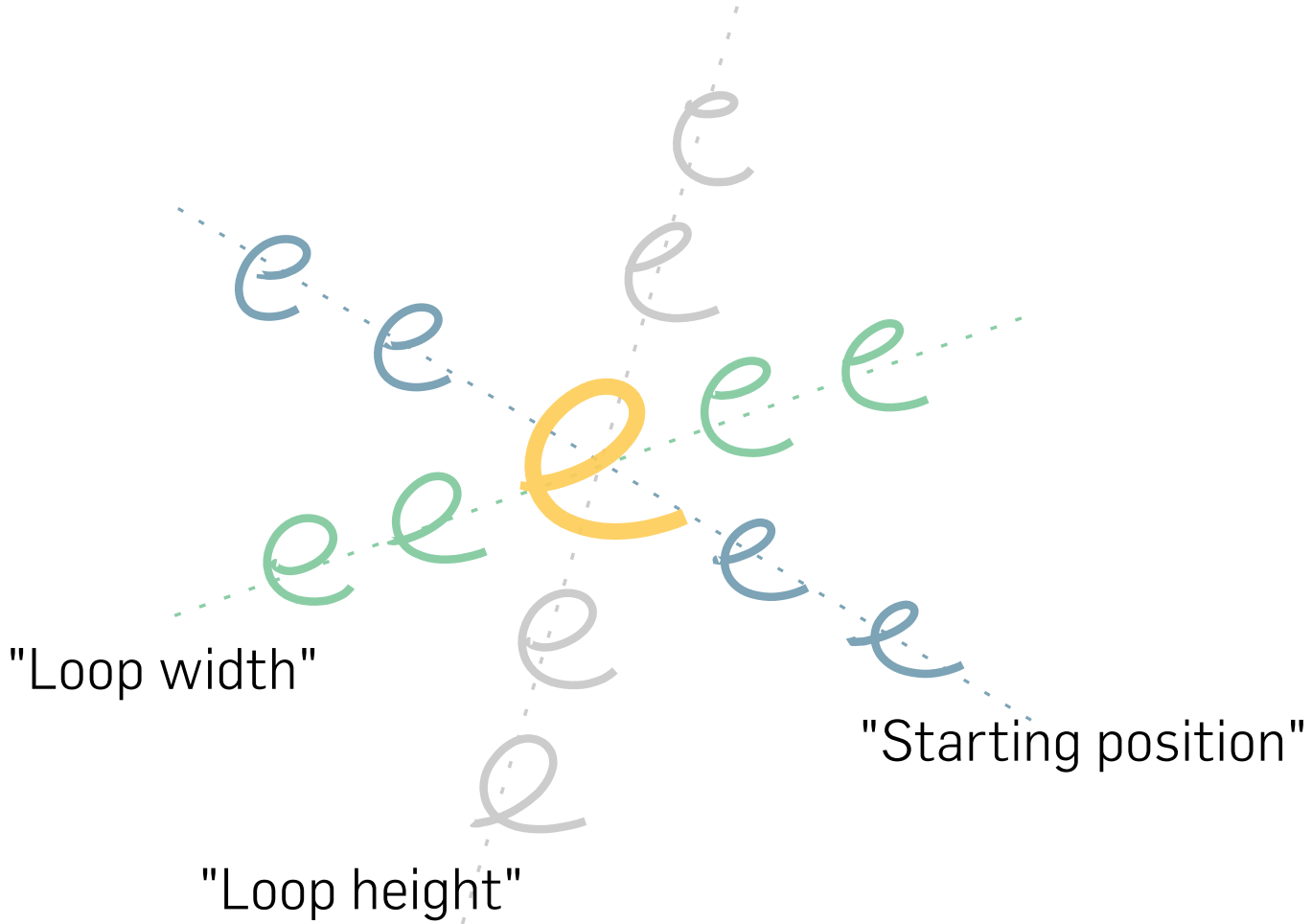


COWRITER IMPLEMENTATION



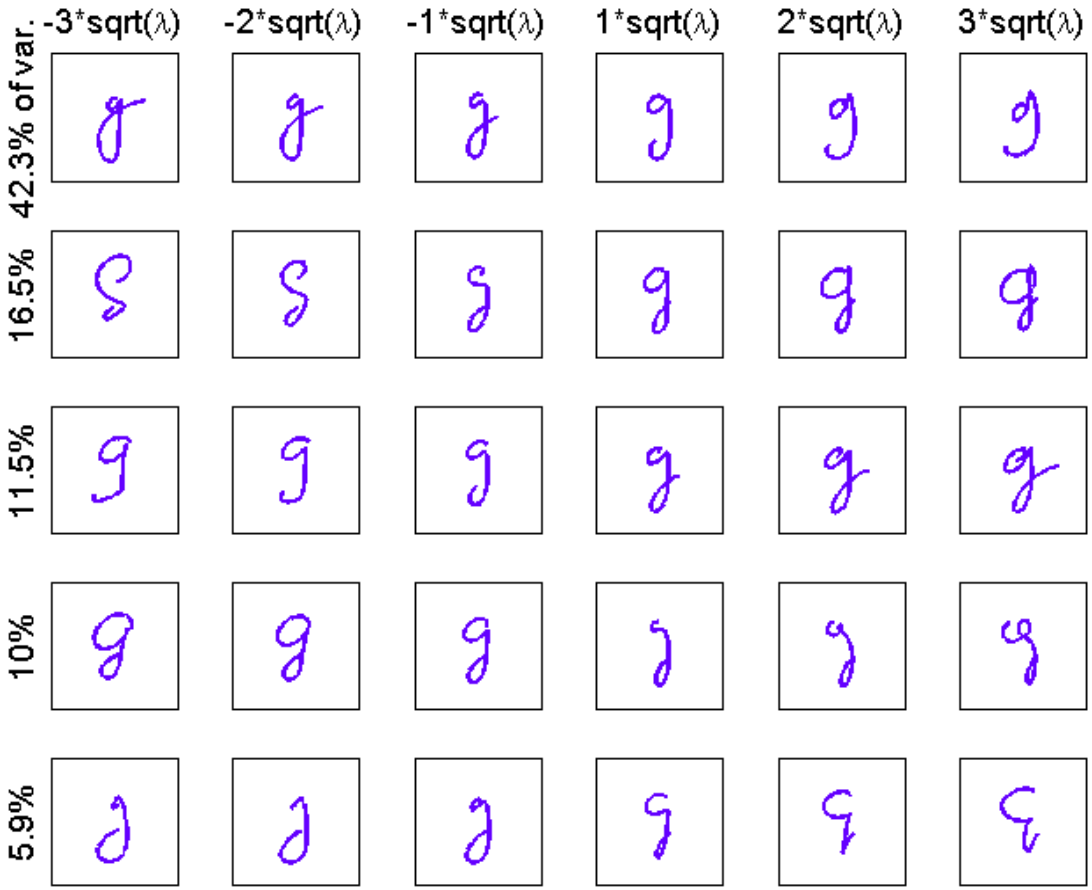
[Hood, Lemaignan et al. **When Children Teach a Robot to Write: An Autonomous Teachable Humanoid [...]** – HRI 2015]

COWRITER IMPLEMENTATION



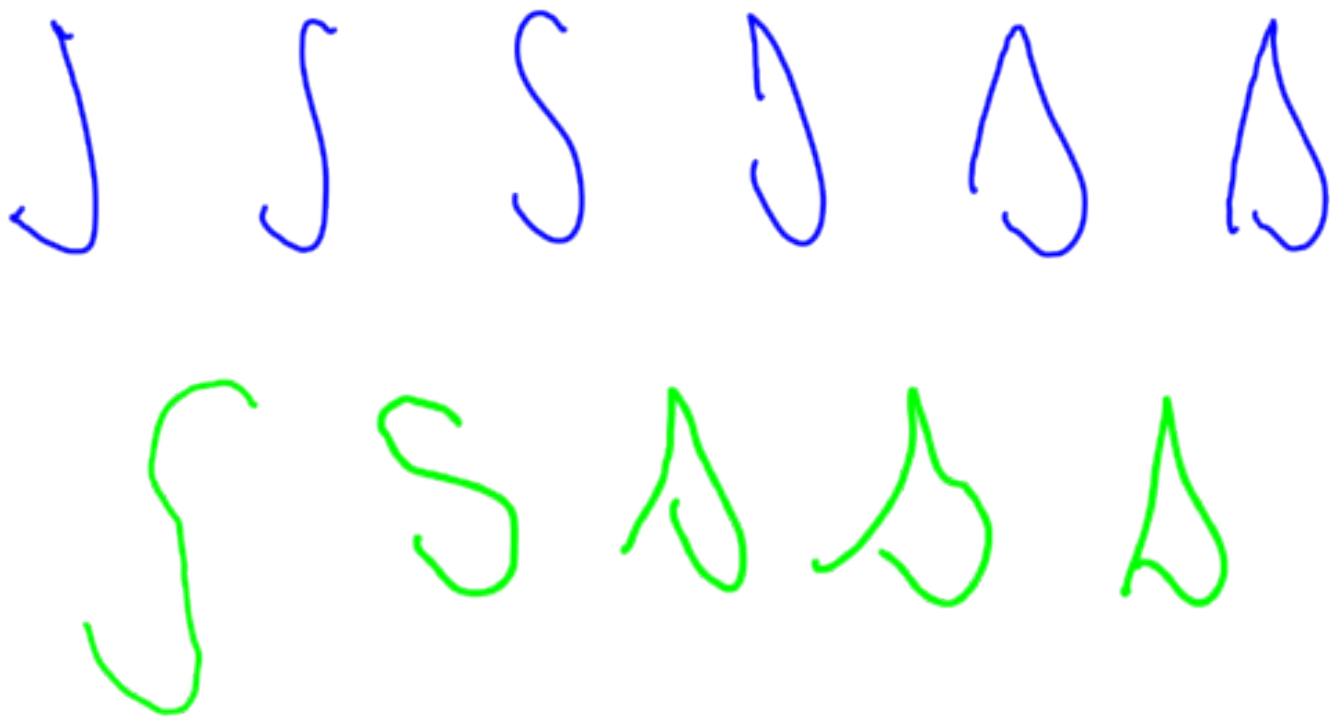
[Hood, Lemaignan et al. **When Children Teach a Robot to Write: An Autonomous Teachable Humanoid [...]** – HRI 2015]

COWRITER IMPLEMENTATION

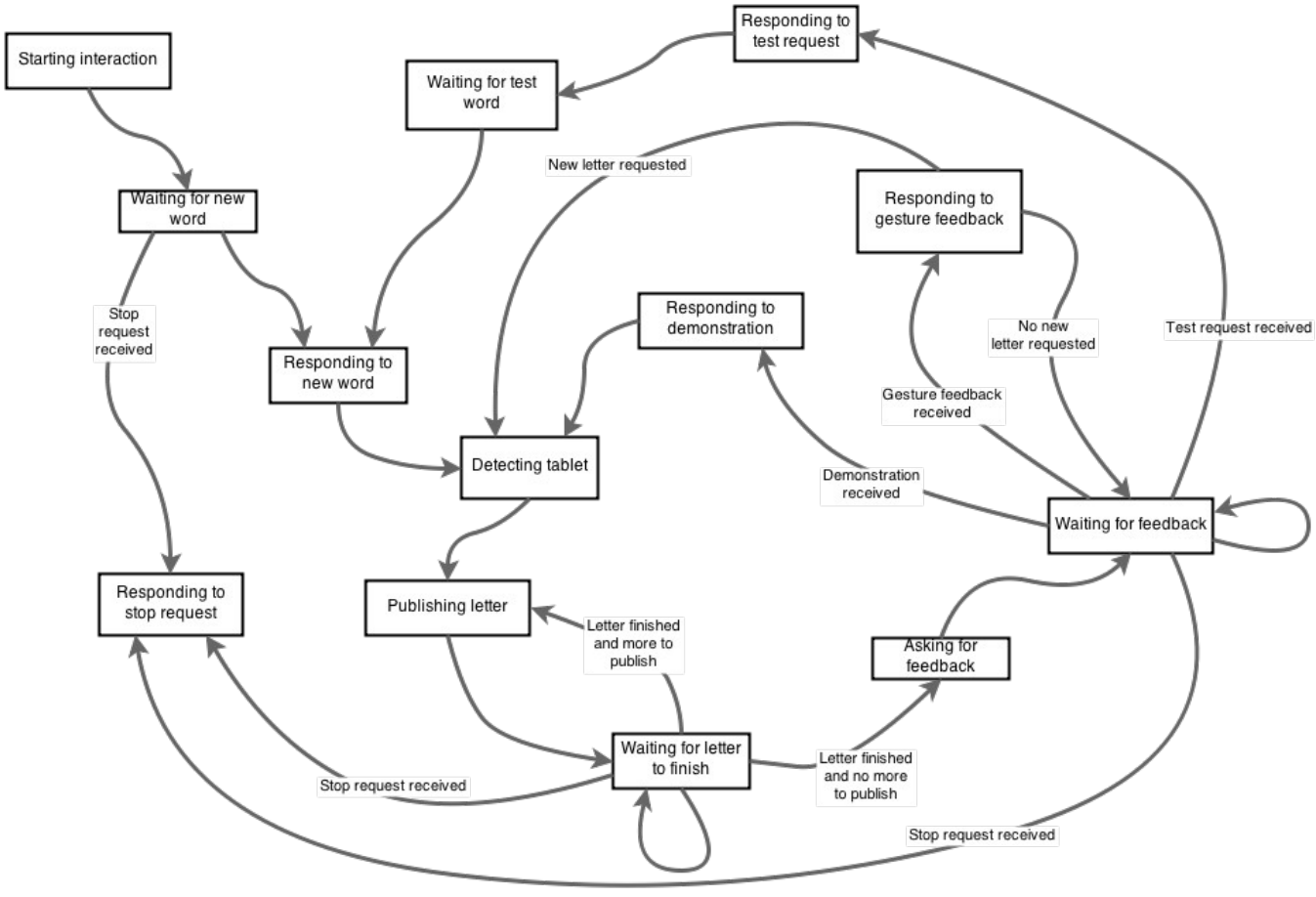


[Hood, Lemaignan et al. **When Children Teach a Robot to Write: An Autonomous Teachable Humanoid [...]** – HRI 2015]

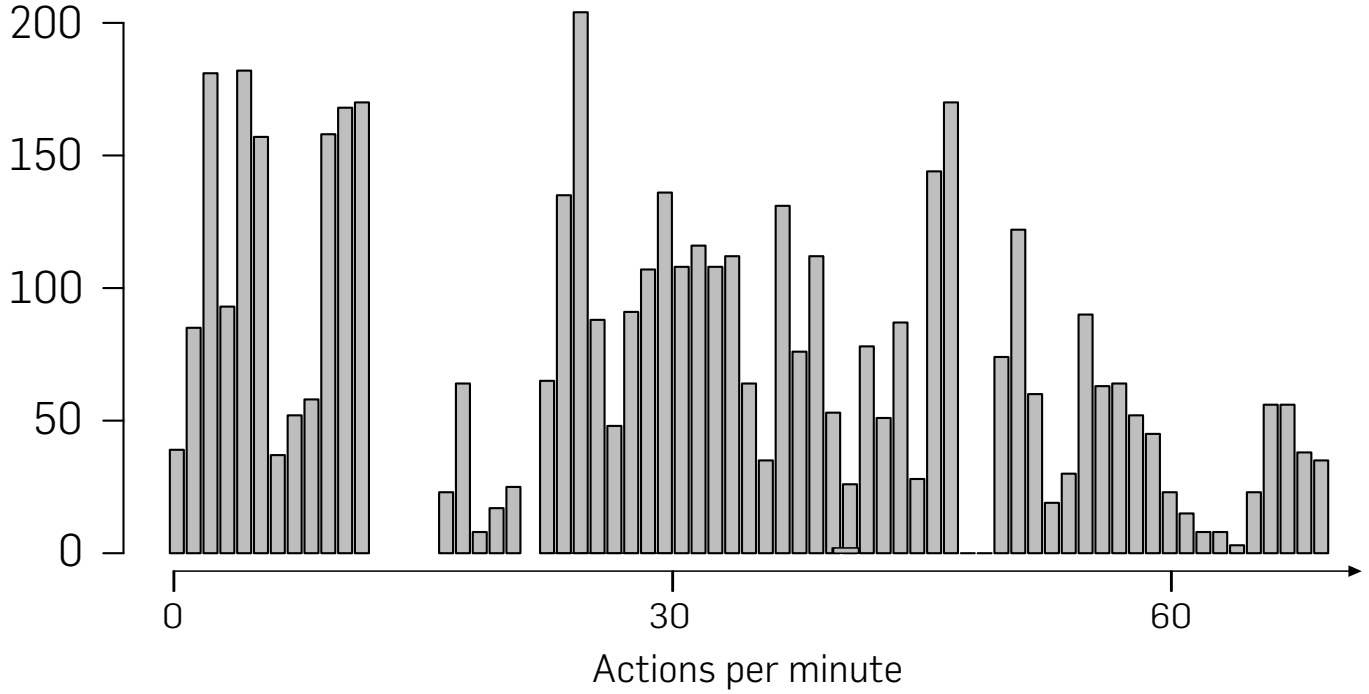
COWRITER IMPLEMENTATION



[Hood, Lemaignan et al. **When Children Teach a Robot to Write: An Autonomous Teachable Humanoid [...]** – HRI 2015]


Starting interaction
Responding to test request
Waiting for test word
Waiting for new word
New letter requested
Responding to gesture feedback
Stop request received
Responding to demonstration
No new letter requested
Test request received
Responding to new word
Gesture feedback received
Detecting tablet
Demonstration received
Waiting for feedback
Responding to stop request
Publishing letter
Letter finished and more to publish
Asking for feedback
Waiting for letter to finish
Letter finished and no more to publish
Stop request received
Stop request received

PRACTICAL CRI



lightbar	on_bumped	active_wait
on_toy_added	up_down_row	closeeyes
move	wakeup	lightpattern
background_blink	look_at_caresses	turn
undock	on_toy_removed	idle
pulse_row	sneak_in	playsound
blink	on_lolette_removed	blush
on_lolette	fall_asleep	
placeeyes	look_at_lolette	

◂ Supplementary material

Can we make the analysis of child-robot interaction **practical**?

- (surface) engagement
- cognitive perception/anthropomorphism
- child speech recognition

WITH-ME-NESS

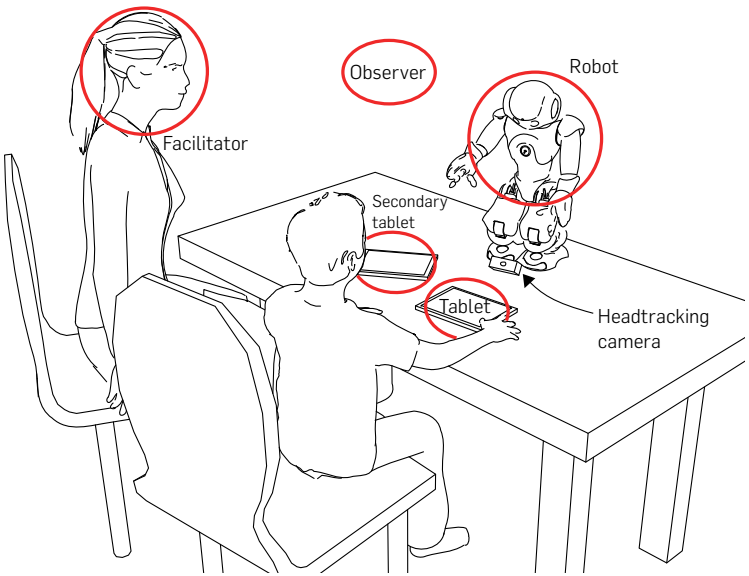
“With-me-ness”: real-time estimation of surface engagement

[Lemaignan et al. **From Real-time Attention Assessment to “With-me-ness” in Human-Robot Interaction** HRI 2016]

EXPECTED FOCUS

Example for the CoWriter task:

Interaction Phase	Expected targets
Presentation	robot
Waiting for word	secondary tablet
Writing word	tablet robot
Waiting for feedback	tablet secondary tablet
Story telling	robot
Bye	robot



WITH-ME-NESS

[Lemaignan et al. **From Real-time Attention Assessment to "With-me-ness" in Human-Robot Interaction** HRI 2016]

WITH-ME-NESS IS…

With-me-ness is…

- An **objective** & **quantitative** precursor of engagement…
- …based on matching the **user's focus of attention** with a set of **prior expectations**
- Can be computed **on-line** by the robot…
- …and **sensitive to** the (task-dependent) **set of expectations**
- ⇒ **relative** metric!

◂ Supplementary material

CONSTRUCTS FOR COGNITIVE PERCEPTION ANALYSIS

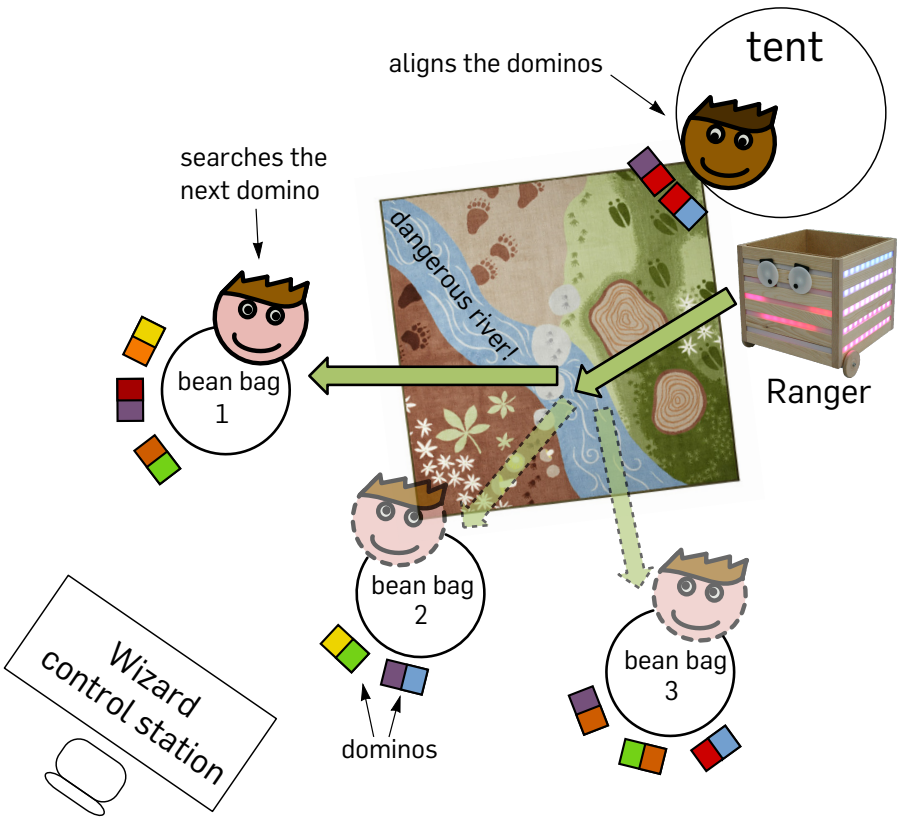
Expectations
How do you imagine a robot? *What could it look like?* *Have you ever seen a robot before?*
Impression
When you first saw R, what did you think? *Is R a robot? How do you know?* *Did you expect R would come over to you when you call it?* *What happened when you put the domino in the box?*
Ascribe intention
Do you think R could go out the door all by itself? *Does R always obey / come over to you?* *Could R do something silly?* *Why did R not come over to you when you called it?*

Ascribe perceptual capabilities
Here is a domino. Do you think R can see it? *When I say "Hello R!", do you think R can hear it?*
Ascribe emotional state
Does R have feelings? Can R be happy or sad sometimes?
Social acceptance
Do you like R? Why (not)? *What do you (not) like about it?* *Would you like to have R at home?*
Companionship
Could R be your friend? Why (not)?
Ascribe moral standing
Assume you go on a holiday for two weeks. Is it alright to leave R alone at home? Why (not)?

◂ Supplementary material

BEHAVIOUR VS PERCEPTION?

Any relation between the behavioural and perceptual measurements?



[Lemaignan, Fink, Mondada, Dillenbourg **You're Doing It Wrong! Studying Unexpected Behaviors in CRI** ICSR 2015]

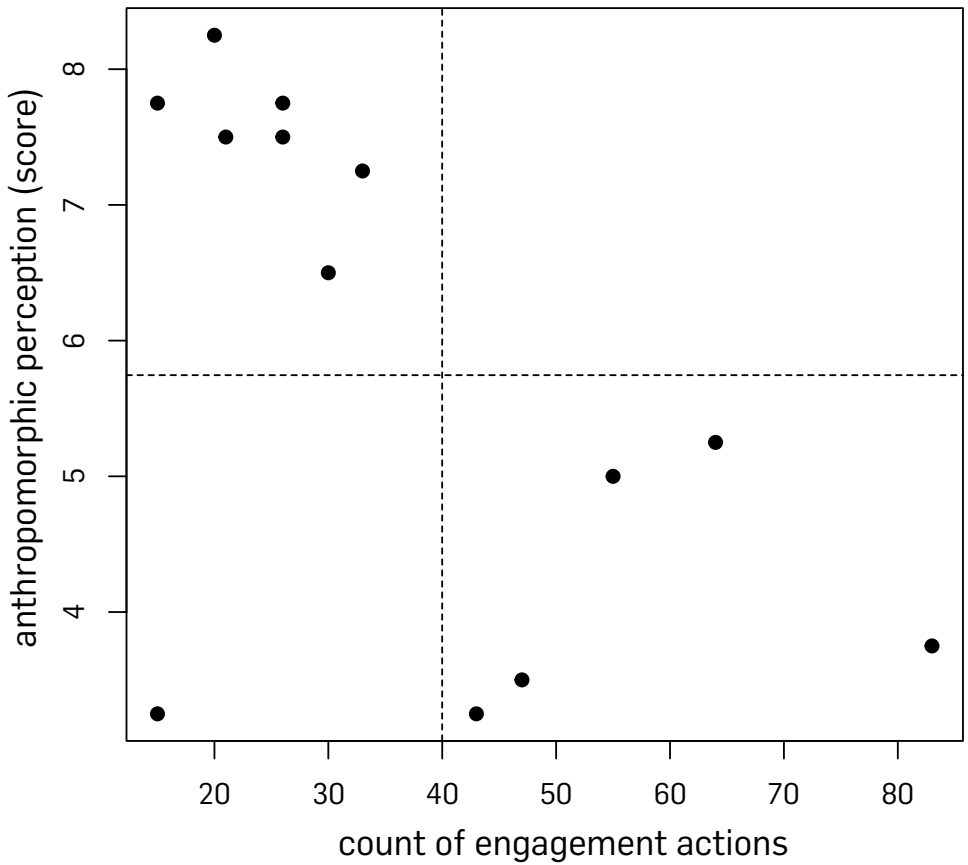
BEHAVIOUR VS PERCEPTION?

Any relation between the behavioural and perceptual measurements?

We can compute for each pair an “anthropomorphic perception” score based on the cognitive ascriptions, and…

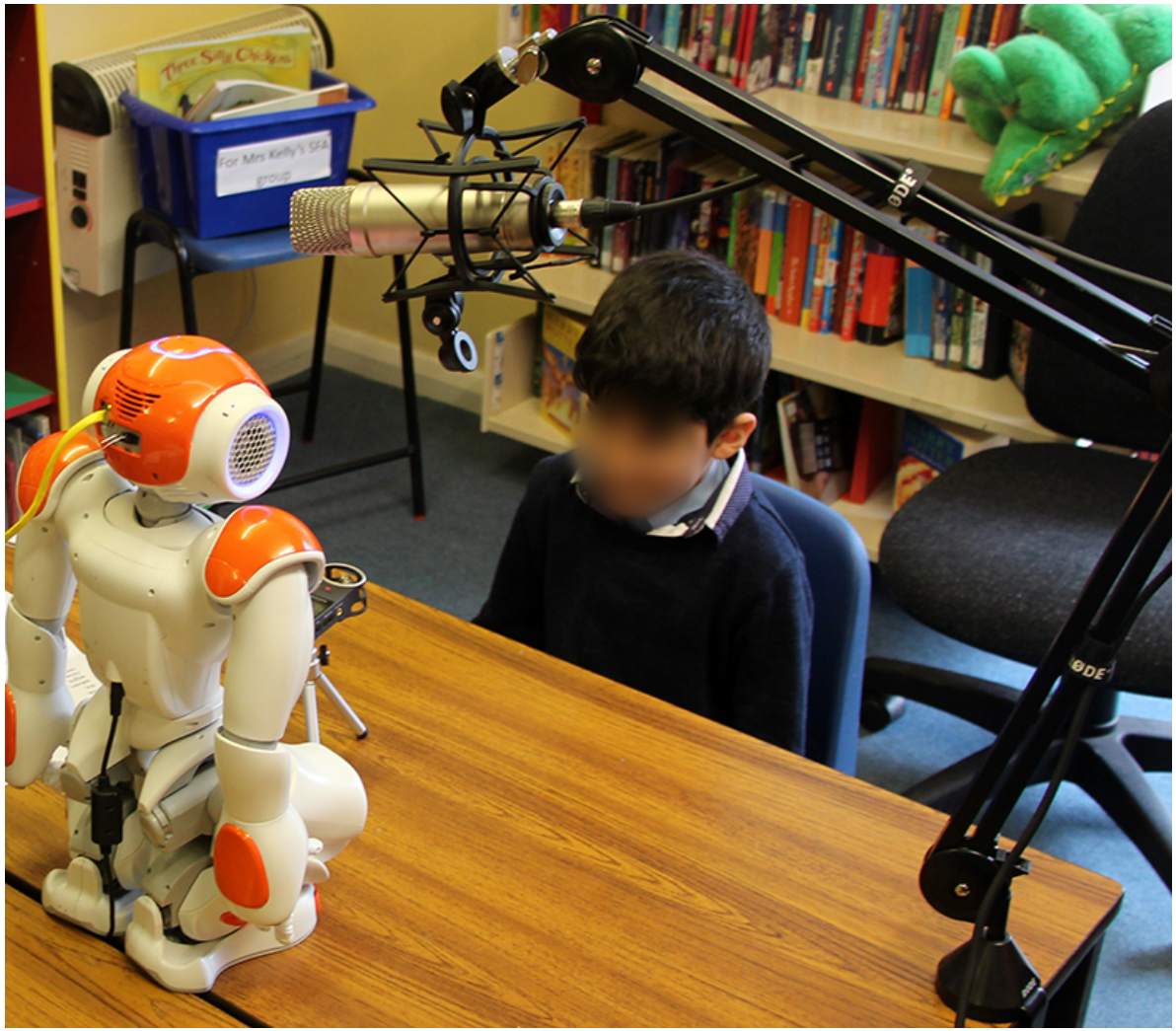
[Lemaignan, Fink, Mondada, Dillenbourg **You’re Doing It Wrong! Studying Unexpected Behaviors in CRI** ICSR 2015]

ANTHROPOMORPHISM != ENGAGEMENT



◂ Supplementary material

AUTOMATIC SPEECH RECOGNITION WITH CHILDREN



[Kennedy, Lemaignan et al. **Child Speech Recognition in Human-Robot Interaction[...]** HRI 2017]

AUTOMATIC SPEECH RECOGNITION WITH CHILDREN

	Google		Bing		Sphinx		Nuance	
	M LD	*% rec.*	*M* LD	*% rec.*	*M* LD	*% rec.*	*M* LD	*% rec.*
fixed (*n*=34)	**0.34**	*11.8 [38]*	0.64	*0 [0]*	0.68	*0 [0]*	0.76	*0 [0]*
spontaneous (*n*=222)	**0.39**	*6.8 [17.6]*	0.64	*0.5 [2.4]*	0.80	*0 [0]*	0.80	*0 [0]*
spontaneous clean only (*n*=83)	**0.40**	*6.0 [16.9]*	0.63	*1.2 [1.2]*	0.78	*0 [0]*	0.78	*0 [0]*

M LD: mean Levenshtein distance, at word level.

◂ Supplementary material

REFRAMING THE RESEARCH

OUR STARTING POINT

Symbolic artificial social cognition: works rather well as long as:

- we know what we want to do (in terms of task domain & declarative knowledge)
- interaction mostly relying on symbolic ***perceptual inputs*** (including visual perspective taking) rather abstract or less explicit representations

OUR STARTING POINT

Symbolic artificial social cognition: works rather well as long as:

- we know what we want to do (in terms of task domain & declarative knowledge)
- interaction mostly relying on symbolic *perceptual inputs* (including visual perspective taking) rather abstract or less explicit representations

Good for any practical HRI purposes? **mostly**!

OUR STARTING POINT

Symbolic artificial social cognition: works rather well as long as:

- we know what we want to do (in terms of task domain & declarative knowledge)
- interaction mostly relying on symbolic *perceptual inputs* (including visual perspective taking) rather abstract or less explicit representations

Good for any practical HRI purposes? **mostly**!

However, intuitively, social modeling goes beyond computing what the human perceives or does not perceive → Flavell's *cognitive connections* vs *mental representations*.

Symbolic cognition **does not explain much about how social cognition actually work**. We need a **principled approach** to social cognition for robots

A LONG-TERM DIRECTION

Adapting and unifying the large and disparate set of theories on social cognition to **build a theory of social cognition for robots**

A LONG-TERM DIRECTION

Adapting and unifying the large and disparate set of theories on social cognition to **build a theory of social cognition for robots**

...or rather, **a computational model of social cognition for robots**

A LONG-TERM DIRECTION

Adapting and unifying the large and disparate set of theories on social cognition to **build a theory of social cognition for robots**

...or rather, **a computational model of social cognition for robots**

...or rather, an **embodied** computational model of social cognition?

ONE QUESTION

Can sociality emerge from interaction?

ONE QUESTION

Can sociality emerge from interaction?

Both “emerge” as ***arise from*** and “emerge“ as in ***emergent paradigm of cognition***!

ONE QUESTION

Can sociality emerge from interaction?

Both “emerge” as ***arise from*** and “emerge“ as in ***emergent paradigm of cognition***!

“Social cognition arising in interaction”? certainly looks like a situated & embodied view on cognition

A MODEL?

Models attempt to ***explain***:

> *"identifying the causes for an event or phenomenon of interest"*
>
> *"unifying disparate phenomena"*

A model's value is gained from

> *"predicting facts that, absent the theory, would be antecedently improbable"*

[cited in Lewandowsky and Farrell, **Computational Modeling In Cognition**, 2011]

A MODEL?

Models attempt to ***explain***:

> *"identifying the causes for an event or phenomenon of interest"*
>
> *"unifying disparate phenomena"*

A model's value is gained from

> *"predicting facts that, absent the theory, would be antecedently improbable"*

...we will come back to the predictive power of a model of artifical social cognition.

[cited in Lewandowsky and Farrell, **Computational Modeling In Cognition**, 2011]

SKETCHING A MODEL

A MODEL OF ARTIFICIAL SOCIAL COGNITION

I postulate **two stages**:

1. building models of others' minds
2. exploiting these models to socially act:
   - prediction, reading others' intentions
   - adapting own behaviour, alignment
   - establish join goals
   - ultimately, performing joint actions

→ Social analogs of *perception* & *action*

COGNITIVIST VS EMERGENT PARADIGMS

"building", "exploiting", "reading", "establishing"... my terminolgy denotes a cognitivist approach ('I, the designer of the system, explicitly implement these capabilities')

[see discussion in Vernon, **Artificial Cognitive Systems: A Primer** 2014]

COGNITIVIST VS EMERGENT PARADIGMS

"building", "exploiting", "reading", "establishing"... my terminolgy denotes a cognitivist approach ('I, the designer of the system, explicitly implement these capabilities')

Possible 'emergent paradigm' rephrasing:

1. developing internal states ***connoting*** others' minds
2. perturbing (influencing) actions synthesis with these states

[see discussion in Vernon, **Artificial Cognitive Systems: A Primer** 2014]

COGNITIVIST VS EMERGENT PARADIGMS

"building", "exploiting", "reading", "establishing"... my terminolgy denotes a cognitivist approach ('I, the designer of the system, explicitly implement these capabilities')

Possible 'emergent paradigm' rephrasing:

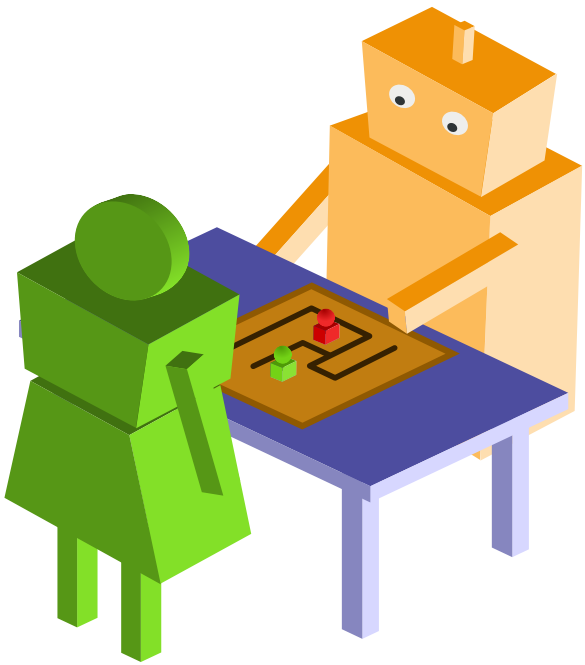
1. developing internal states ***connoting*** others' minds
2. perturbing (influencing) actions synthesis with these states

Hybrid approaches are possible – mapping to "raw phenomenal experience" vs "access counciousness".

[see discussion in Vernon, **Artificial Cognitive Systems: A Primer** 2014]

MODELING OTHERS' MIND?

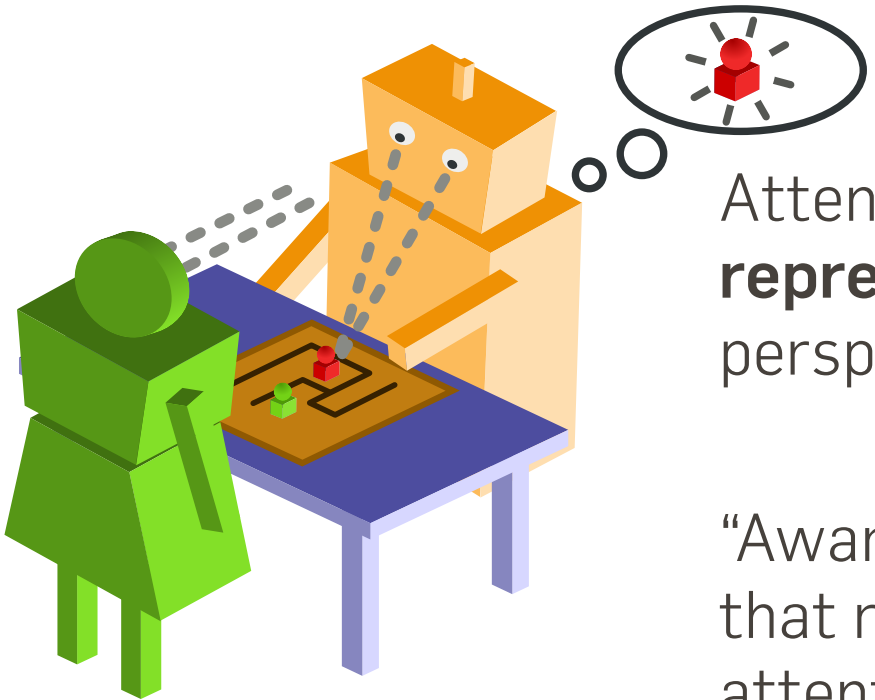
In cognitive neurosciences: Graziano's *Attention Schemata Theory*



[Graziano **Consciousness and the Social Brain** – 2013]

MODELING OTHERS' MIND?

In cognitive neurosciences: Graziano's *Attention Schemata Theory*



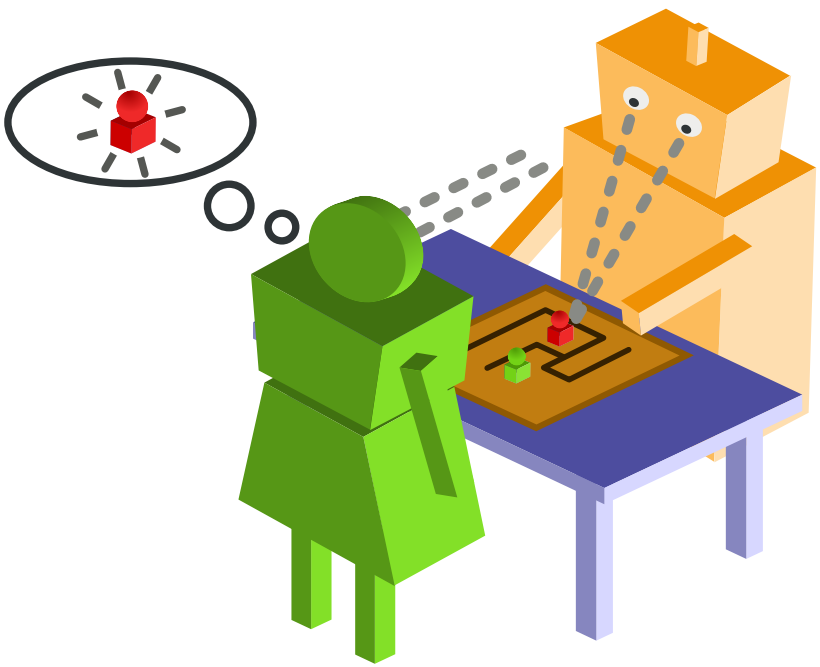
Attention is more about **representation** than visual perspective

"Awareness is a construct that represents the attentional state of a brain"

[Graziano **Consciousness and the Social Brain** – 2013]

MODELING OTHERS' MIND?

In cognitive neurosciences: Graziano's *Attention Schemata Theory*

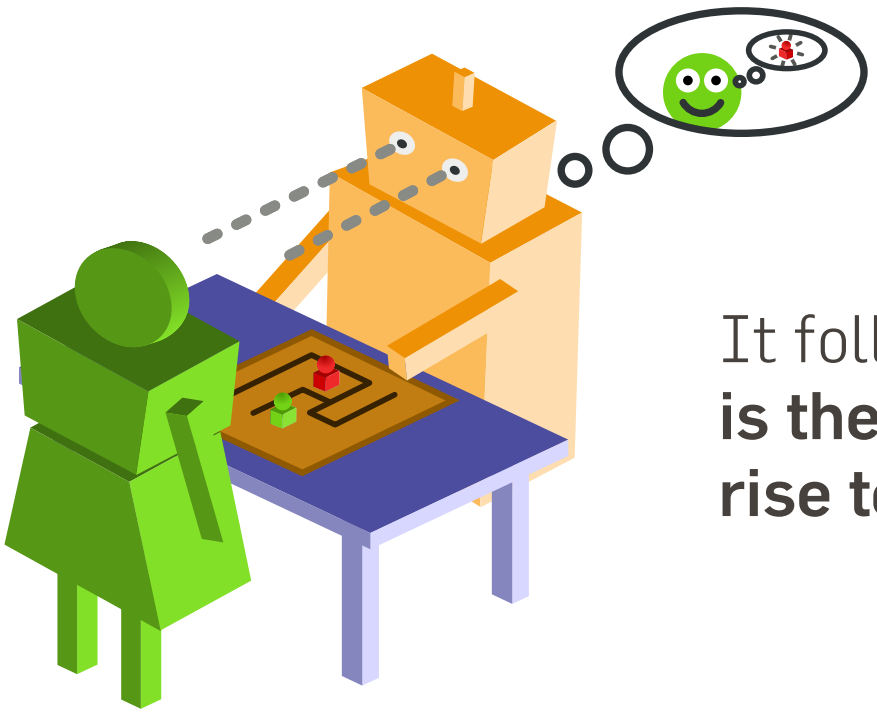


Graziano's postulate that modelling other's state of awareness is **mediated by one's own attentional system**, through joint attention

[Graziano **Consciousness and the Social Brain** – 2013]

MODELING OTHERS' MIND?

In cognitive neurosciences: Graziano's *Attention Schemata Theory*



It follows that **joint attention is the process that gives rise to social awareness**

[Graziano **Consciousness and the Social Brain** – 2013]

SKETCHING A PATH FORWARD: MENTALIZING

Hypothesis 1: Graziano is right: mental representations are snapshots of *awareness*, *awareness* being itself a label for the *memory-mediated process of attention*.

SKETCHING A PATH FORWARD: MENTALIZING

Hypothesis 1: Graziano is right: mental representations are snapshots of *awareness*, *awareness* being itself a label for the *memory-mediated process of attention*.

Hypothesis 2: this can be extended to social cognition. **Modeling one other mental representations equates to taking snapshots of their current state of awareness**.

As we do not have direct access to others' process of attention, it has to be mediated. Following Graziano, we hypotesise that **modelling other's state of awareness is mediated by one's own attentional system, through joint attention mechanisms**.

IN MORE DETAILS

1. **mental representations** are **snapshots of what we are aware of**

[Graziano's Attention Schemata Theory; Desimone and Ducan **Biased Competition Model of Attention** 1995; Block 1996]

IN MORE DETAILS

1. **mental representations** are **snapshots of what we are aware of**
2. **awareness** is the label we conveniently put on the **process of attention**

[Graziano's Attention Schemata Theory; Desimone and Ducan **Biased Competition Model of Attention** 1995; Block 1996]

IN MORE DETAILS

1. **mental representations** are **snapshots of what we are aware of**
2. **awareness** is the label we conveniently put on the **process of attention**
3. attention at time t is the label we put on the set of the **activated units** in a (biased) **associative memory network**

[Graziano's Attention Schemata Theory; Desimone and Ducan **Biased Competition Model of Attention** 1995; Block 1996]

IN MORE DETAILS

1. **mental representations** are **snapshots of what we are aware of**
2. **awareness** is the label we conveniently put on the **process of attention**
3. attention at time t is the label we put on the set of the **activated units** in a (biased) **associative memory network**
4. modelling **others' mental representations** is taking snapshots of their own current state of awareness

[Graziano's Attention Schemata Theory; Desimone and Ducan **Biased Competition Model of Attention** 1995; Block 1996]

IN MORE DETAILS

1. **mental representations** are **snapshots of what we are aware of**
2. **awareness** is the label we conveniently put on the **process of attention**
3. attention at time t is the label we put on the set of the **activated units** in a (biased) **associative memory network**
4. modelling **others' mental representations** is taking snapshots of their own current state of awareness
5. modelling other's state of awareness, i.e. their current attentional process, is **mediated by one's own attentional system**, typically through **joint attention** mechanisms

[Graziano's Attention Schemata Theory; Desimone and Ducan **Biased Competition Model of Attention** 1995; Block 1996]

IN MORE DETAILS

1. **mental representations** are **snapshots of what we are aware of**
2. **awareness** is the label we conveniently put on the **process of attention**
3. attention at time t is the label we put on the set of the **activated units** in a (biased) **associative memory network**
4. modelling **others' mental representations** is taking snapshots of their own current state of awareness
5. modelling other's state of awareness, i.e. their current attentional process, is **mediated by one's own attentional system**, typically through **joint attention** mechanisms
6. Points 1 to 5 essentially refer to a *phenomenal* awareness (a *raw* inner experience). *Phenomenal* awareness can be turned into *access consciousness* (the abstract, cognitive ability to reflect on the inner experience)

[Graziano's Attention Schemata Theory; Desimone and Ducan **Biased Competition Model of Attention** 1995; Block 1996]

IN MORE DETAILS

1. **mental representations** are **snapshots of what we are aware of**
2. **awareness** is the label we conveniently put on the **process of attention**
3. attention at time t is the label we put on the set of the **activated units** in a (biased) **associative memory network**
4. modelling **others' mental representations** is taking snapshots of their own current state of awareness
5. modelling other's state of awareness, i.e. their current attentional process, is **mediated by one's own attentional system**, typically through **joint attention** mechanisms
6. Points 1 to 5 essentially refer to a *phenomenal* awareness (a *raw* inner experience). *Phenomenal* awareness can be turned into *access consciousness* (the abstract, cognitive ability to reflect on the inner experience)
7. In AI, *phenomal awareness* maps to connectionist approaches, while *access consciousness* maps to **symbolic representations**

[Graziano's Attention Schemata Theory; Desimone and Ducan **Biased Competition Model of Attention** 1995; Block 1996]

AN HYBRID MODEL OF COGNITION

- ***phenomenal experience*** modelled in a connectionist (sub-symbolic) fashion (associative memory network)
- ***access conciousness*** in a cognitivist (symbolic) fashion (typically, an ontology or epistemic logics)

$\Rightarrow$ bottom-up, from raw percepts to ***accessible*** representations

AN HYBRID MODEL OF COGNITION

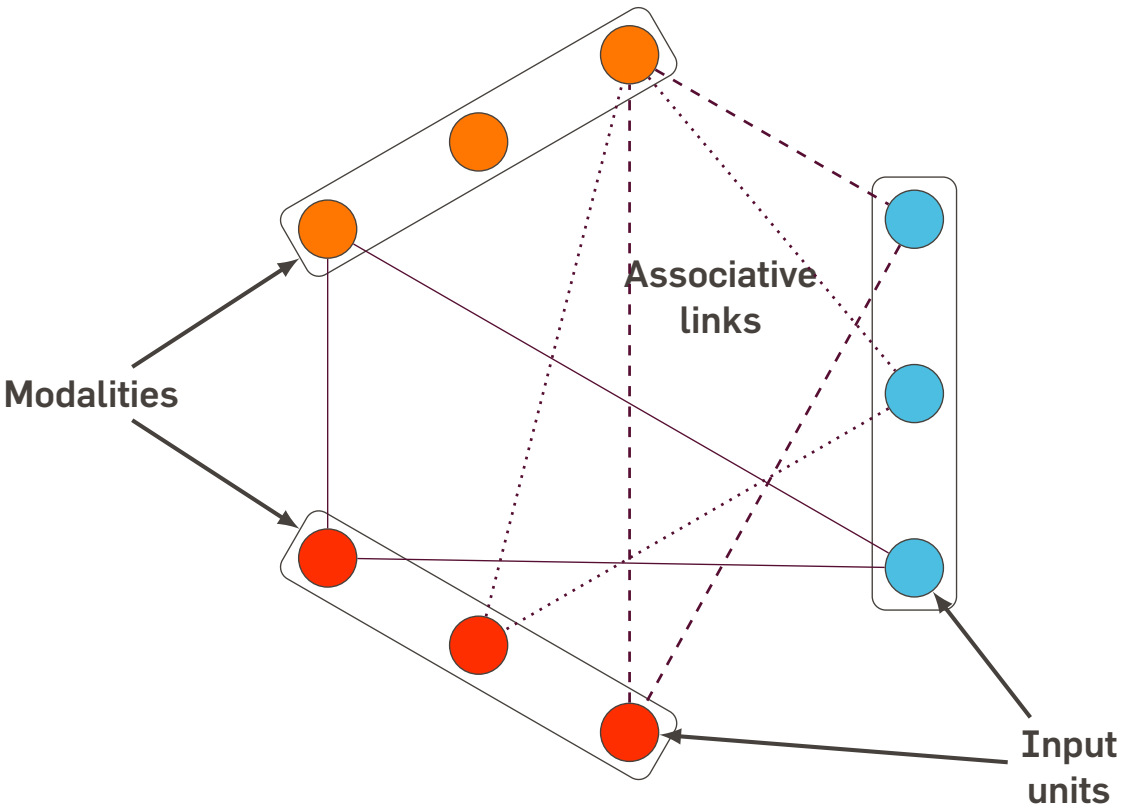
- *phenomenal experience* modelled in a connectionist (sub-symbolic) fashion (associative memory network)
- *access conciousness* in a cognitivist (symbolic) fashion (typically, an ontology or epistemic logics)

$\Rightarrow$ bottom-up, from raw percepts to *accessible* representations

The *Biased Competition Model of Attention* supports bottom-up as well as **top-down biasing mechanisms**:

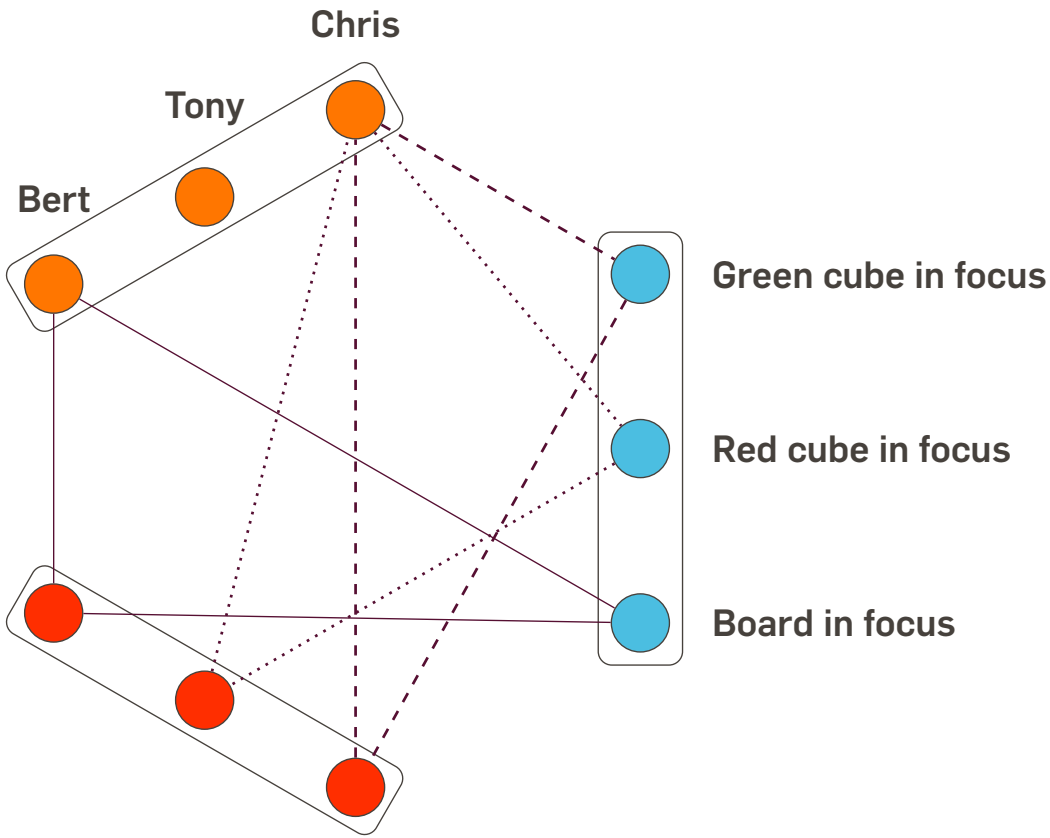
- *bottom-up*: if a unit is activated longer/stronger, it biases the resulting attention to this unit.
- *top-down*: abstract cognitive processes can influence the memory network at symbolic level to bias the attention process. Practically less clear, but also potentially very interesting as it **closes the loop between the emergent and cognitivist paradigms**

ASSOCIATIVE MEMORY NETWORK



[Baxter et al., Cognitive Architecture for HRI: Towards Behavioural Alignment, 2013]

ASSOCIATIVE MEMORY NETWORK



[Baxter et al., Cognitive Architecture for HRI: Towards Behavioural Alignment, 2013]

MANY GAPS CALLING FOR INVESTIGATION!

- mental representations as instantaneous 'snapshots' is a **strong assumption** (and what about representation of lasting situations?)

MANY GAPS CALLING FOR INVESTIGATION!

- mental representations as instantaneous 'snapshots' is a **strong assumption** (and what about representation of lasting situations?)
- if attention is the set of activated units in a memory network, **what are the units?** physical entities? social events? situations? what is the right level of abstraction: from raw perceptual inputs to high-level units like objects, joint gazing, etc.

MANY GAPS CALLING FOR INVESTIGATION!

- mental representations as instantaneous 'snapshots' is a **strong assumption** (and what about representation of lasting situations?)
- if attention is the set of activated units in a memory network, **what are the units?** physical entities? social events? situations? what is the right level of abstraction: from raw perceptual inputs to high-level units like objects, joint gazing, etc.
- **the mapping** between 'phenomenal consciousness' and 'access consciousness' and, respectively, sub-symbolic and symbolic computational structures **might be naive**. This need to be better evidenced, if only because the frontier between phenomenal and access consciousness is all but clear (as pointed by Graziano).

MANY GAPS CALLING FOR INVESTIGATION!

- mental representations as instantaneous 'snapshots' is a **strong assumption** (and what about representation of lasting situations?)
- if attention is the set of activated units in a memory network, **what are the units?** physical entities? social events? situations? what is the right level of abstraction: from raw perceptual inputs to high-level units like objects, joint gazing, etc.
- **the mapping** between 'phenomenal consciousness' and 'access consciousness' and, respectively, sub-symbolic and symbolic computational structures **might be naive**. This need to be better evidenced, if only because the frontier between phenomenal and access consciousness is all but clear (as pointed by Graziano).
- what is the **social motivation** for the robot to carry over this modeling? What social drives?

MANY GAPS CALLING FOR INVESTIGATION!

- mental representations as instantaneous 'snapshots' is a **strong assumption** (and what about representation of lasting situations?)
- if attention is the set of activated units in a memory network, **what are the units?** physical entities? social events? situations? what is the right level of abstraction: from raw perceptual inputs to high-level units like objects, joint gazing, etc.
- **the mapping** between 'phenomenal consciousness' and 'access consciousness' and, respectively, sub-symbolic and symbolic computational structures **might be naive**. This need to be better evidenced, if only because the frontier between phenomenal and access consciousness is all but clear (as pointed by Graziano).
- what is the **social motivation** for the robot to carry over this modeling? What social drives?
- at epistemic level, if access to other's mental representations is *mediated* by one's own attentional system, these mental representations are subjective. **Can we equate humans' and robots' subjectivities?**

SKETCHING A PATH FORWARD: SOCIAL BEHAVIOURS

Hypothesis 3: together, representations of one's and others' minds are *necessary* and *sufficient* for social behaviours to emerge

Hypothesis 3': representations of one's and others' minds, along with a situation (physical and temporal environment), are *necessary* and *sufficient* conditions for social behaviours to emerge

SKETCHING A PATH FORWARD: SOCIAL BEHAVIOURS

Hypothesis 3: together, representations of one's and others' minds are *necessary* and *sufficient* for social behaviours to emerge

Hypothesis 3': representations of one's and others' minds, along with a situation (physical and temporal environment), are *necessary* and *sufficient* conditions for social behaviours to emerge

THE FREEPLAY SANDBOX

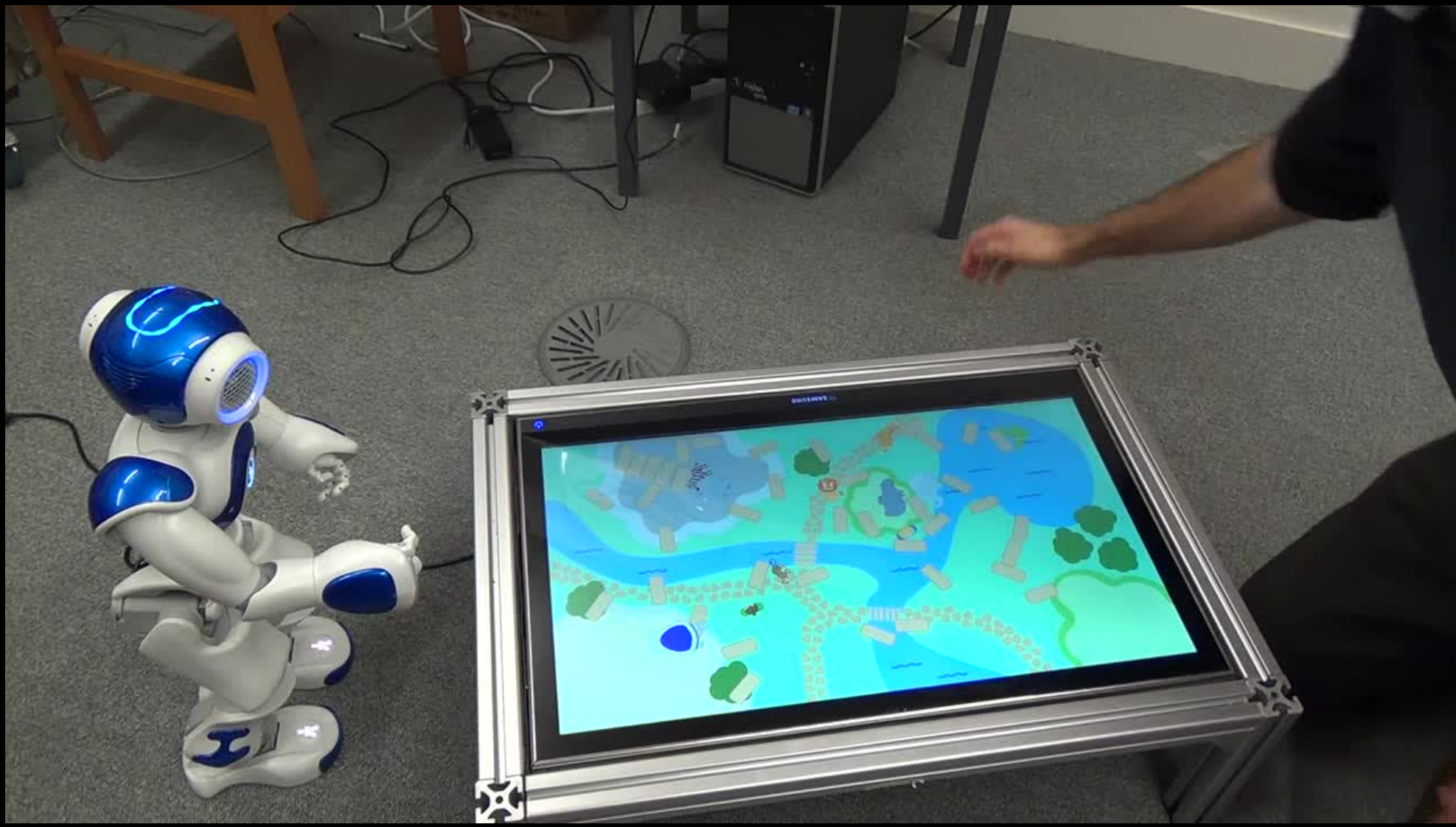
AN EXPERIMENTAL FRAMEWORK

- The “Zoo design” play situation
- **Free play** with the following constraints:

  initial prompt (“Let’s build a zoo!”)
  limited set of tokens (cubes, Lego animals)
  spatially limited playground

AN EXPERIMENTAL FRAMEWORK

- The “Zoo design” play situation
- **Free play** with the following constraints:
  - initial prompt (“Let’s build a zoo!”)
  - limited set of tokens (cubes, Lego animals)
  - spatially limited playground
- to make it technically tractable with robots, the physical playground is **replaced by a large touchscreen** (sandtray): entirely skips the difficult problem of perception and manipulation in a dense & cluttered scene
- the touchscreen strictly replace the perception of objects on the playground (exports ROS TF frames of each object) and their manipulation (receives virtual ’touches’ from the robot)
- importantly, perception of the partner and of the global scene geometry is genuine



AN EXPERIMENTAL FRAMEWORK

Open-ended task: more an **experimental framework** than a task.

- free play, yet sufficiently well-defined to be reproducible
- focus on abstract socio-cognitive facets (perception is simplified; manipulation is mostly avoided)

Besides, well suited for interaction analysis, with tools like:

- behavioural alignment between partners: for instance, using Słowinski's *Individual Motor Signature*
- Ballard's (and Anderson's extension) coding of children's free-play interactions
- *With-me-ness* as a metric of co-engagment

[Lemaignan et al., **From Real-time Attention Assessment to "With-me-ness" in Human-Robot Interaction** HRI 2016]

ONLY THE START OF IT!

Which cognitive model? which cognitive architecture?

$\rightarrow$ will likely draw from hybrid architectures (CLARION), internal simulation (HAMMER), sub-symbolic cognitive architecture (ERA)

...but not many cognitive architectures model social interactions! (on BICA website, about 0 actually!)

ONLY THE START OF IT!

Which cognitive model? which cognitive architecture?

→ will likely draw from hybrid architectures (CLARION), internal simulation (HAMMER), sub-symbolic cognitive architecture (ERA)

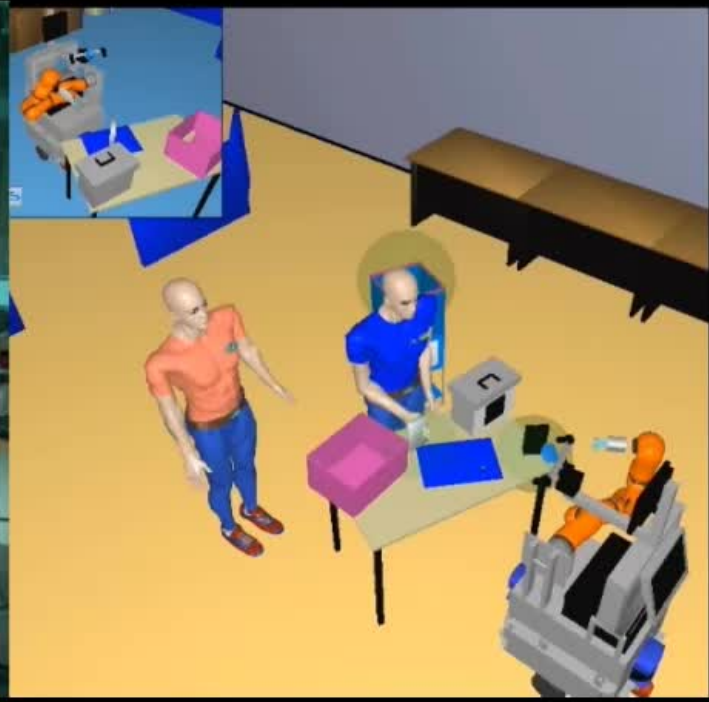
...but not many cognitive architectures model social interactions! (on BICA website, about 0 actually!)

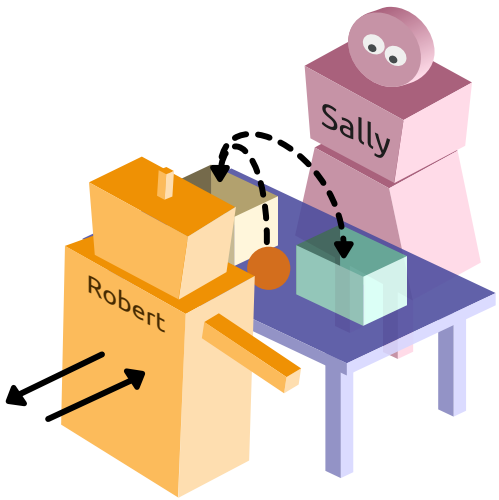
What inputs for a connectionist take on social interactions? low-level? high-level? To reconstruct someone else's attentional state, Graziano suggests:

- gaze direction
- facial expression
- body language
- prior knowledge of person
- location of salient objects

Probably not the end of it, though!

THEORY OF MIND

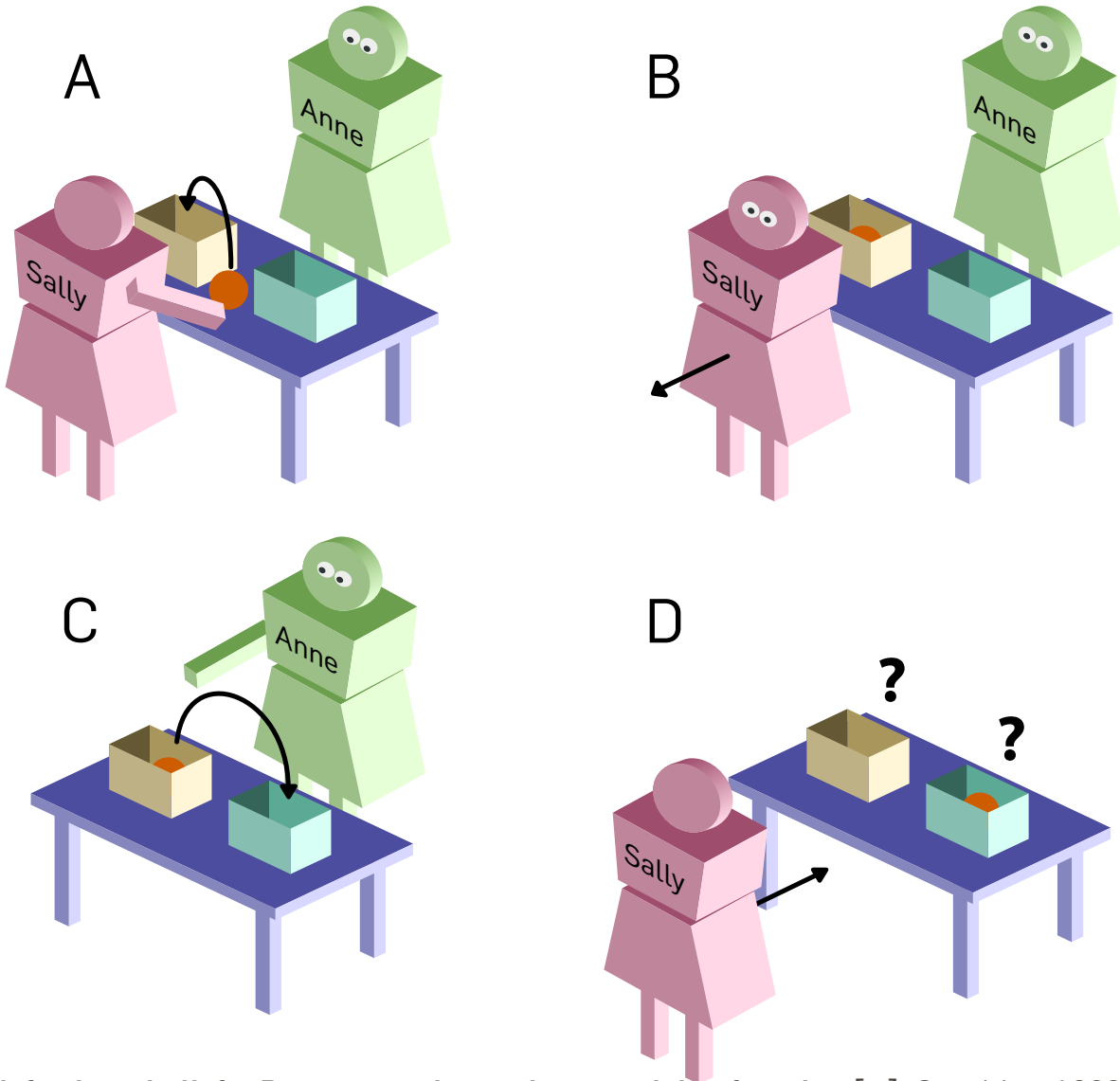




$$\begin{array}{ccc} \mathcal{M}^\circ(Sally, R, X) & \xleftrightarrow{\quad\Delta_4\quad} & \mathcal{M}^\circ(R, R, X) \\ \Big\updownarrow \Delta_5 & & \Big\updownarrow \\ \mathcal{M}^\circ(Sally, R, Y) & \longleftrightarrow & \mathcal{M}^\circ(R, R, Y) \end{array}$$

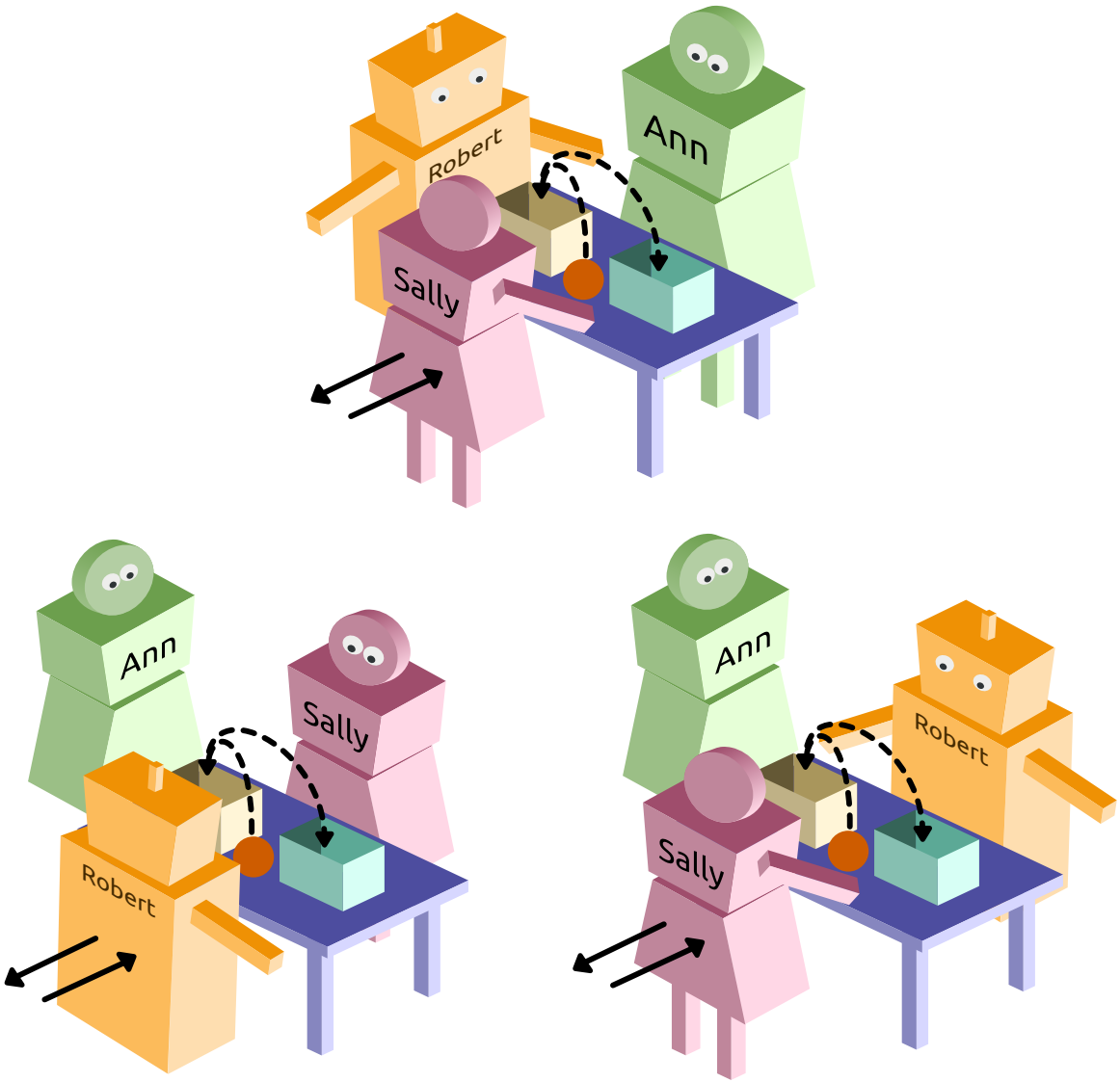
”Theory of Mind tasks”

1ST ORDER TOM: THE FALSE-BELIEF EXPERIMENT



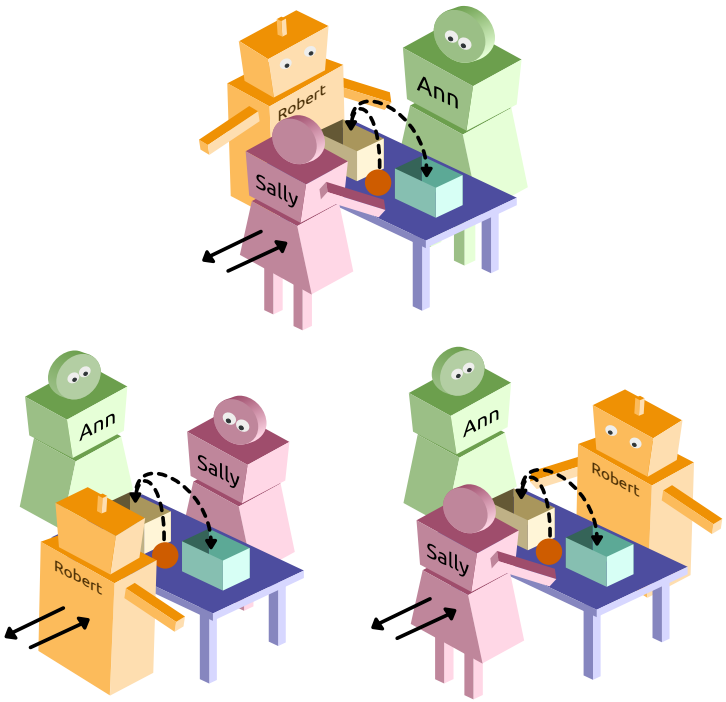
[Wimmer and Perner, **Beliefs about beliefs: Representation and constraining function [...]**, Cognition, 1983]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

THE FALSE-BELIEF EXPERIMENT, RELOADED



[Dillenbourg, Lemaignan et al., **The Symmetry of Partner Modelling**, Intl. J. of Computer Supported Collaborative Learning 2016] 120

THE FALSE-BELIEF EXPERIMENT, RELOADED



- $\mathcal{M}(A, B, X)$
- $\mathcal{M}^{\circ}(A, B, X)$

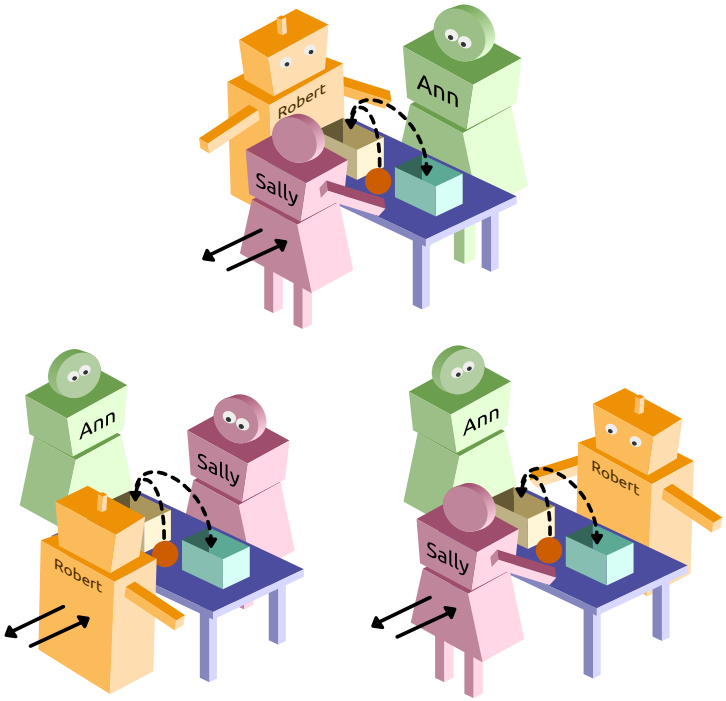
e.g. $\mathcal{M}(\text{robot}, \text{Sally}, \textit{plans})$

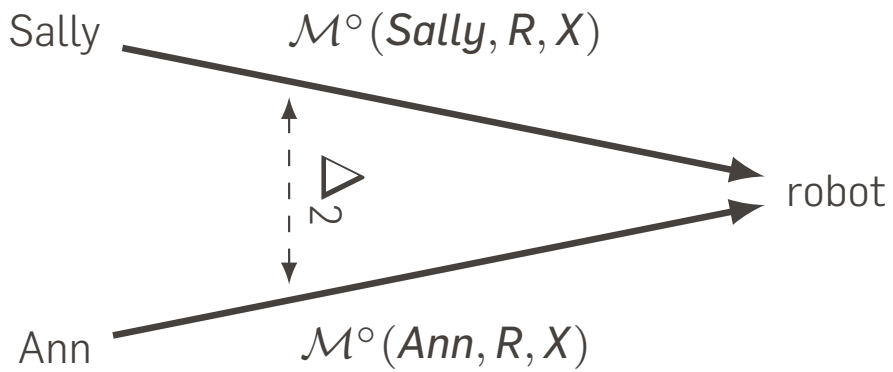
[Dillenbourg, Lemaignan et al., **The Symmetry of Partner Modelling**, Intl. J. of Computer Supported Collaborative Learning 2016]

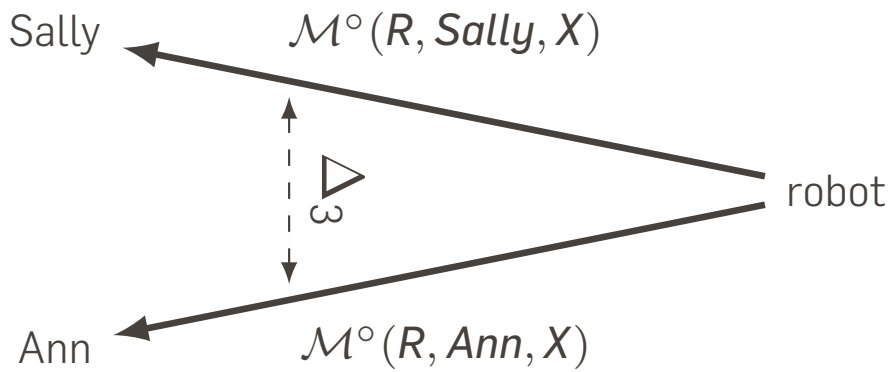
THE FALSE-BELIEF EXPERIMENT, RELOADED

- **Robot is the observer**
  $\mathcal{M}^{\circ}(R, \text{Sally}|\text{Ann}, \textit{plans})$? can the human verbalise it? i.e. $\mathcal{M}(H, R, \mathcal{M}(R, H, \textit{plans}))$?
- **Robot is an active participant**
  $\mathcal{M}(H, R, \textit{knowledge}|\textit{plans}|\textit{goals})$? i.e. How Ann interprets the behaviour of a robot who moves the ball from the beige box to the blue box while Sally is away?

THE FALSE-BELIEF EXPERIMENT, RELOADED







THE FALSE-BELIEF EXPERIMENT, RELOADED

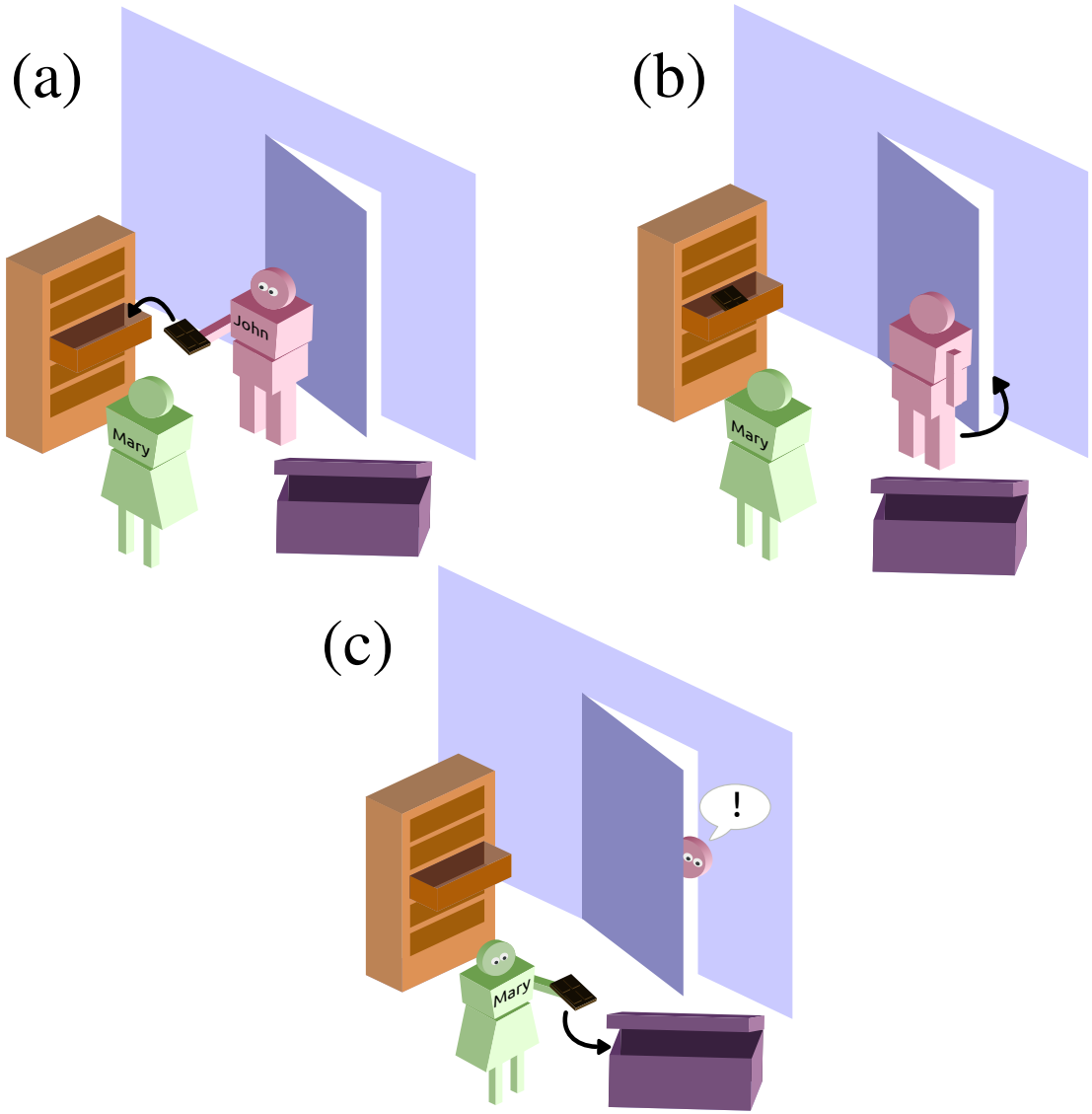
Do Sally and Ann have the same accuracy when modelling the robot?

$$\Delta_2 = \Delta(\mathcal{M}(\text{Sally}, R, X), \mathcal{M}(\text{Ann}, R, X))$$

Conversely, what may lead the robot to model more accurately Sally or Ann?

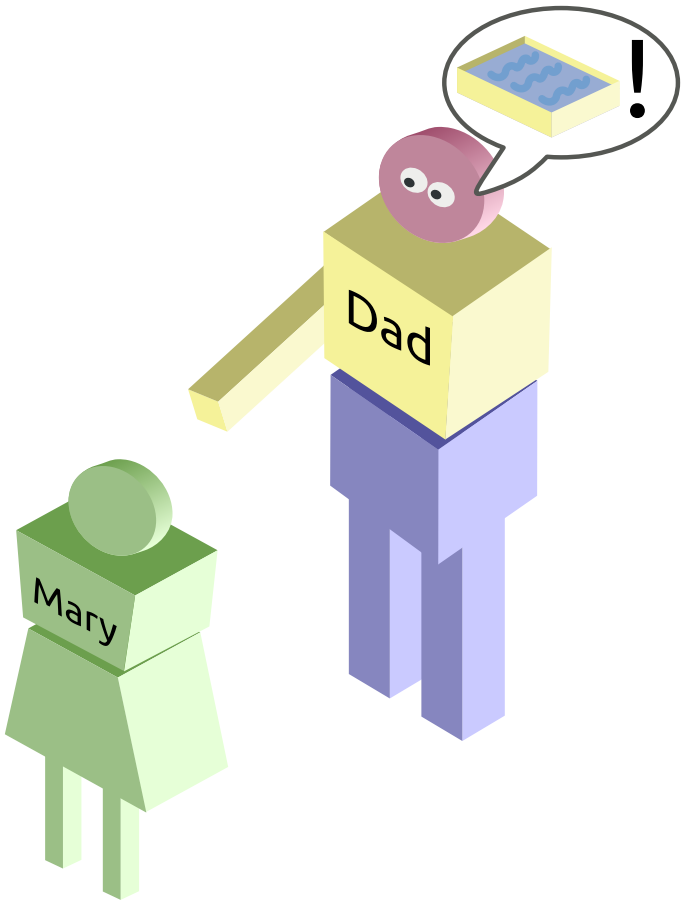
$$\Delta_3 = \Delta(\mathcal{M}(R, \text{Sally}, X), \mathcal{M}(R, \text{Ann}, X))$$

2ND ORDER TOM: THE CHOCOLATE BAR EXPERIMENT



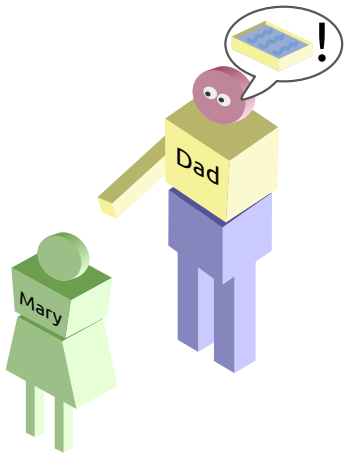
[Flobbe et al. **Children's application of theory of mind in reasoning and language**, J. of Logic, Language and Information, 2008]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

AGREEMENT AS ∞-ORDER TOM



[Verbrugge, **Logic and social cognition**, Journal of Philosophical Logic, 2009]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

AGREEMENT AS ∞-ORDER TOM

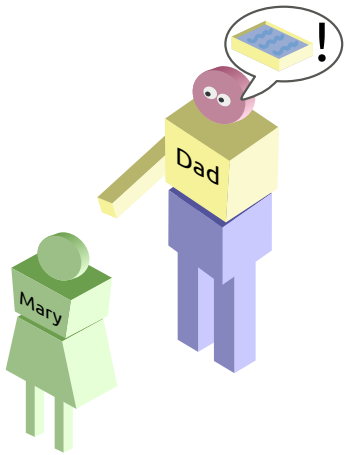


Shared knowledge

$$EK_J\varphi \leftrightarrow \bigwedge_{i \in J} K_i\varphi$$

[Verbrugge, **Logic and social cognition**, Journal of Philosophical Logic, 2009]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

AGREEMENT AS ∞-ORDER TOM



Shared knowledge

$$EK_J\varphi \leftrightarrow \bigwedge_{i \in J} K_i\varphi$$

Common knowledge

$$CK_J\varphi \leftrightarrow EK_J\varphi \wedge EK_JEK_J\varphi \wedge EK_JEK_JEK_J\varphi \wedge \ldots$$

[Verbrugge, **Logic and social cognition**, Journal of Philosophical Logic, 2009]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

SHOPPING LIST FOR HRI?

Already in the HRI fridge	**To buy...**
Instrumental gestures	Expressive gestures
Using person as tool	Using person as receiver of information
Talking about desires and emotions	Talking about beliefs and ideas
Showing "active" sociability	Showing "interactive" sociability
Elicited structured play	Spontaneous pretend play

[Frith and Happé **Autism: Beyond "theory of mind"** – Cognition, 1994]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

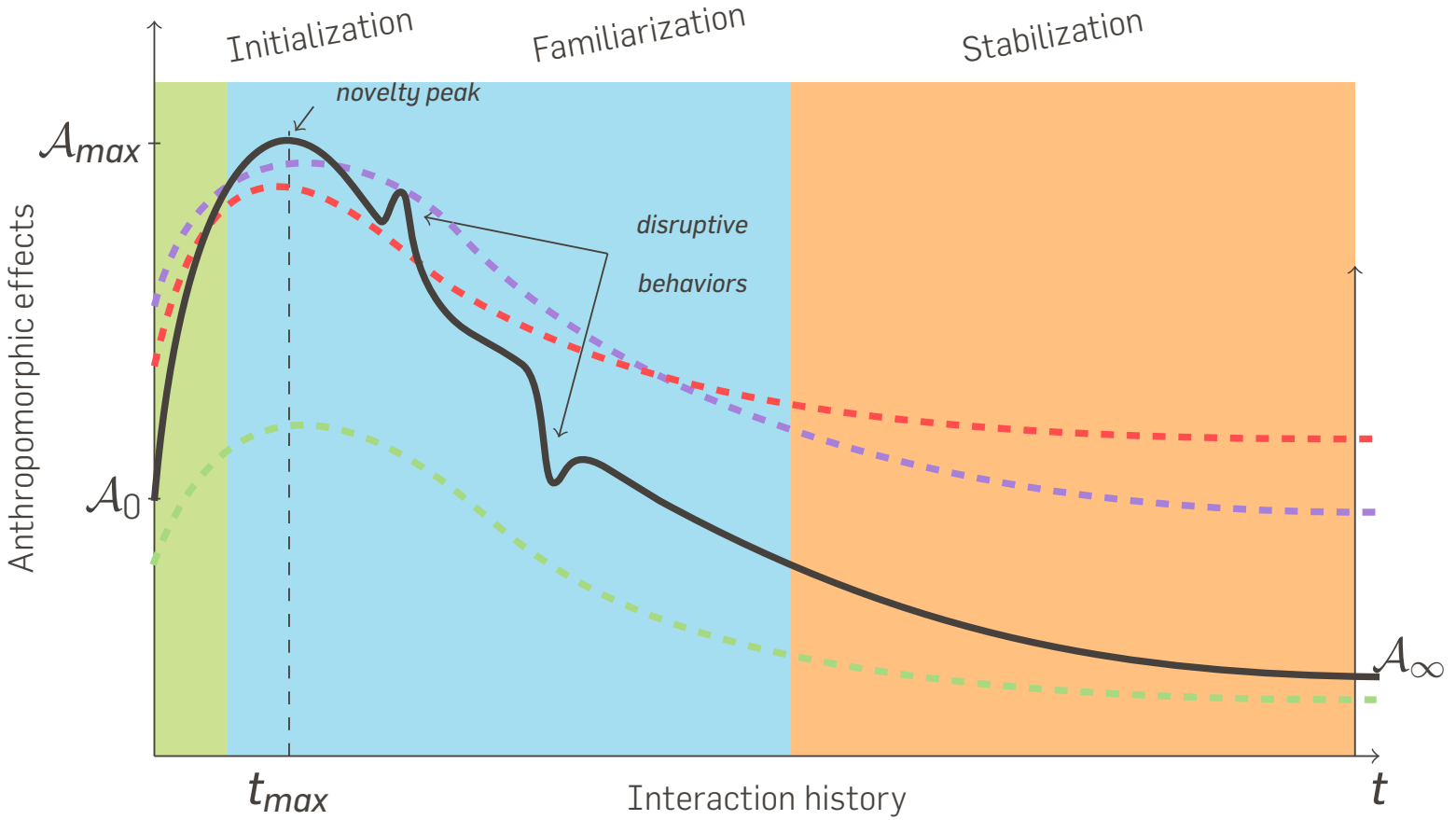
AUTISTIC ASSETS AND DEFICITS OBSERVED IN REAL LIFE

Assets	**Deficits**
Instrumental gestures	Expressive gestures
Using person as tool	Using person as receiver of information
Talking about desires and emotions	Talking about beliefs and ideas
Showing "active" sociability	Showing "interactive" sociability
Elicited structured play	Spontaneous pretend play

[Frith and Happé **Autism: Beyond "theory of mind"** – Cognition, 1994]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

DYNAMICS OF INTERACTION

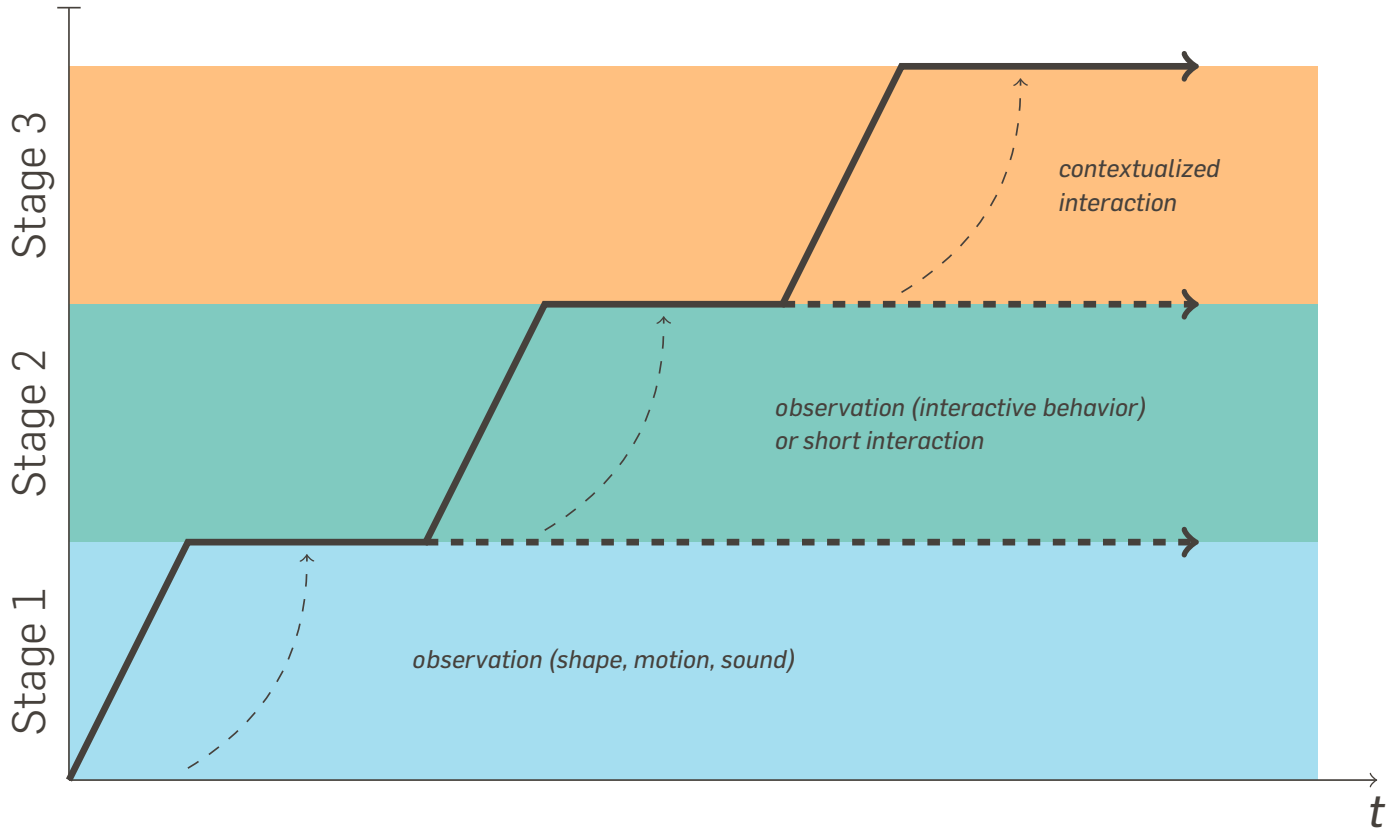
HOW DO WE PERCEIVE ROBOT OVER TIME?



◂ Supplementary material

[Lemaignan, Fink, Dillenbourg, Braboszcz **The Cognitive Correlates of Anthropomorphism**, HRI 2014]

COGNITIVE INTERPRETATION?



adapted mental model
→ adapted interaction
(anthropomorphic or not)

projection of existing mental models
(familiarity)

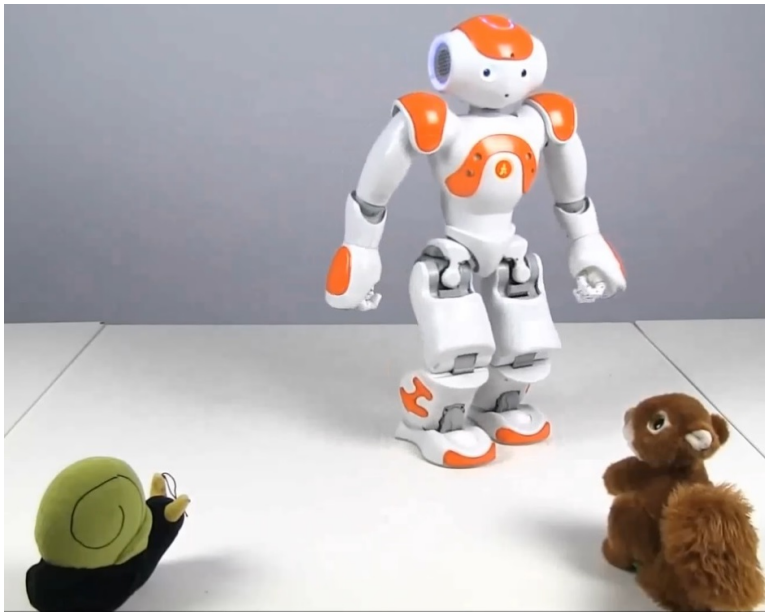
pre-cognitive anthropomorphism

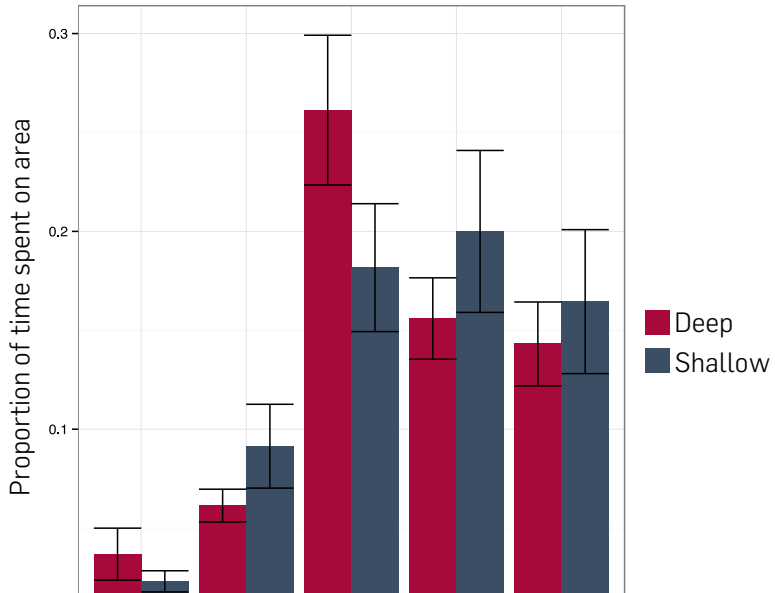
[Lemaignan, Fink, Dillenbourg, Braboszcz **The Cognitive Correlates of Anthropomorphism**, HRI 2014]

UNEXPECTED BEHAVIOURS

	Unplanned by the robot	Planned by the robot
Perceived as non-intentional	A	B
Perceived as intentional	C	D

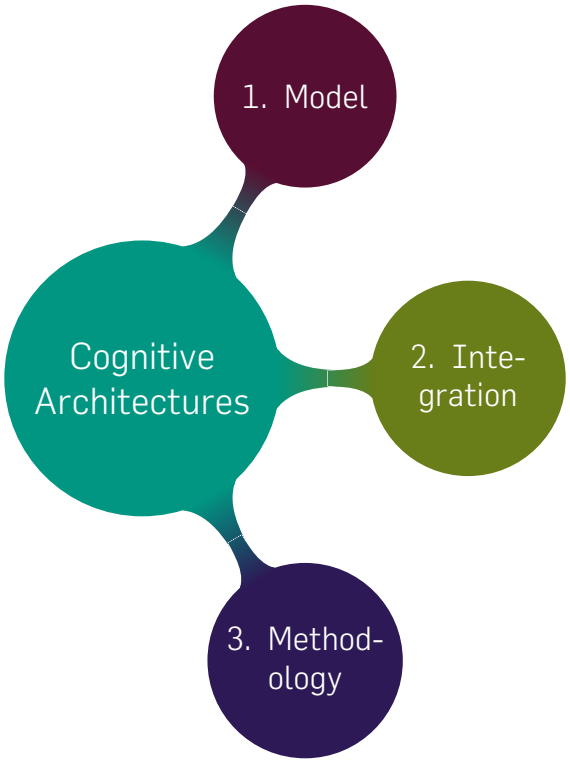
COGNITIVE CONTEXT AND ANTHROPOMORPHISM





COGNITIVE ARCHITECTURES

COGNITIVE ARCHITECTURES FOR SOCIAL HRI



1. Models of Human Cognition

– Modelling (aspects of) human cognition

– Subsequent application to robots

2. Technical Integration

– Define required functionality of robots

– Implement algorithms (etc) necessary

3. A Methodology

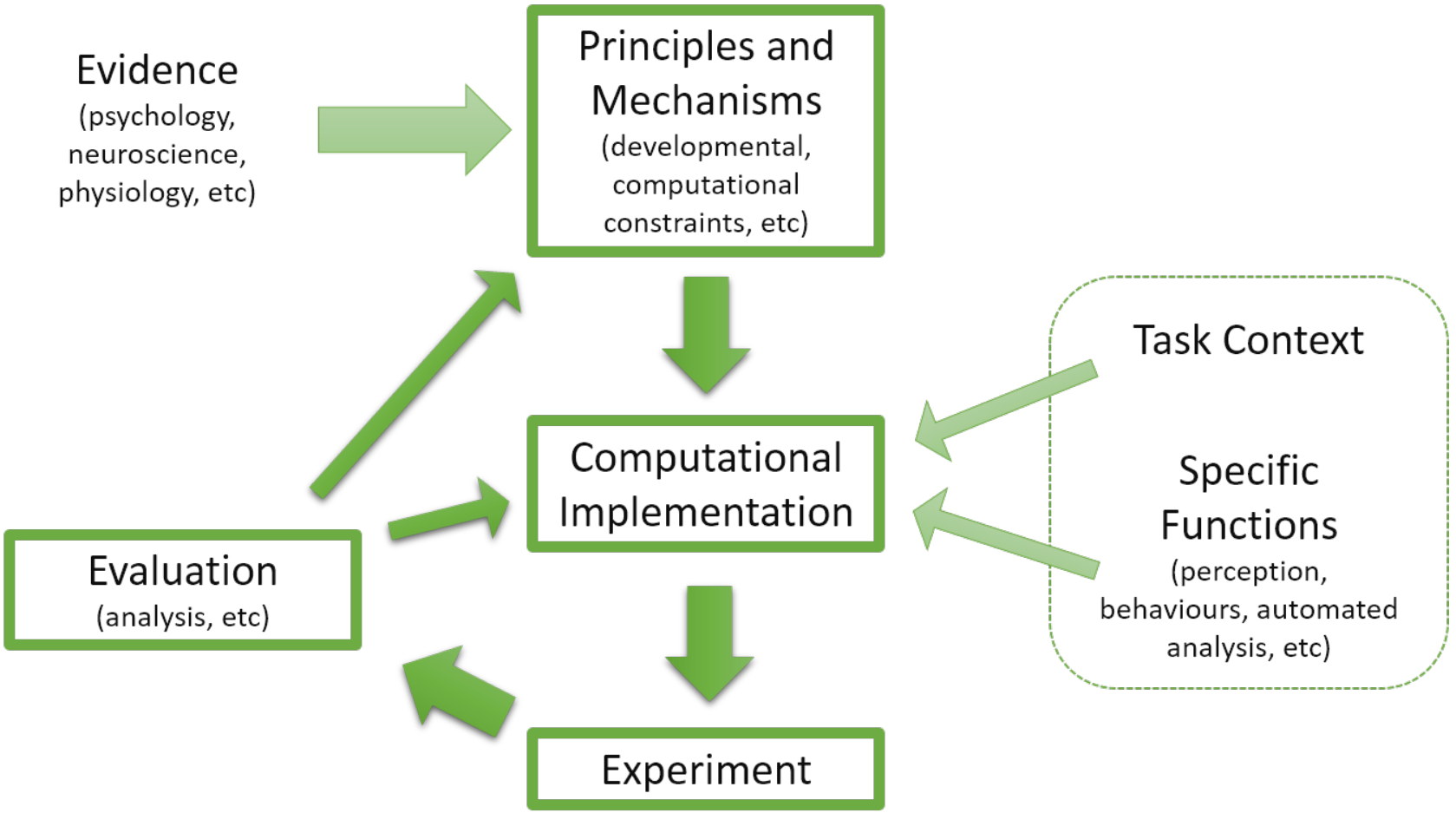
– Formalising assumptions

– Integrating knowledge from multiple disciplines

– Iteratively updating architecture

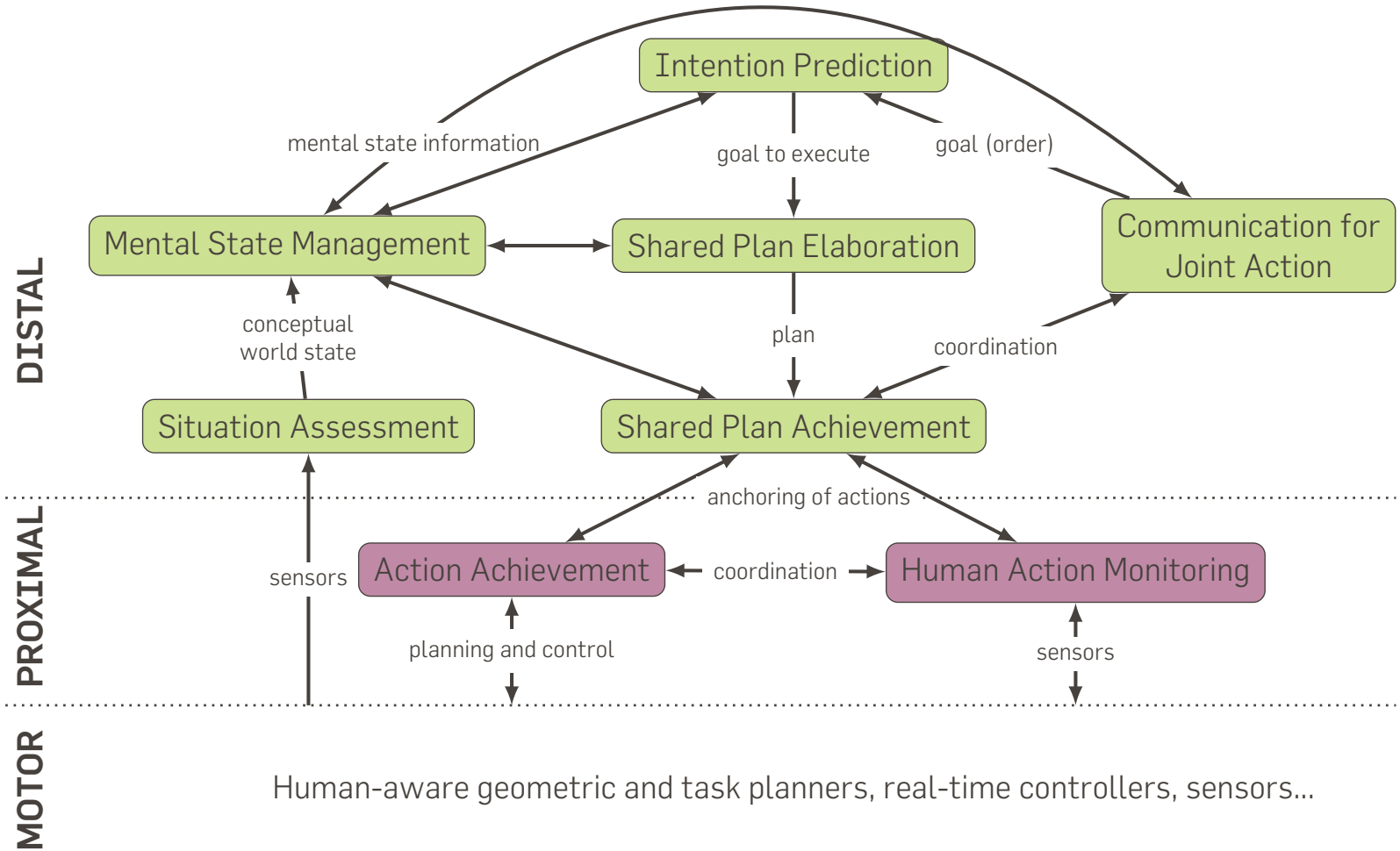
[Baxter, Lemaignan, Trafton **Workshop on Cognitive Architectures for Social HRI** – HRI 2016]

COGNITIVE ARCHITECTURE AS A METHODOLOGY



[Baxter, Lemaignan, Trafton **Workshop on Cognitive Architectures for Social Human-Robot Interaction** HRI 2016]

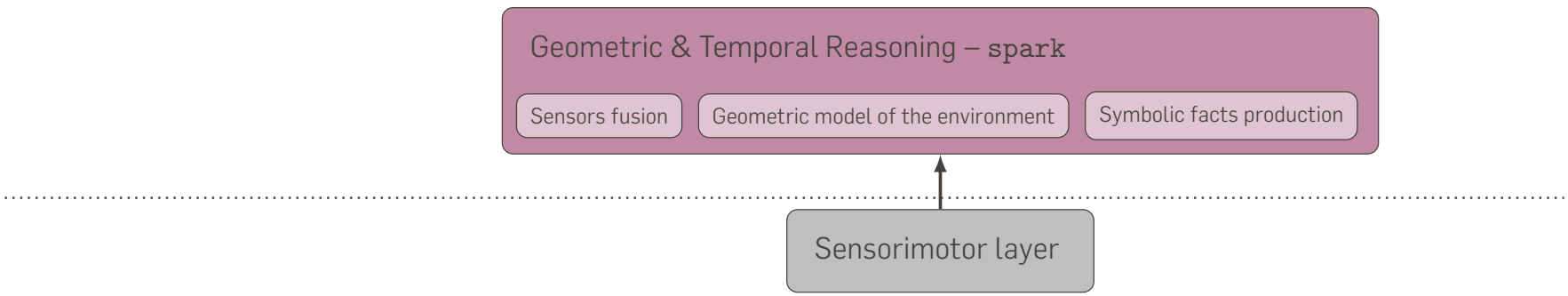
“MODELING OF HUMAN COGNITION”...



INTO AN CONTROL ARCHITECTURE

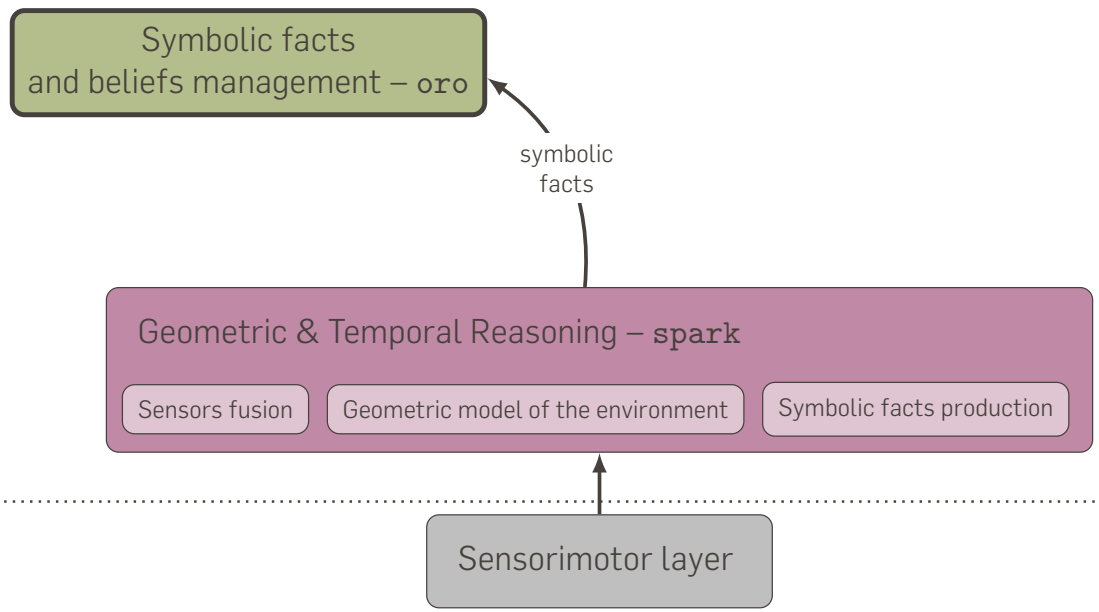
Sensorimotor layer

INTO AN CONTROL ARCHITECTURE



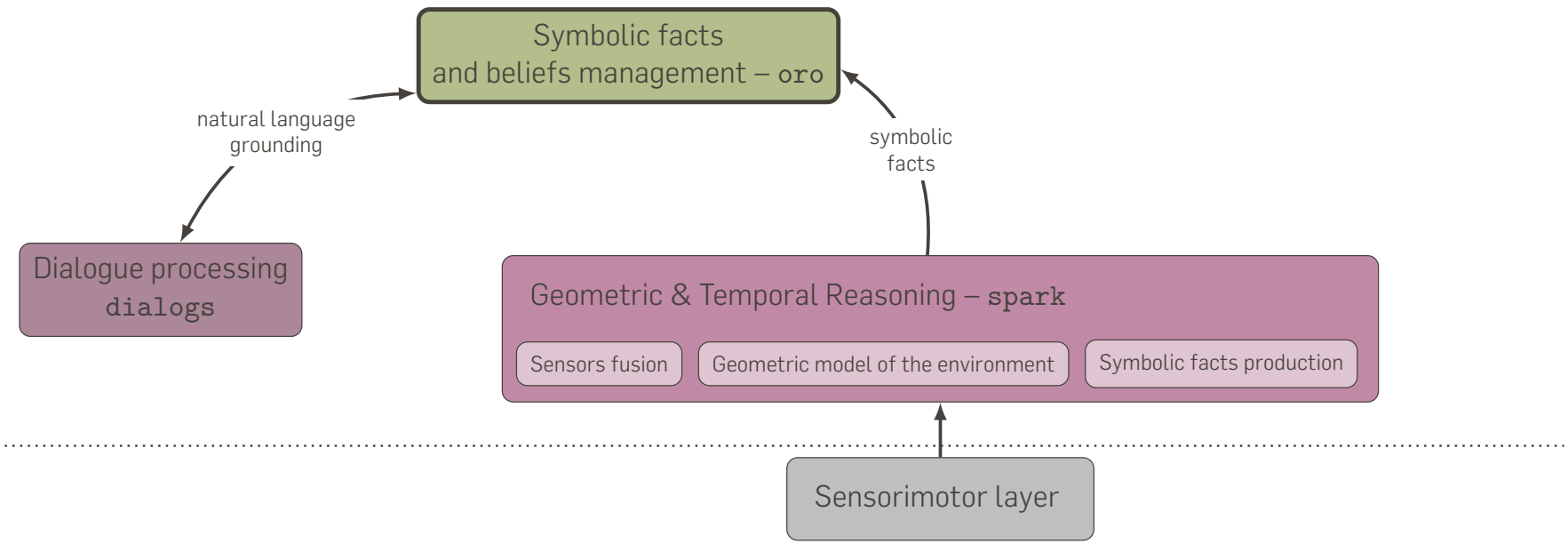
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

INTO AN CONTROL ARCHITECTURE



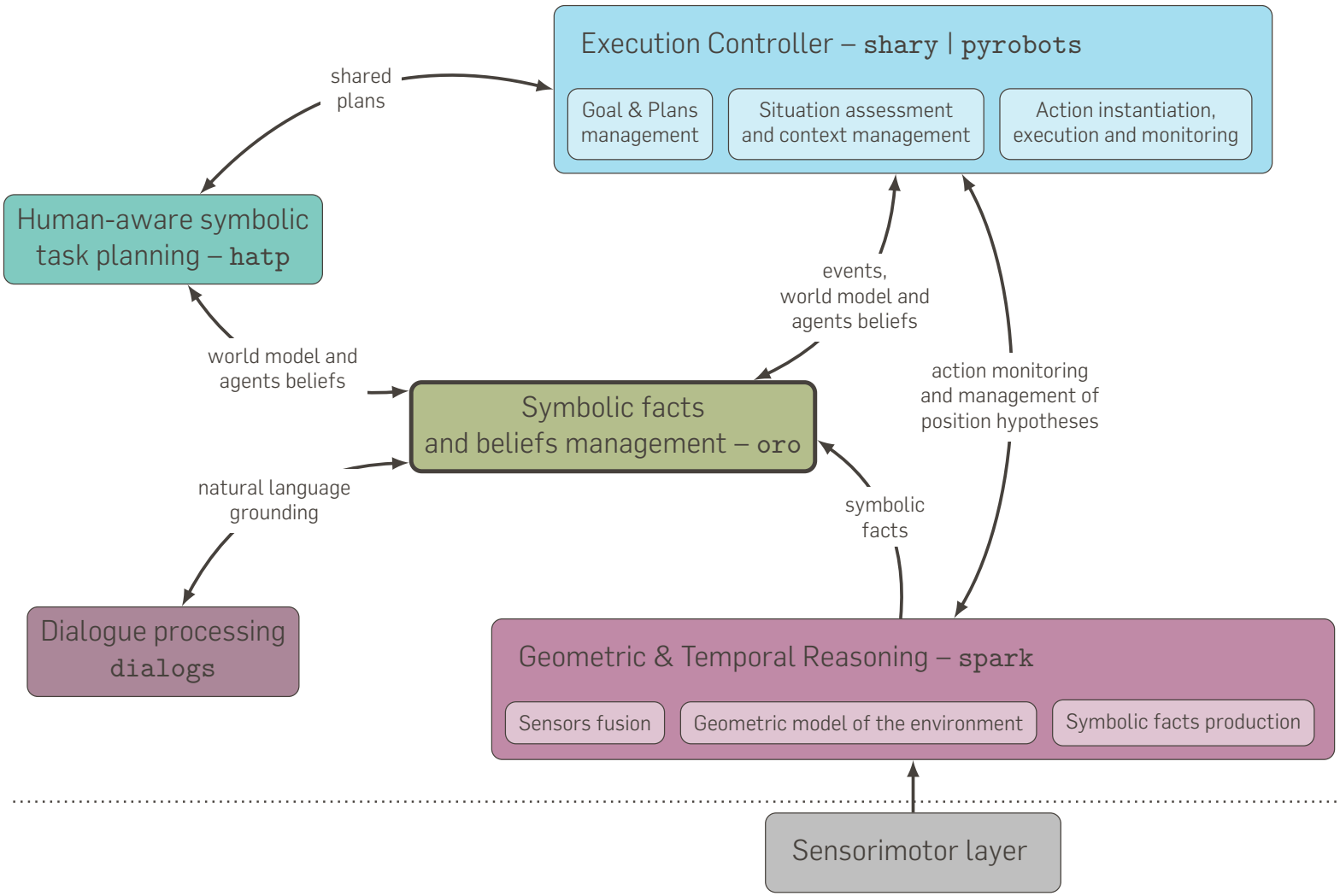
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

INTO AN CONTROL ARCHITECTURE



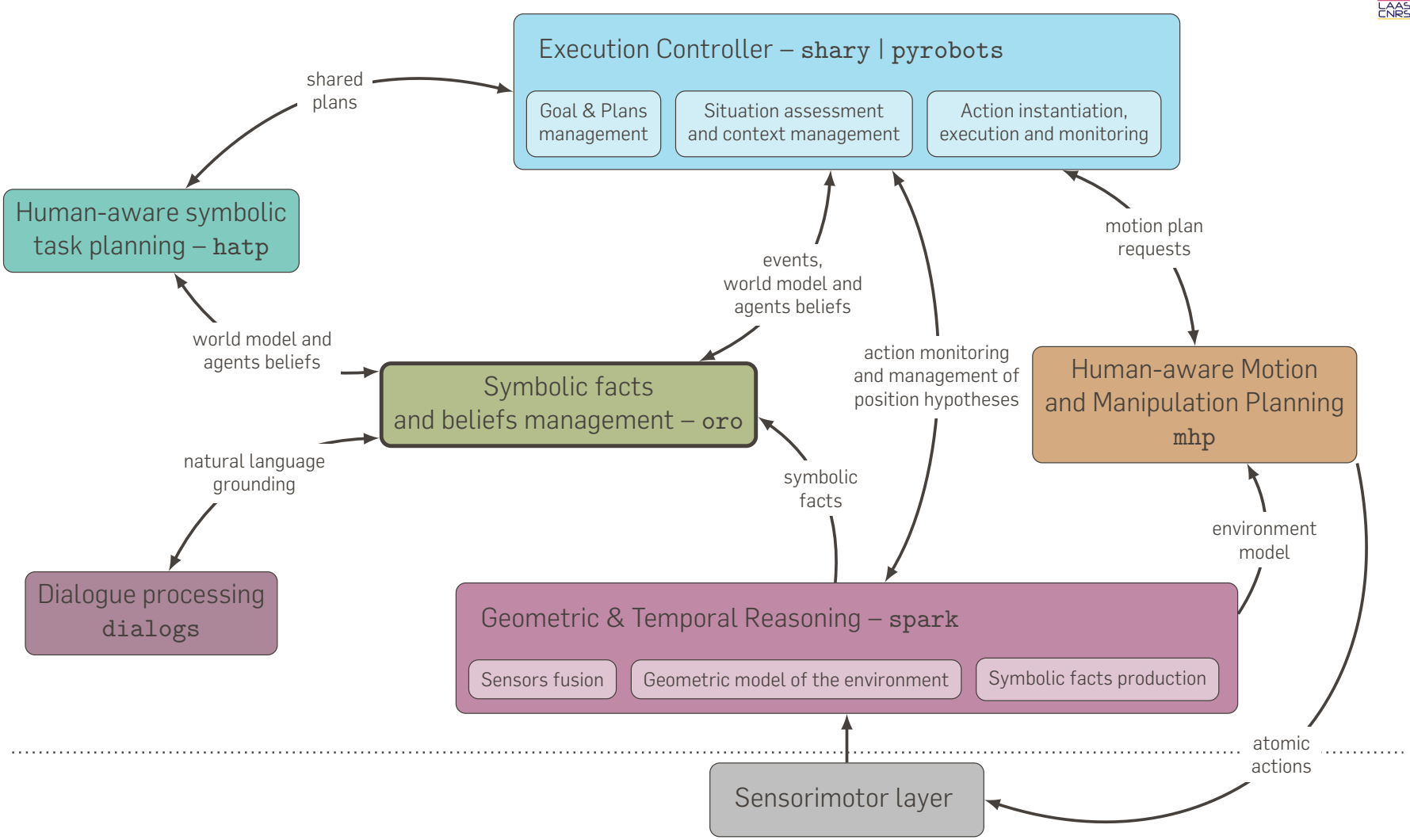
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

INTO AN CONTROL ARCHITECTURE



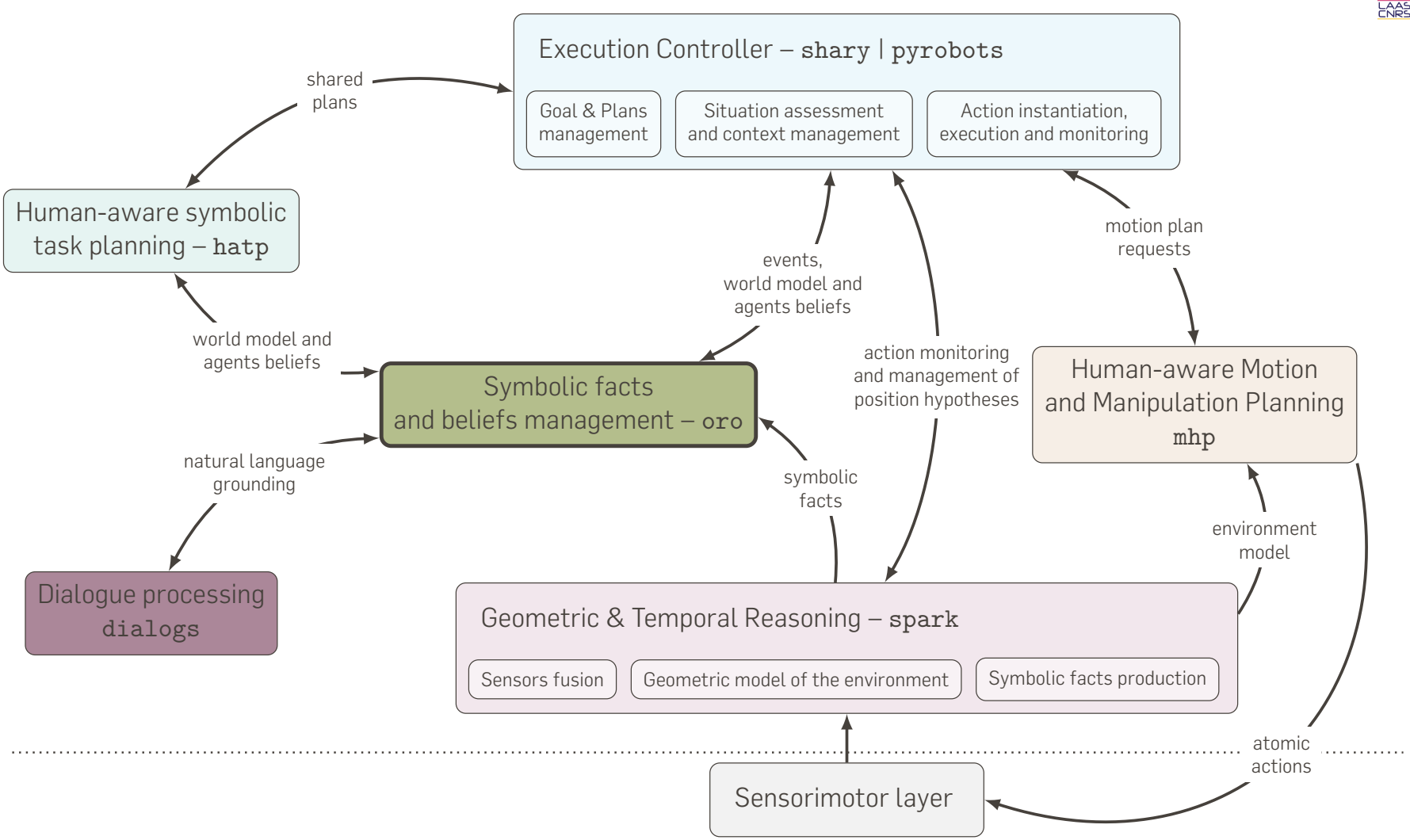
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

INTO AN CONTROL ARCHITECTURE



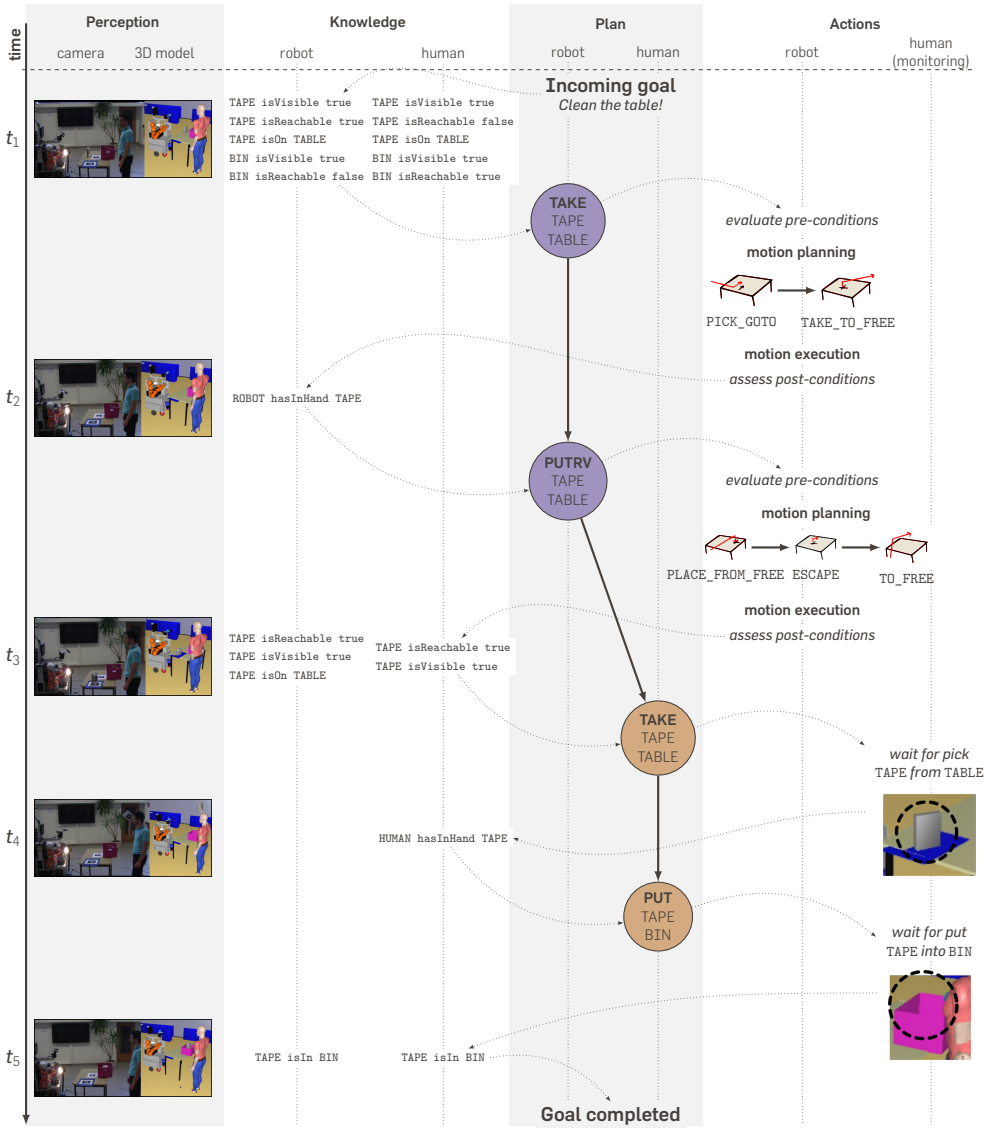
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

INTO AN CONTROL ARCHITECTURE



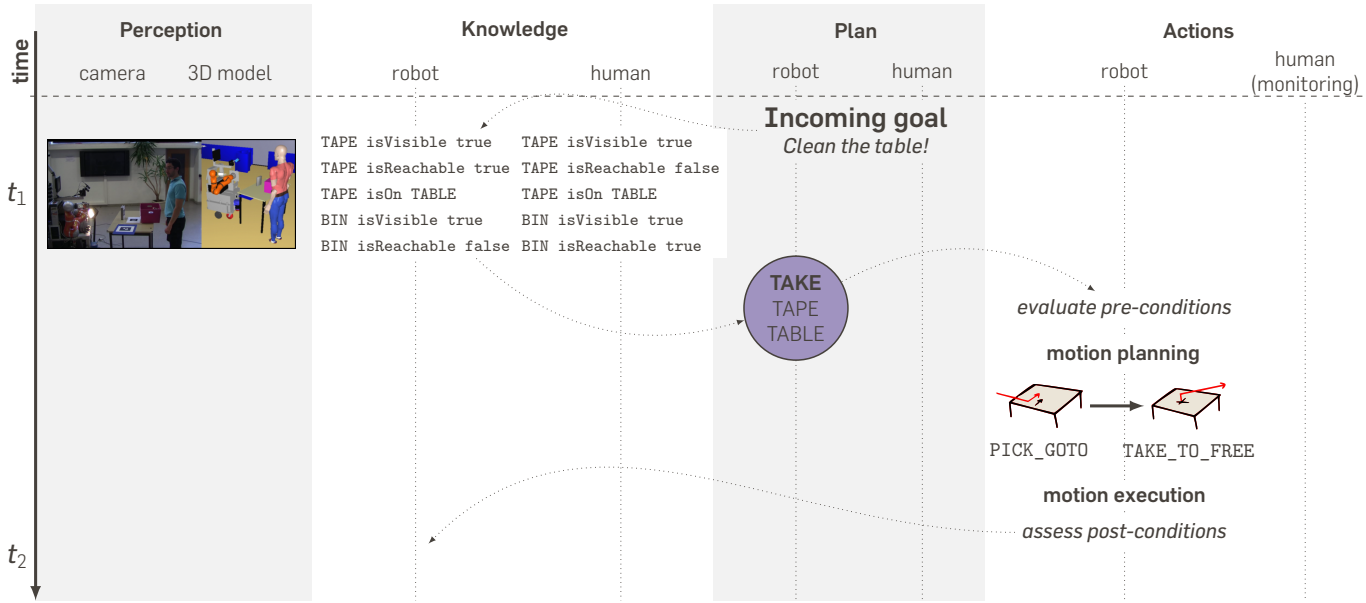
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

“CLEANING THE TABLE”…



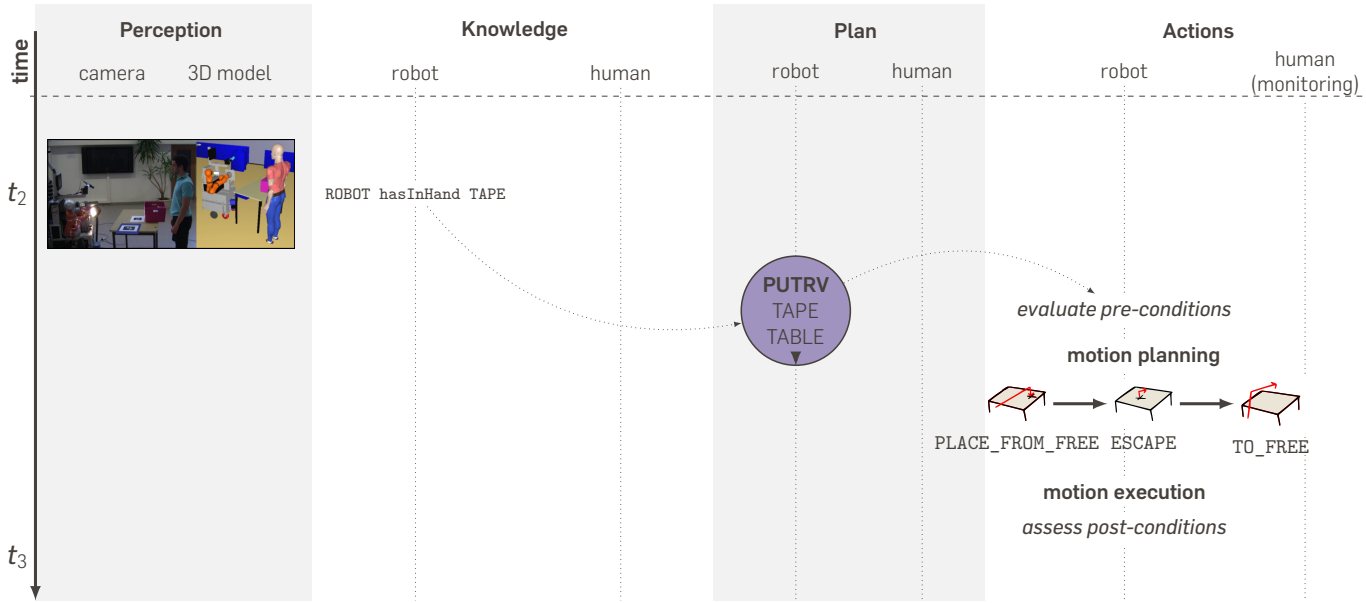
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

“CLEANING THE TABLE”...

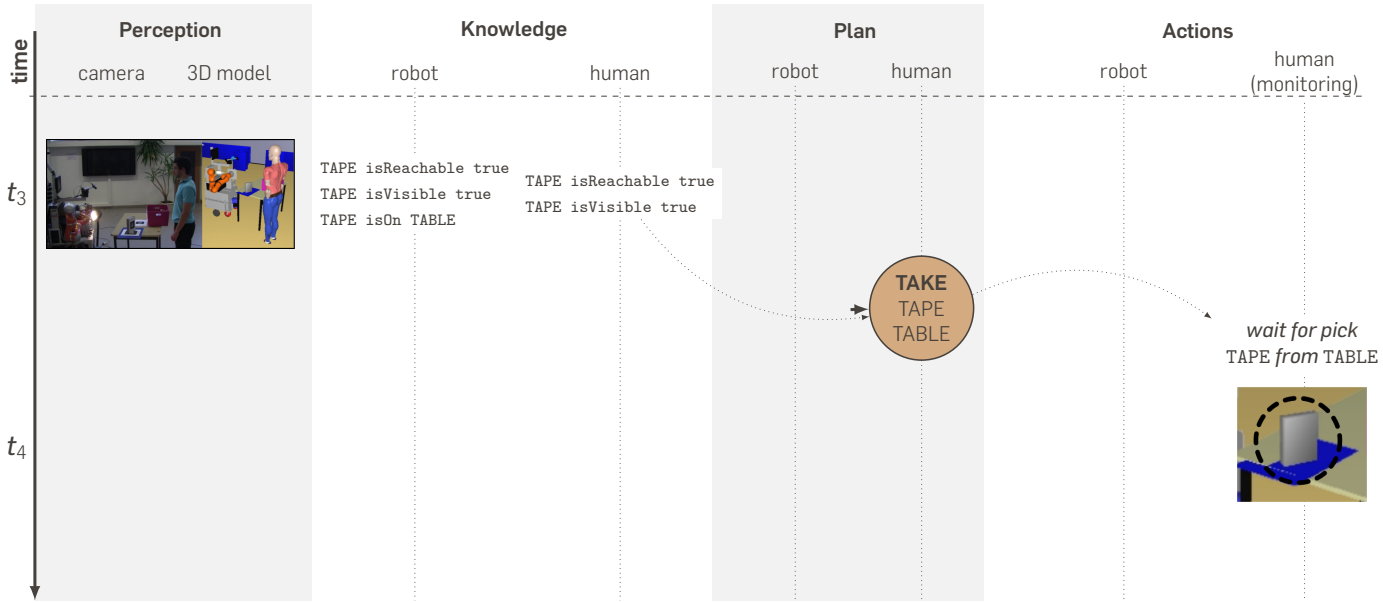


[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

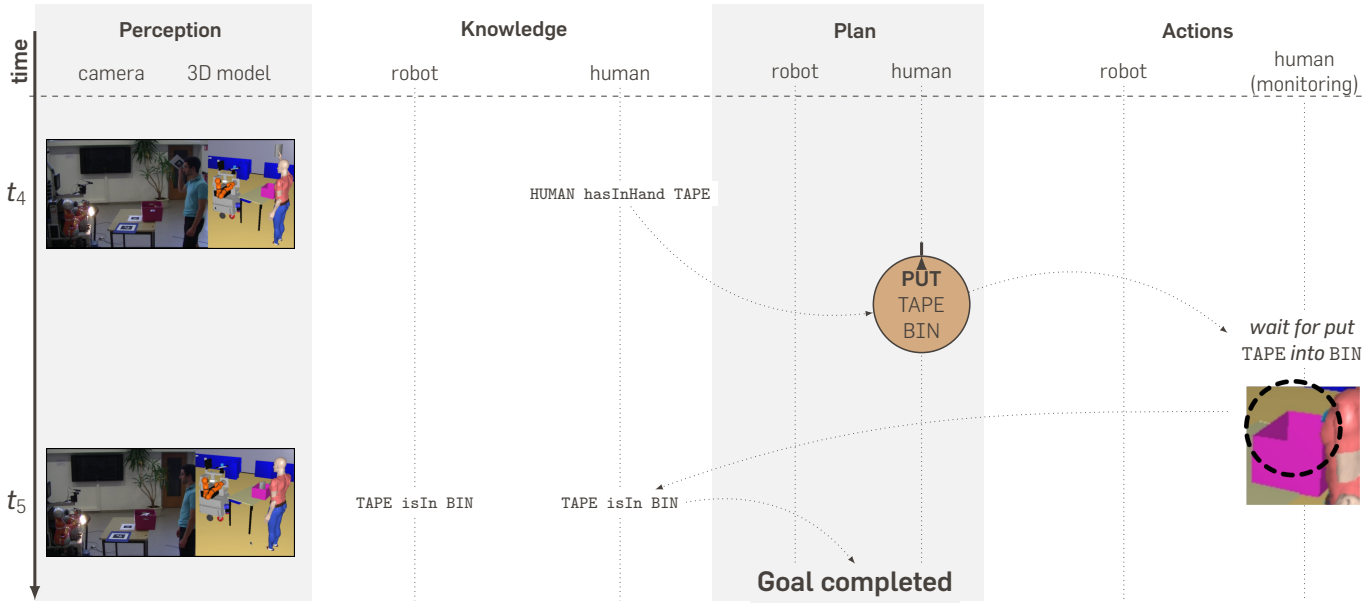
“CLEANING THE TABLE”...



“CLEANING THE TABLE”...

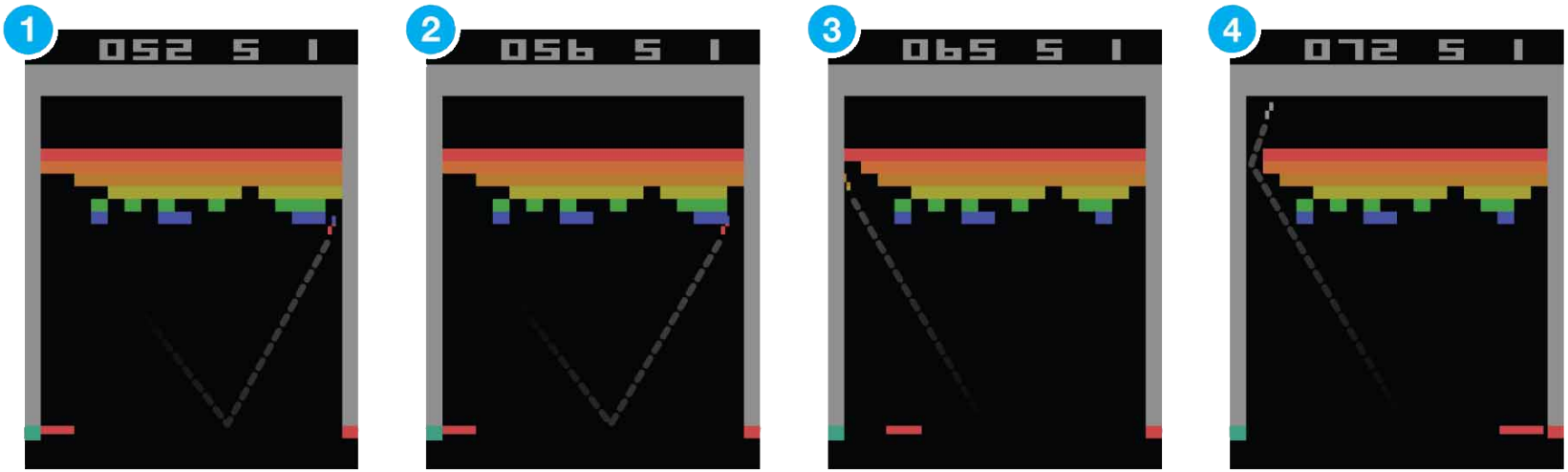


“CLEANING THE TABLE”...



DEEP LEARNING AND ROBOTICS

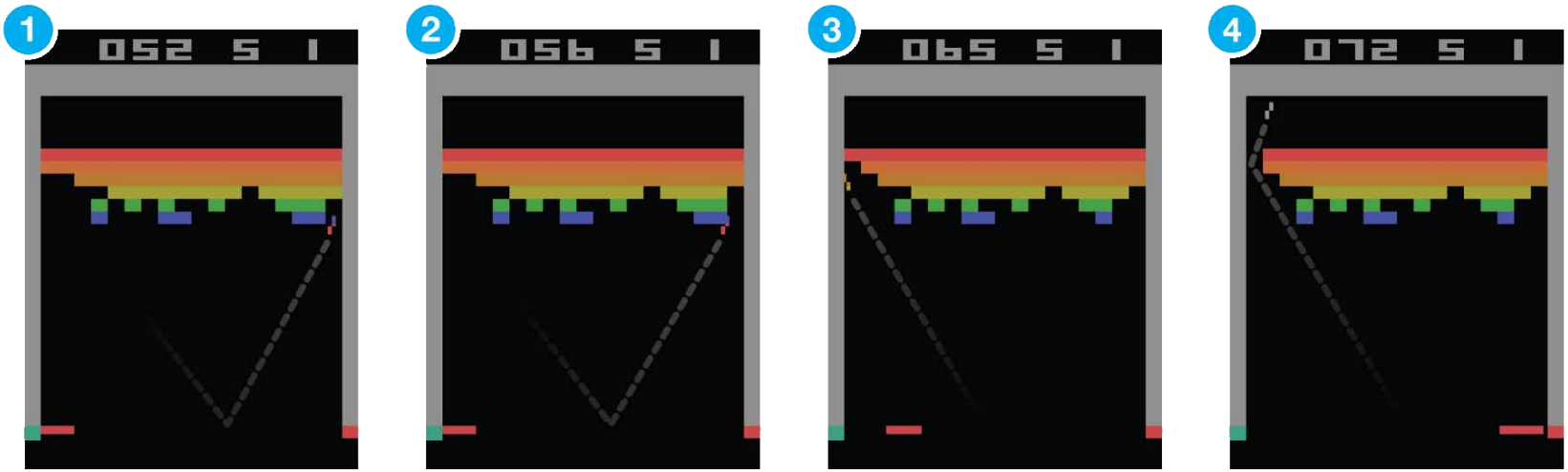
LEARNING COMPLEX BEHAVIOURS



- Inputs: raw screen image + score
- from the outside, looks like planning

[Mnih et al. **Human-level control through deep reinforcement learning** – Nature 2015]

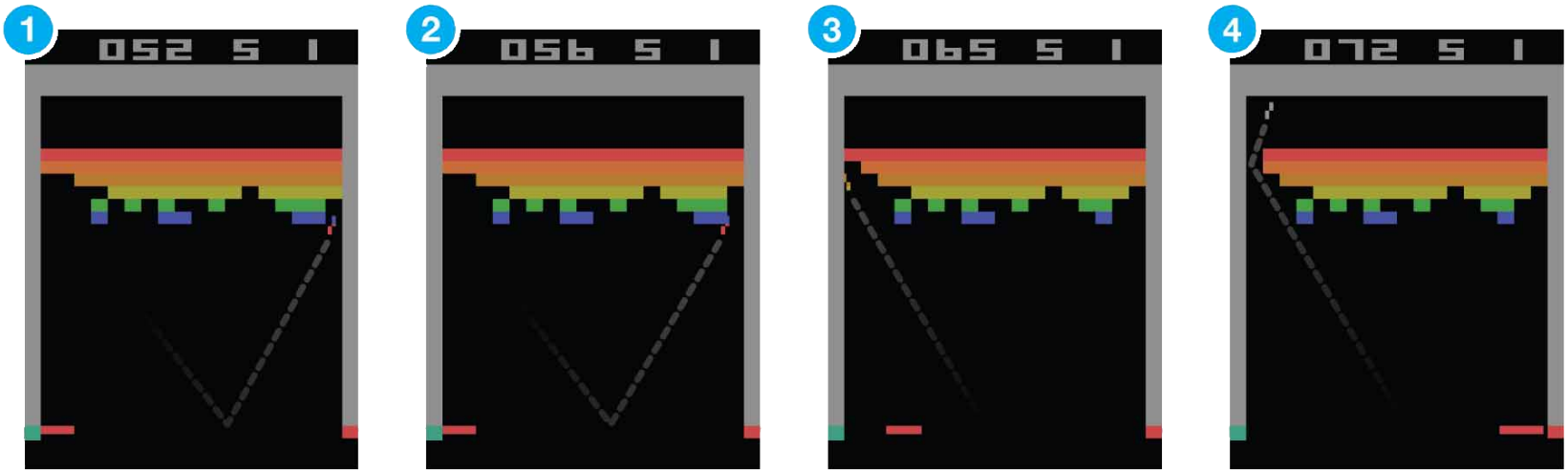
LEARNING COMPLEX BEHAVIOURS



- Inputs: raw screen image + score
- from the outside, looks like planning
- ~~1.000.000~~ **500** games to play a good human-level

[Mnih et al. **Human-level control through deep reinforcement learning** – Nature 2015]

LEARNING COMPLEX BEHAVIOURS



- Inputs: raw screen image + score
- from the outside, looks like planning
- ~~1.000.000~~ **500** games to play a good human-level

Could we also learn social dynamics?

[Mnih et al. **Human-level control through deep reinforcement learning** – Nature 2015]