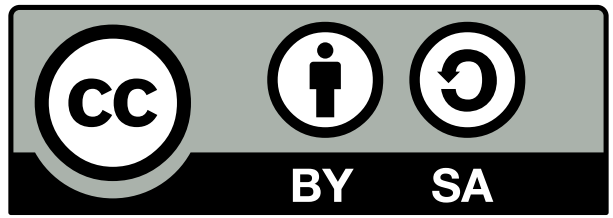


You can download the sources of this presentation here:
https://github.com/severin-lemaignan/presentation-cognitive-robotics

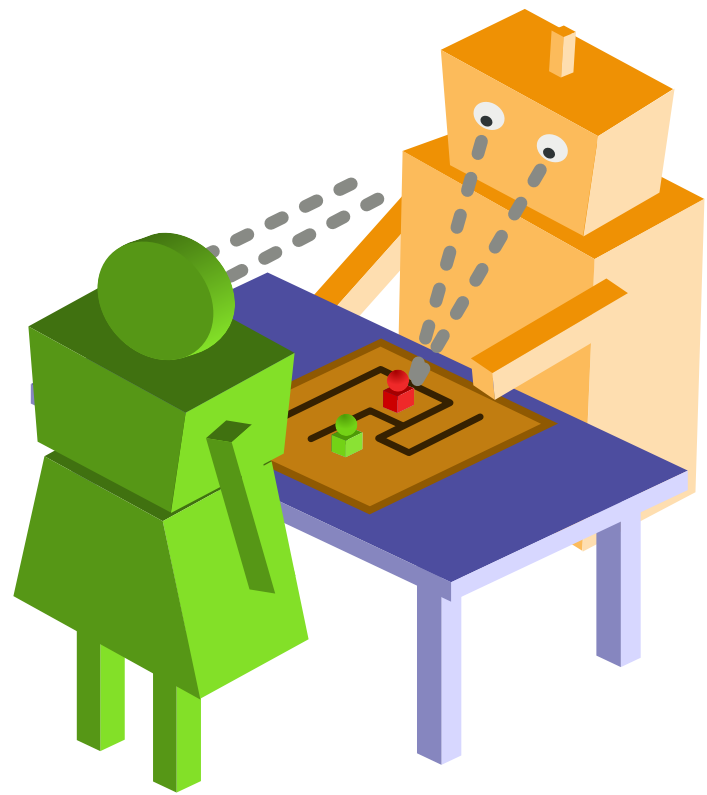
WITH PLYMOUTH UNIVERSITY

MARIE CURIE ACTIONS

Of Cognition and Social Robots

towards a theory of artificial social cognition

Heriot-Watt University – **3 February 2017**

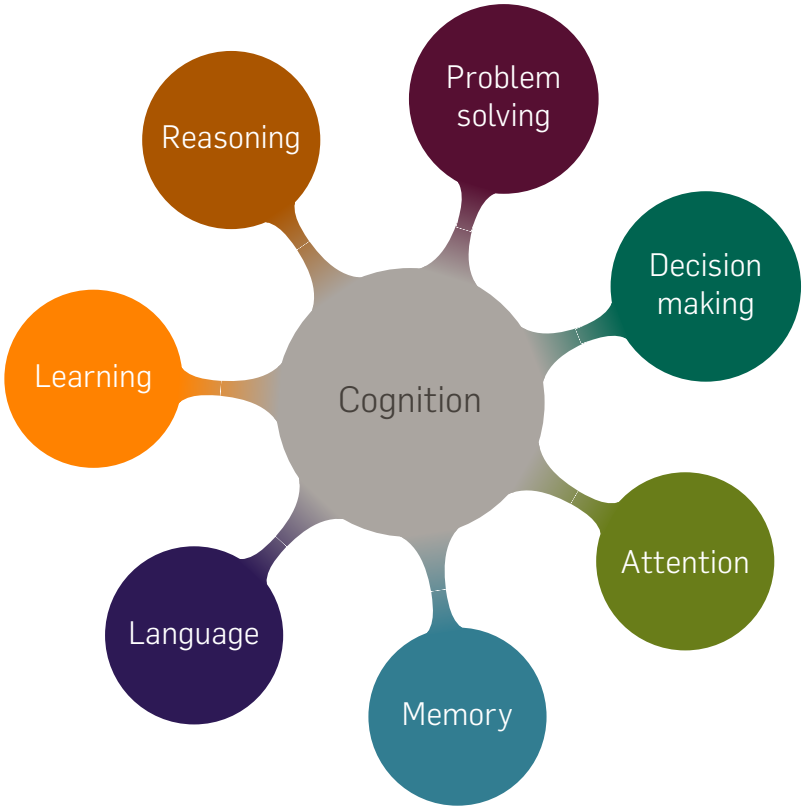


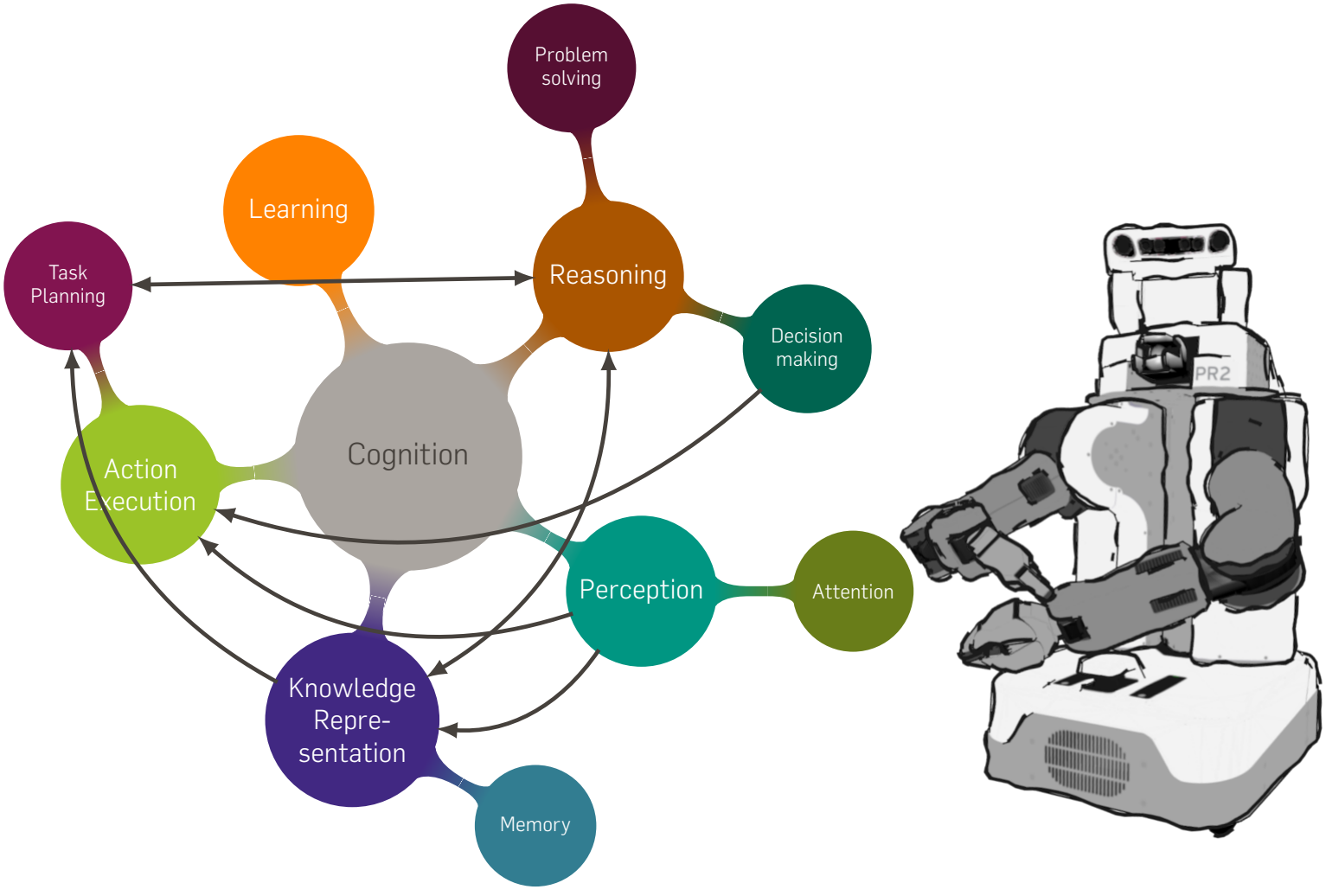
Séverin Lemaignan

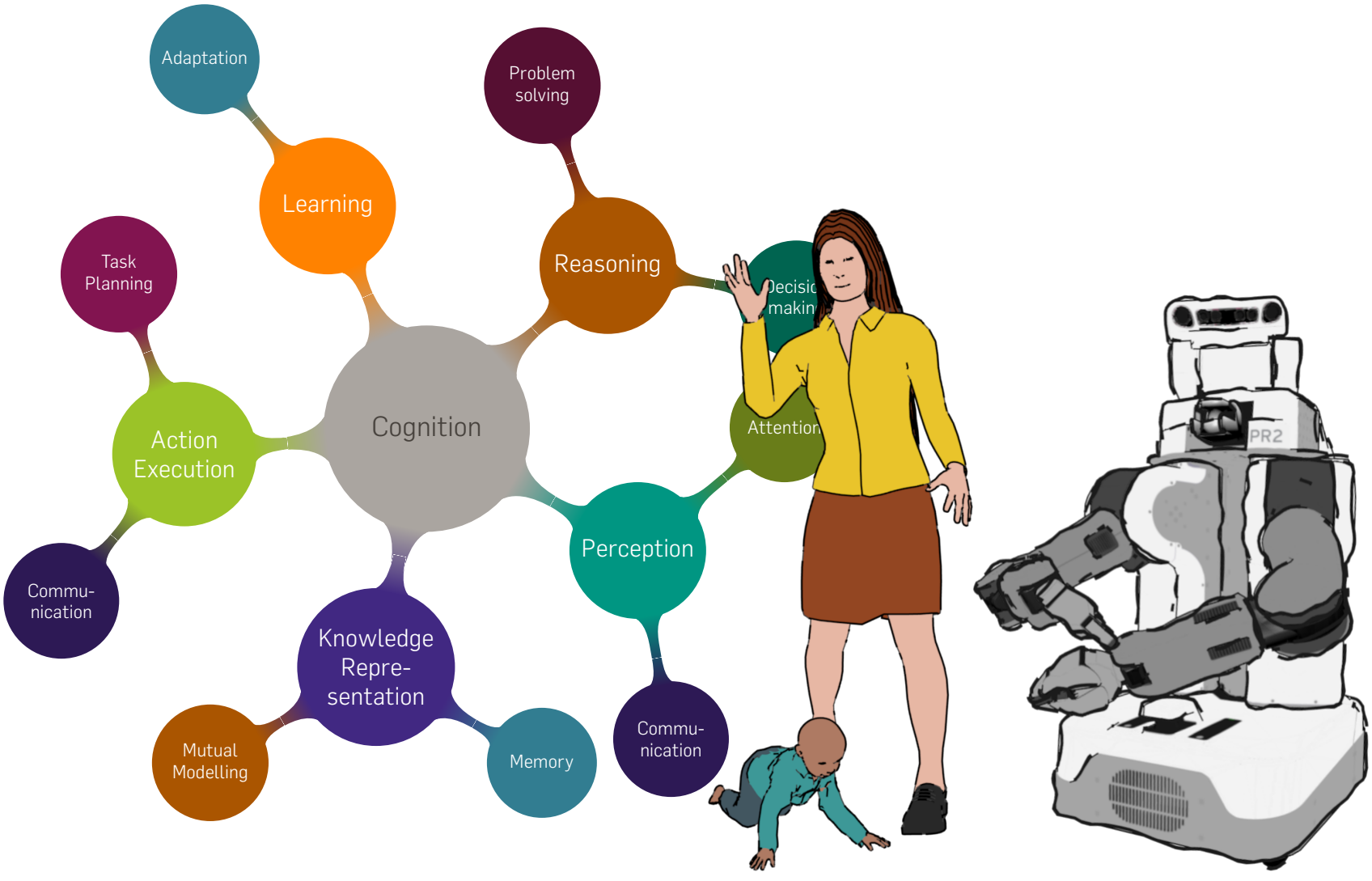
Centre for Robotics and Neural Systems
Plymouth University

MARIE CURIE ACTIONS

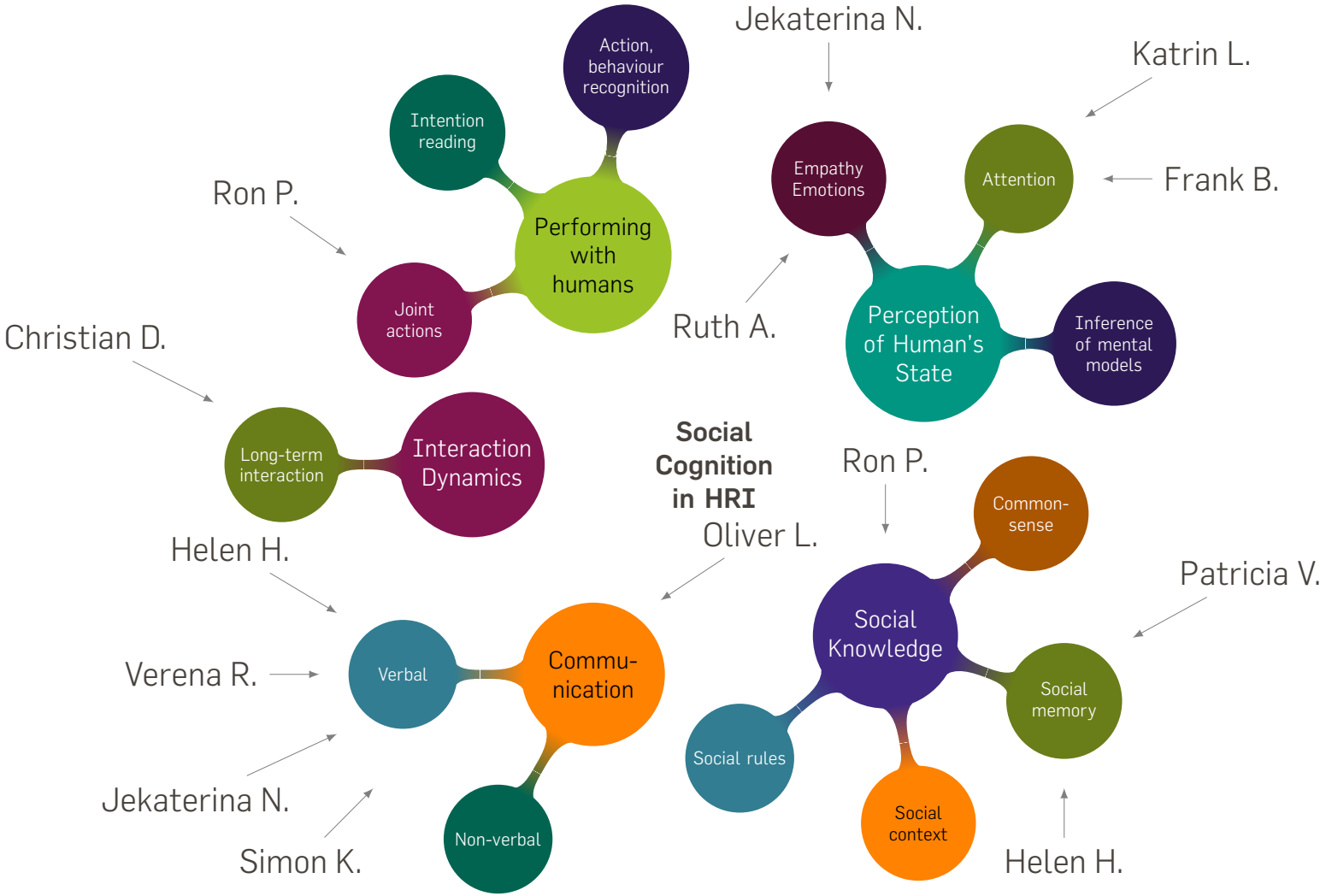
“Cognition is a group of mental processes that includes **attention**, **memory**, producing and understanding **language**, **learning**, **reasoning**, **problem solving**, and **decision making**.”


Problem solving
Reasoning
Decision making
Learning
Cognition
Attention
Language
Memory

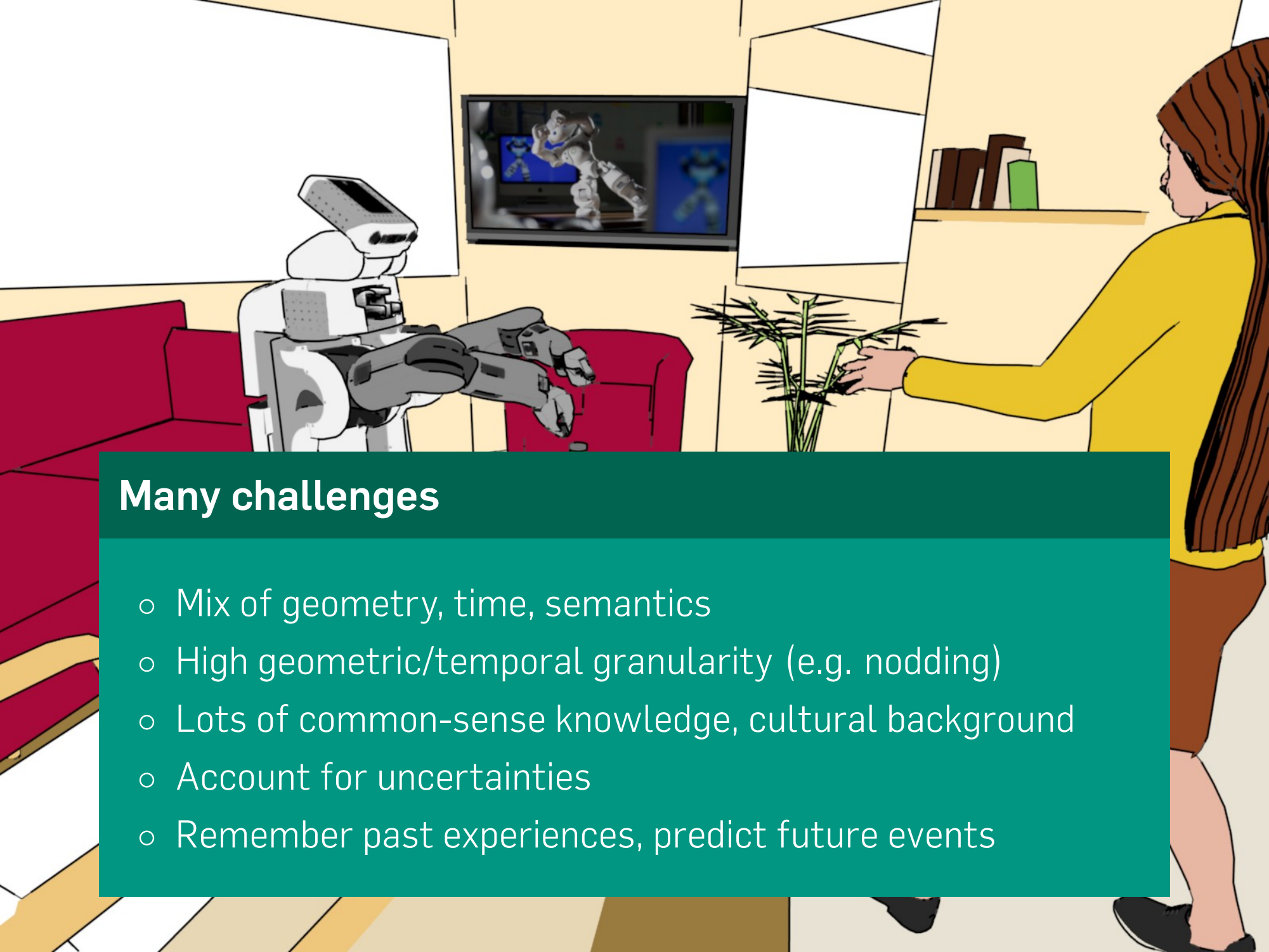

Problem solving
Learning
Task Planning
Reasoning
Decision making
Cognition
Action Execution
Perception
Attention
Knowledge Repre-sentation
Memory
PR2


Adaptation
Learning
Problem solving
Reasoning
Task Planning
Action Execution
Cognition
Attention
Perception
Commu-nication
Knowledge Repre-sentation
Mutual Modelling
Memory
Commu-nication
PR2

SURFACE FUNCTIONS FOR SOCIAL COGNITION



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



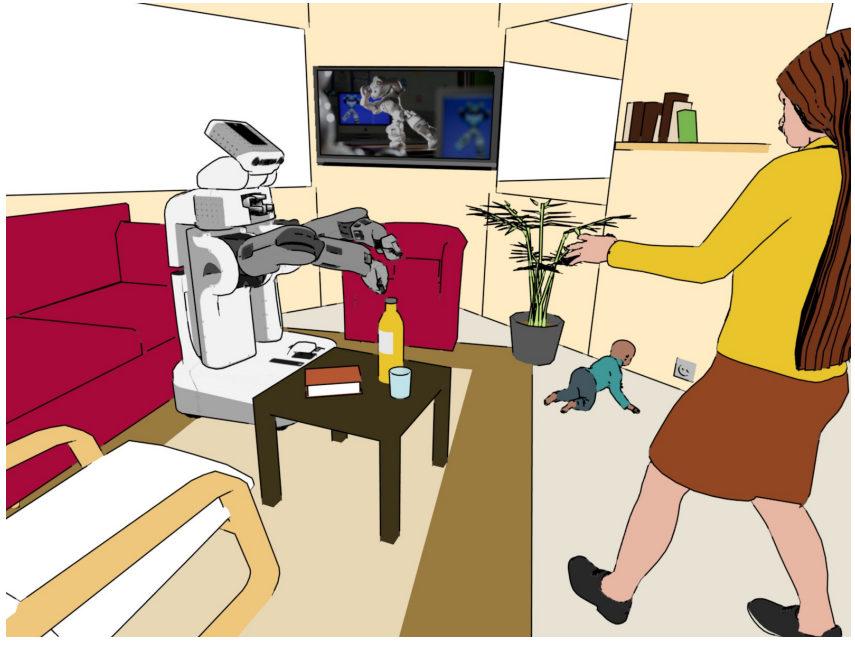
Many challenges

- Mix of geometry, time, semantics
- High geometric/temporal granularity (e.g. nodding)
- Lots of common-sense knowledge, cultural background
- Account for uncertainties
- Remember past experiences, predict future events

Situated dialogue evidences nicely these challenges

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

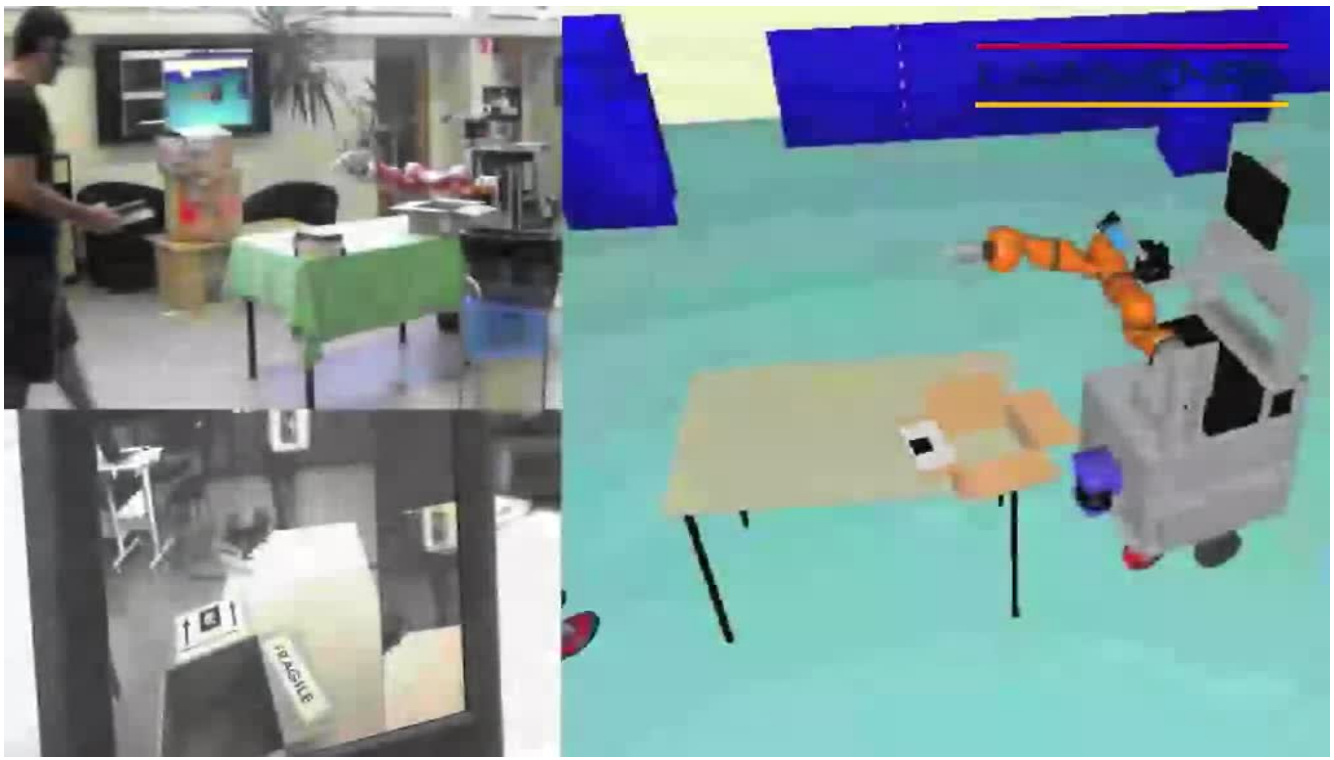
“Can you give me that book?”



My PhD: a symbolic approach to this problem

SYMBOLIC SOCIAL COGNITION

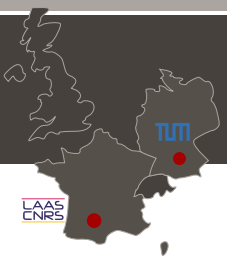
SITUATION ASSESSMENT

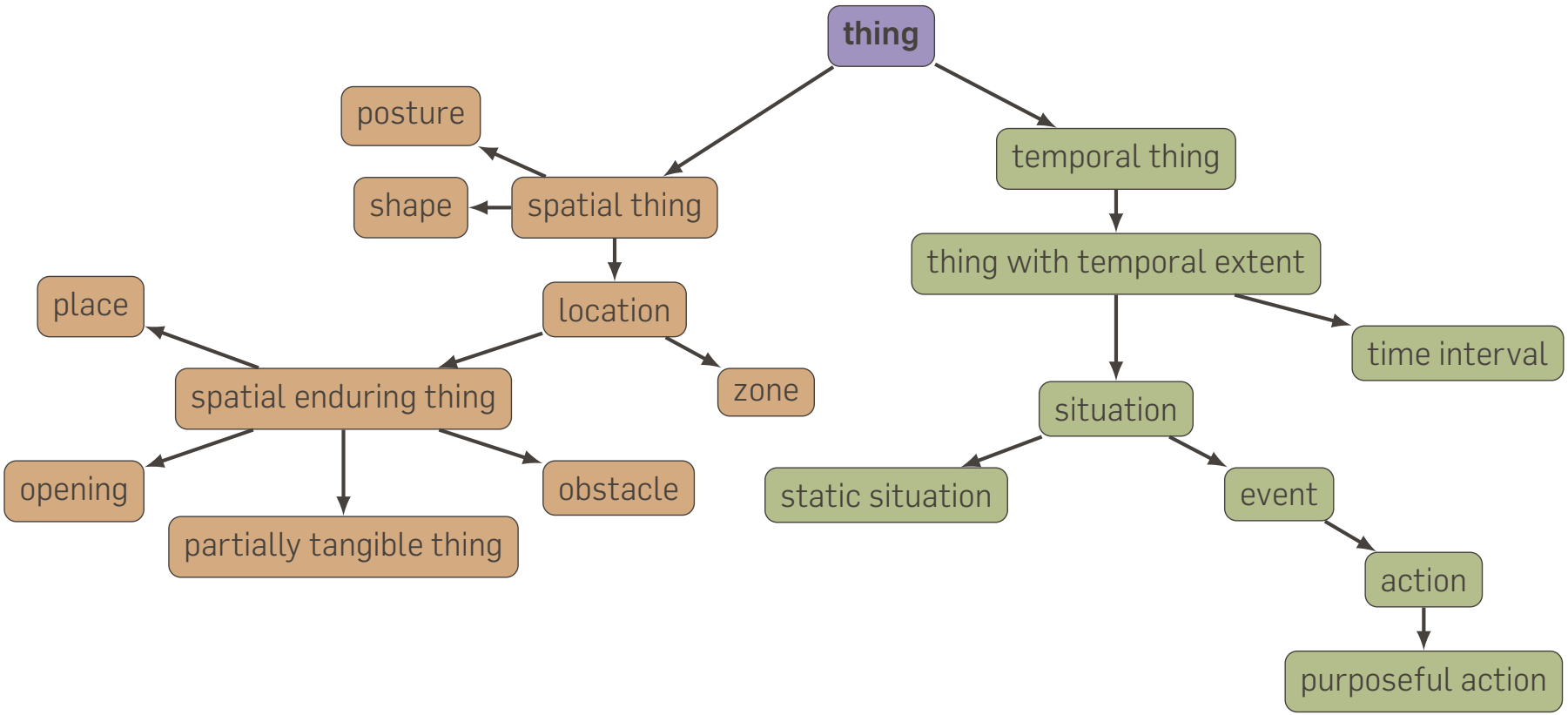


Subject	Predicate	Object
Location	isAt ≡ cyc:objectFoundInLocation	Location
	→ isOn ≡ cyc:above_Touching	
	→ isIn	
	→ isNextTo	
Location	isAbove ≡ cyc:above-Generally	Location
Location	isBelow	Location

Location	hasRelativePosition	Location
	→ behind ≡ cyc:behind-Generally	
	→ inFrontOf ≡ cyc:inFrontOf-Generally	
	→ leftOf	
	→ rightOf	
Object	cyc:farFrom	Agent
Object	cyc:near	Agent
Agent	looksAt	SpatialThing
Agent	sees	SpatialThing
SpatialThing	isInFieldOfView	xsd:boolean
Agent	pointsAt ≡ cyc:pointingToward	SpatialThing
Agent	focusesOn	SpatialThing
Agent	seesWithHeadMovement	SpatialThing
Agent	canReach	Object

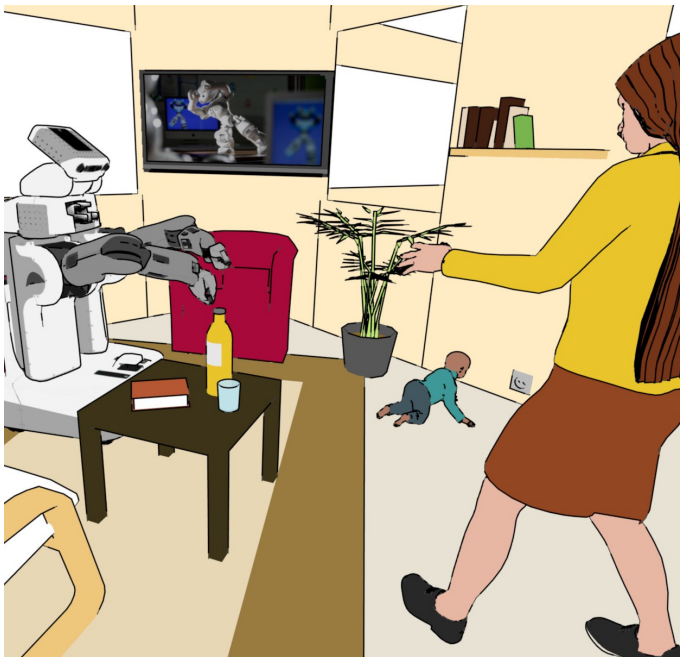
FROM SPATIAL MODEL TO SYMBOLIC MODEL

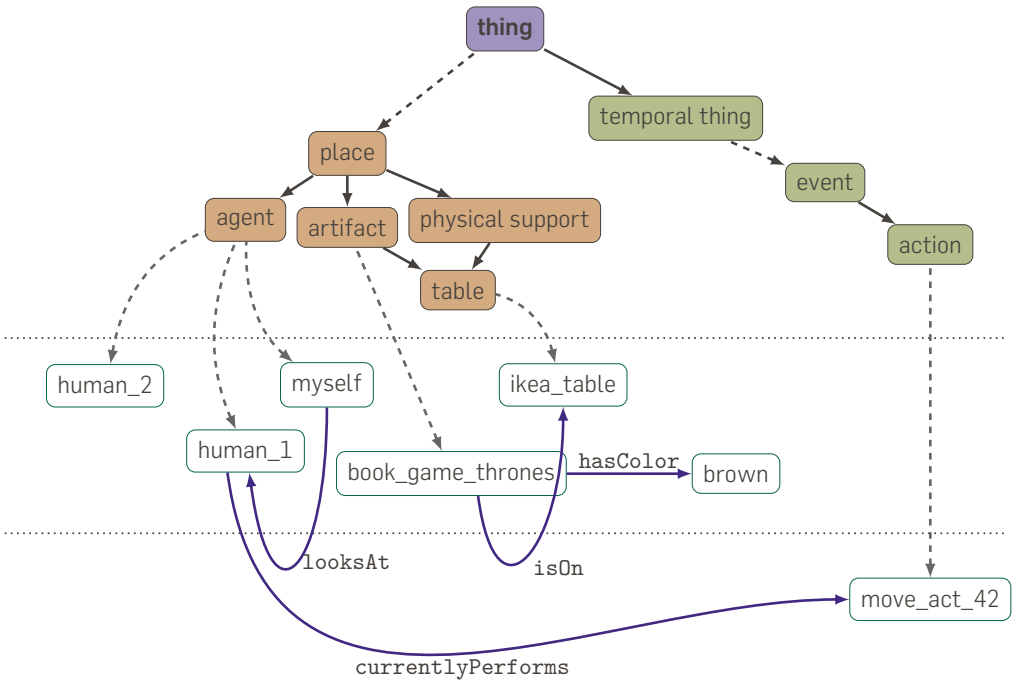




[Lemaignan et al. **ORO, a Knowledge Management Module for Cognitive Architectures in Robotics**, IROS 2010]

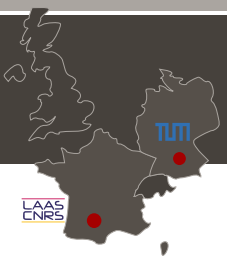
ONLINE INSTANTIATION





[Lemaignan, Alami, **Explicit Knowledge and the Deliberative Layer: Lessons Learned**, IROS 2013]

DIALOGUE GROUNDING



I keep the NLP details for the questions, but:

"Give me the book on the table"

↓

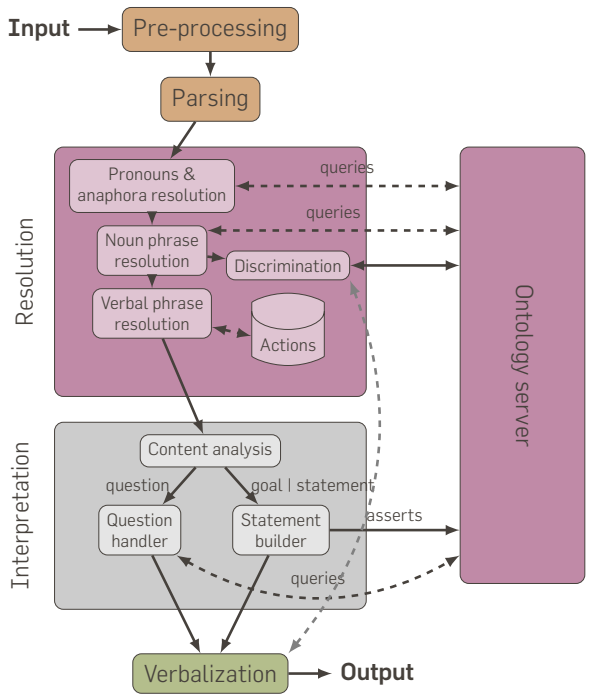
`me` → `human_1`

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

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

↓

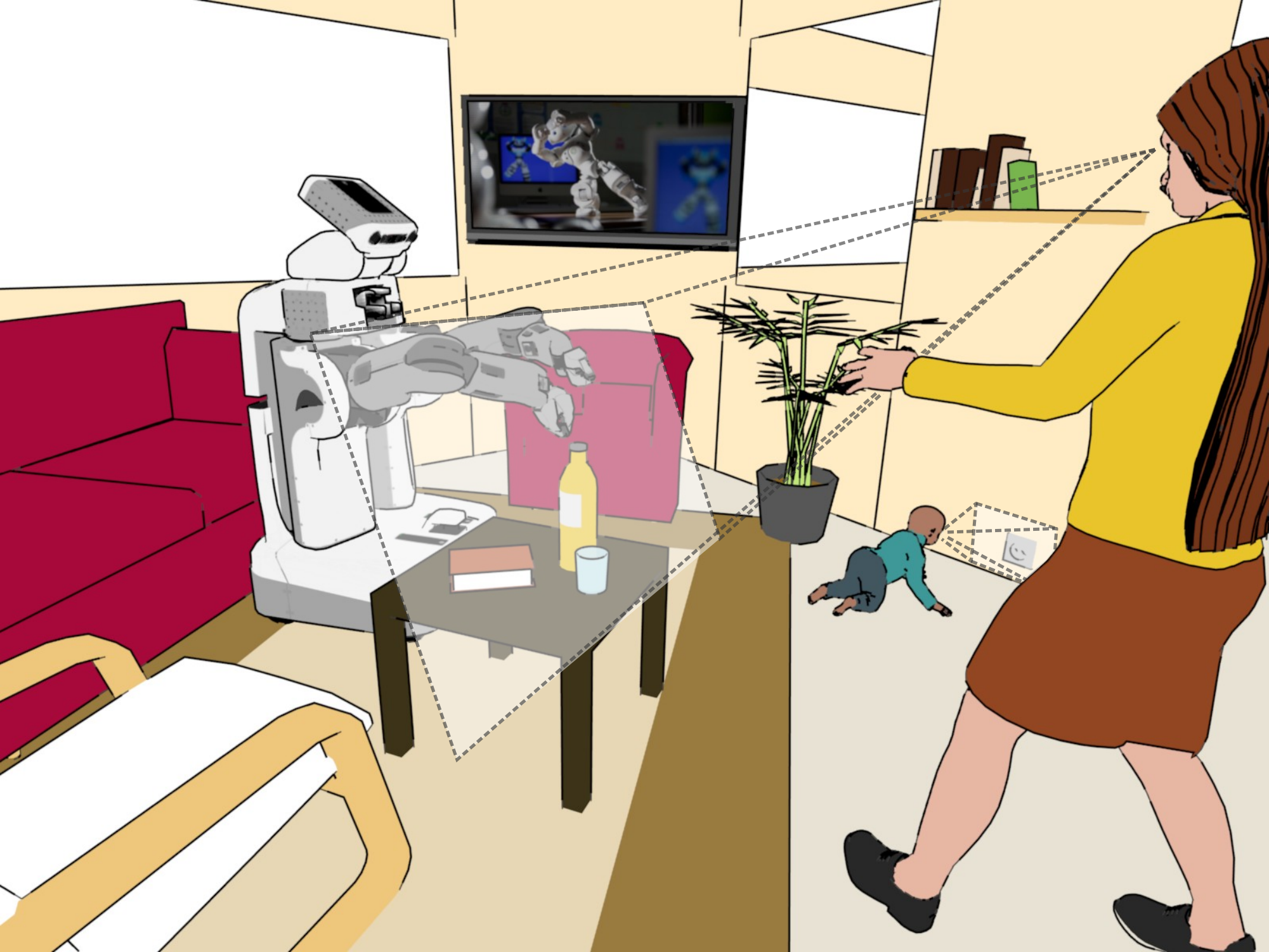
```
human_1 desires give_act_1,
give_act_1 type Give,
give_act_1 performedBy myself,
give_act_1 actsOnObject book,
give_act_1 receivedBy human_1
```



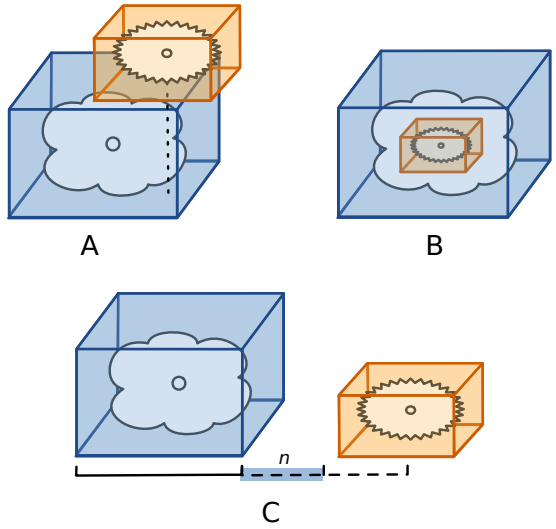
[Lemaignan et al. **Grounding the Interaction: Anchoring Situated Discourse in Everyday HRI** Intl. J. of Social Robotics 2011]

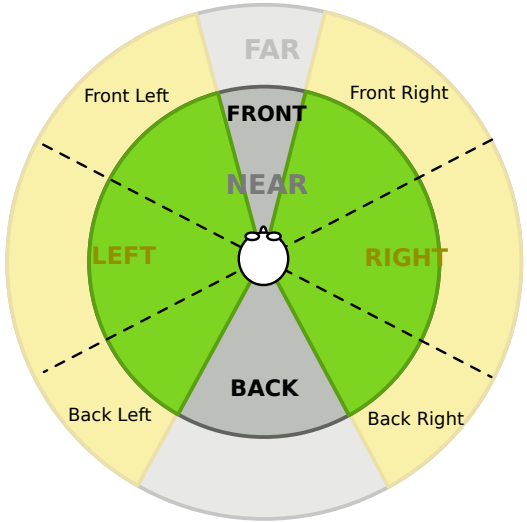
What about

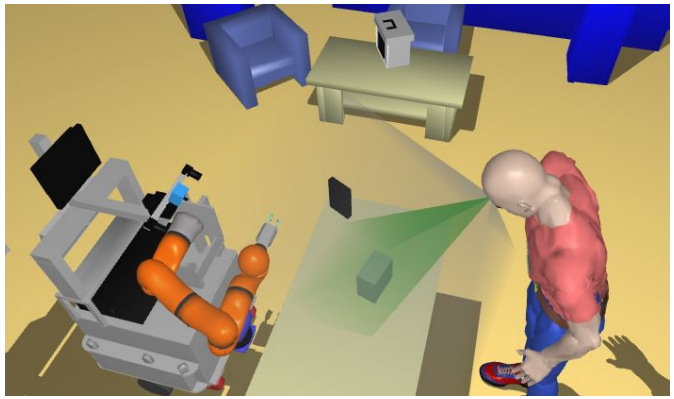
“Give me that book”?

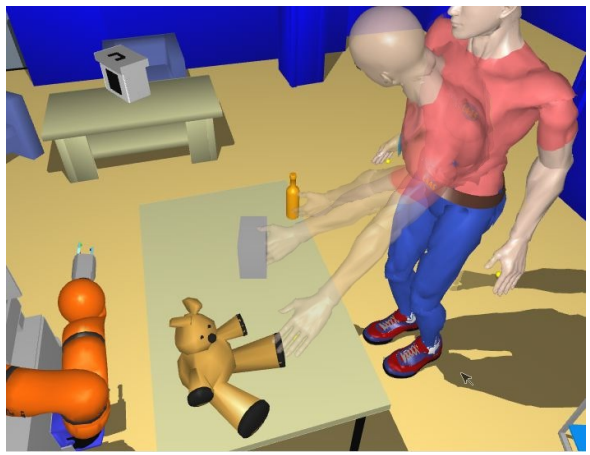


VISUAL PERSPECTIVE TAKING







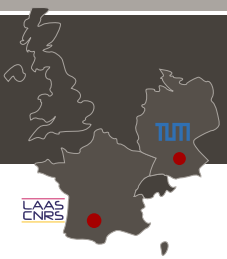


Subject	Predicate	Object
Location	isAt ≡ cyc:objectFoundInLocation	Location
	→ isOn ≡ cyc:above_Touching	
	→ isIn	
	→ isNextTo	
Location	isAbove ≡ cyc:above-Generally	Location
Location	isBelow	Location

Location	hasRelativePosition	Location
	→ behind ≡ cyc:behind-Generally	
	→ inFrontOf ≡ cyc:inFrontOf-Generally	
	→ leftOf	
	→ rightOf	
Object	cyc:farFrom	Agent
Object	cyc:near	Agent
Agent	looksAt	SpatialThing
Agent	sees	SpatialThing
SpatialThing	isInFieldOfView	xsd:boolean
Agent	pointsAt ≡ cyc:pointingToward	SpatialThing
Agent	focusesOn	SpatialThing
Agent	seesWithHeadMovement	SpatialThing
Agent	canReach	Object


LAAS-CNRS

DIALOGUE GROUNDING



"Where is the other tape?"

↓

find(`?obj isAt ?loc, ?obj type VideoTape, ?obj differentFrom TAPE1`)

Symbolic approaches effective at dealing with this kind of constraints

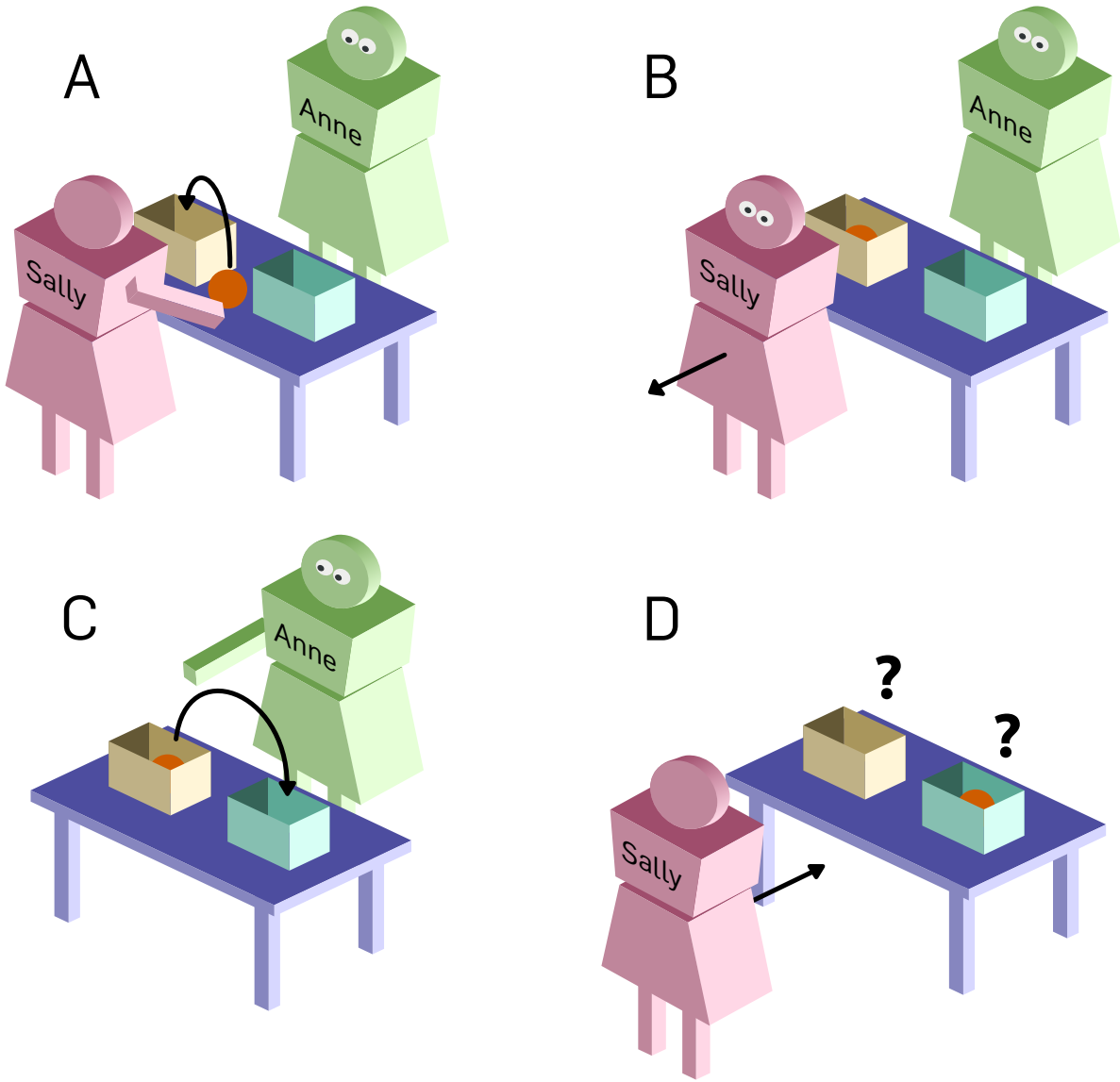
[Lemaignan et al. **Grounding the Interaction: Anchoring Situated Discourse in Everyday HRI** Intl. J. of Social Robotics 2011]

ONE STEP FURTHER: THEORY OF MIND

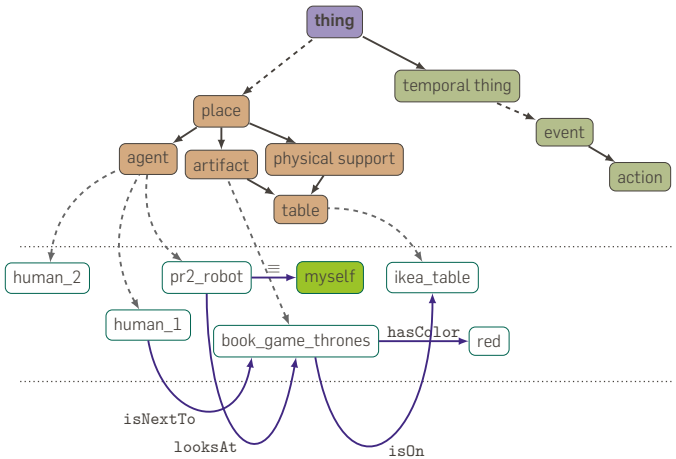
What if I ask for the video tape in the box, but the robot previously moved it somewhere else?

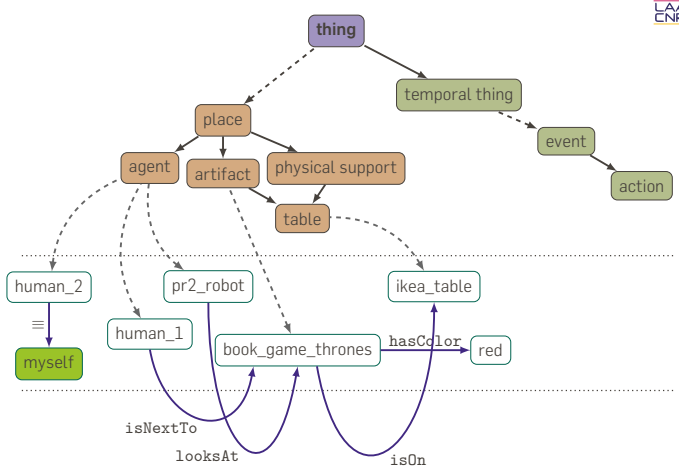
False-belief situation

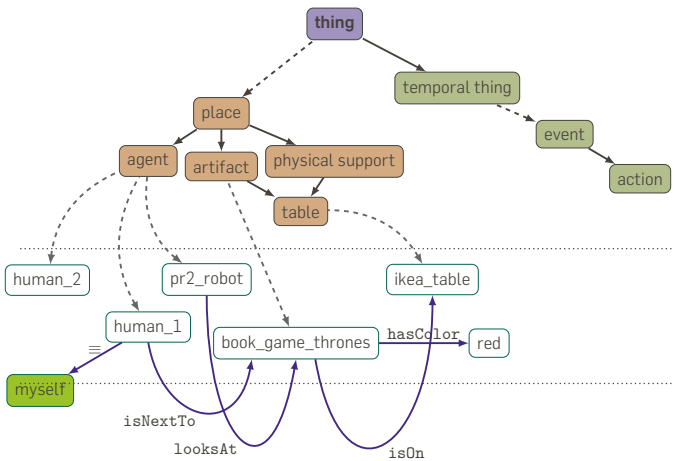
THE FALSE-BELIEF EXPERIMENT



MULTIPLE SYMBOLIC MODELS



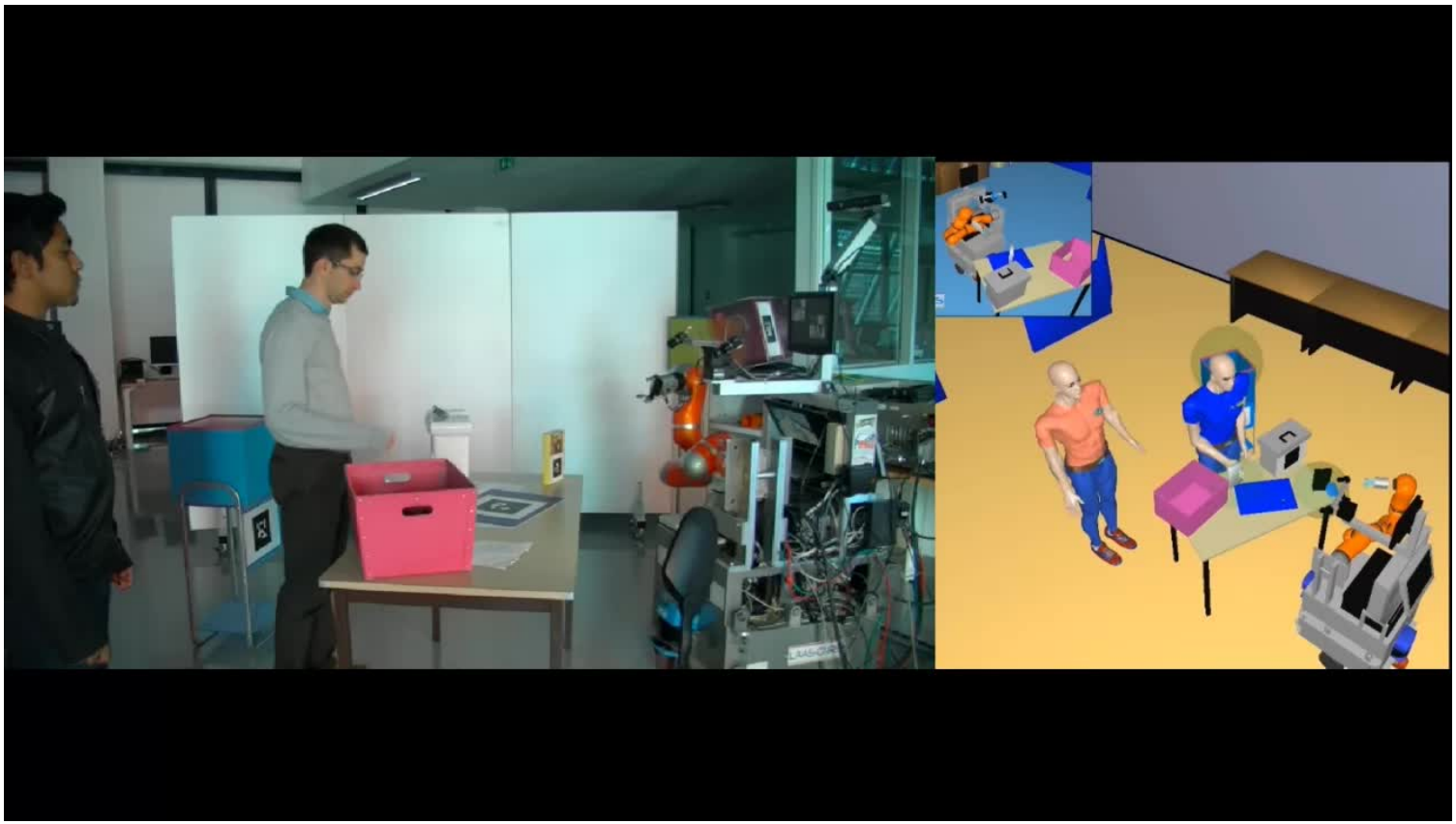




. . .

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

THE FALSE-BELIEF EXPERIMENT



[Warnier, Guitton, Lemaignan, Alami **When the Robot Puts Itself in Your Shoes** – ROMAN 2012]


LAAS-CNRS

"Give me the can!"

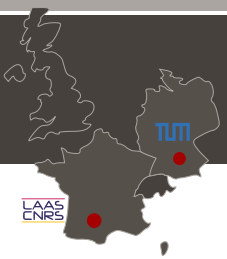
↓

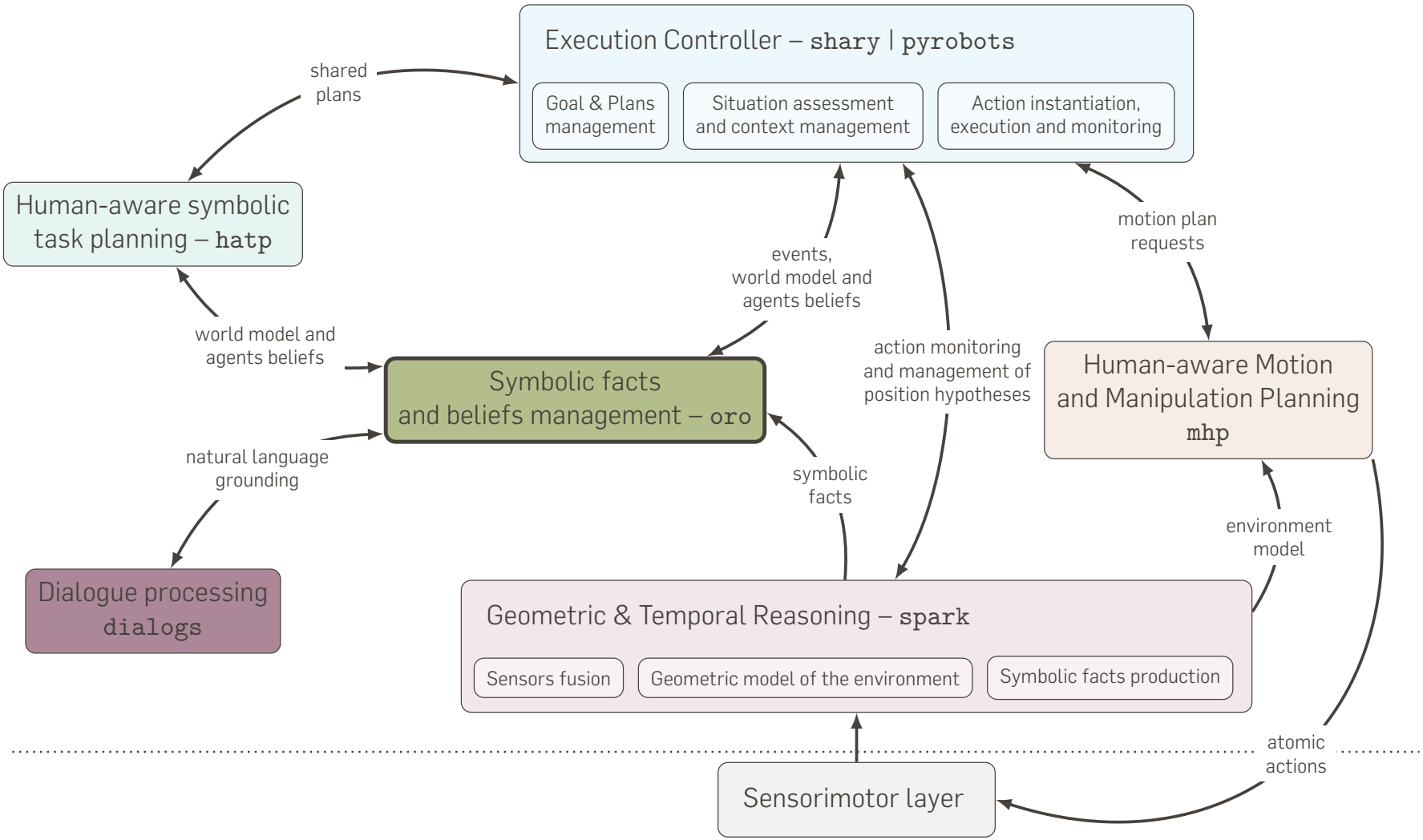
find(`?obj type Can, model='severin'`)

How to integrate all these?


Emotions
Theory of Mind
Natural Language
Gestures
Joint Actions
Attention

INTO AN CONTROL ARCHITECTURE

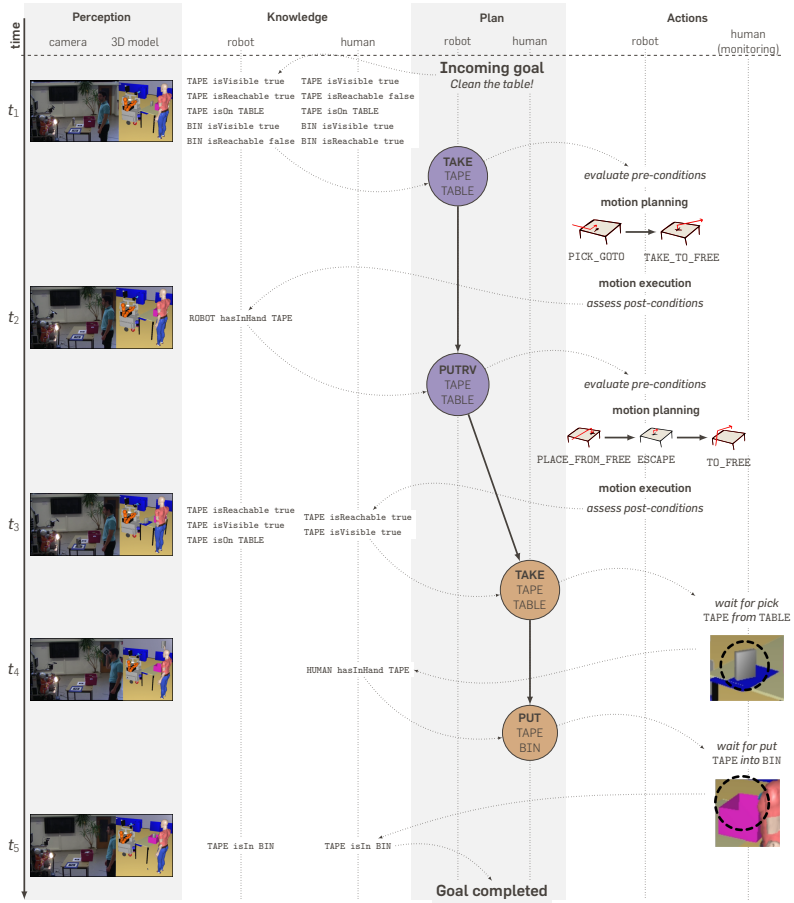




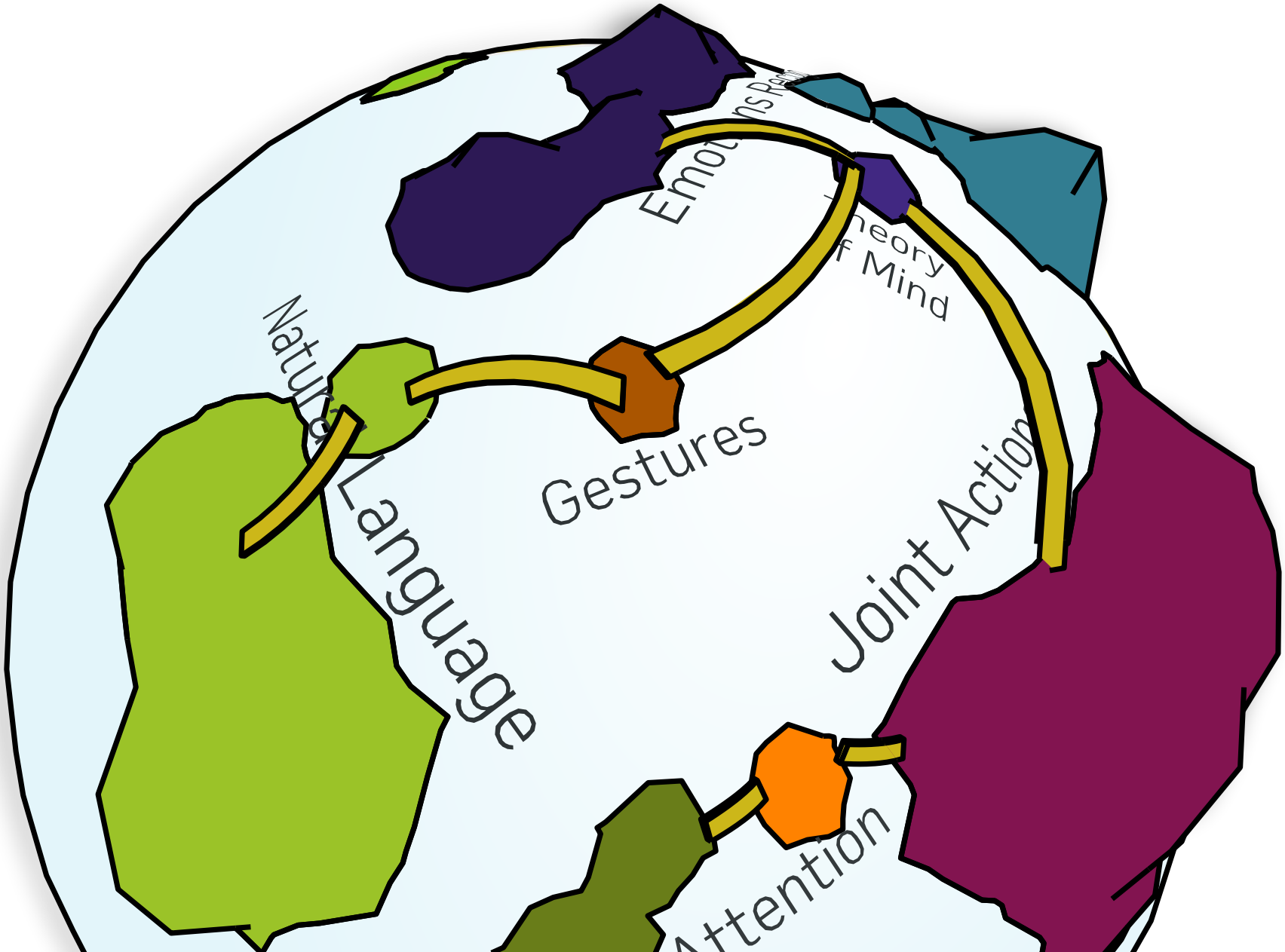
[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

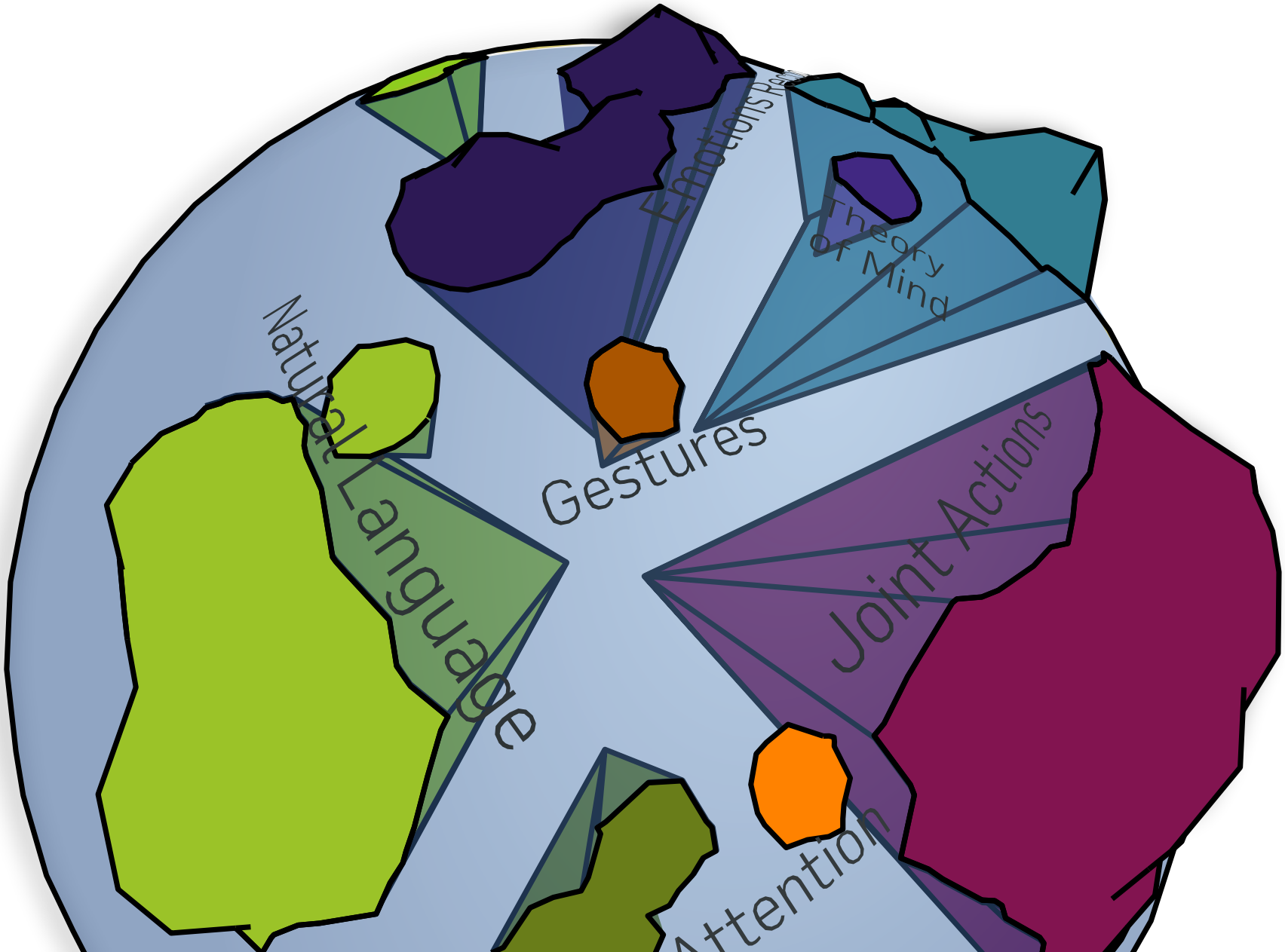
ONE EXAMPLE

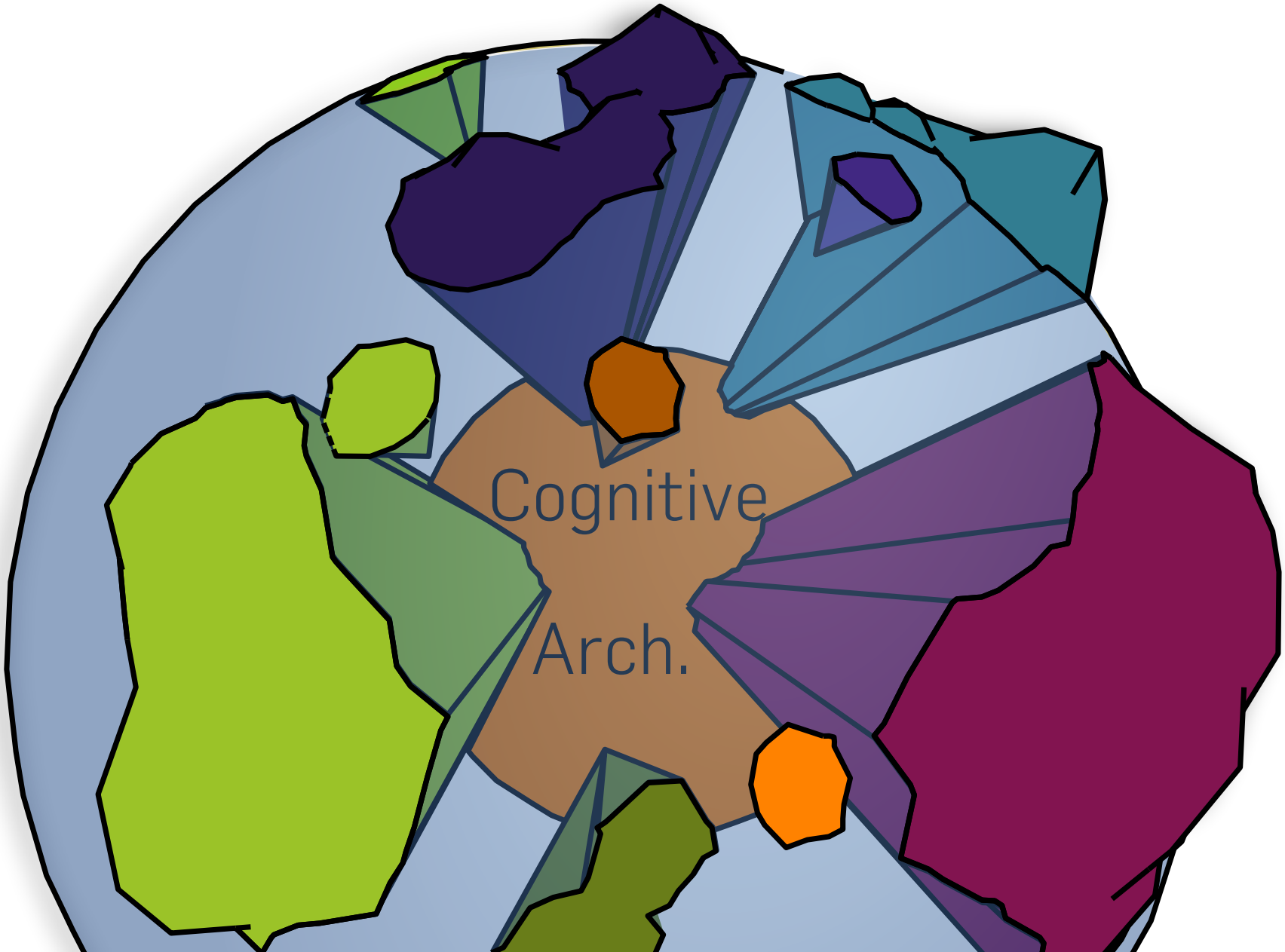
Walkthrough one full-stack example in the question



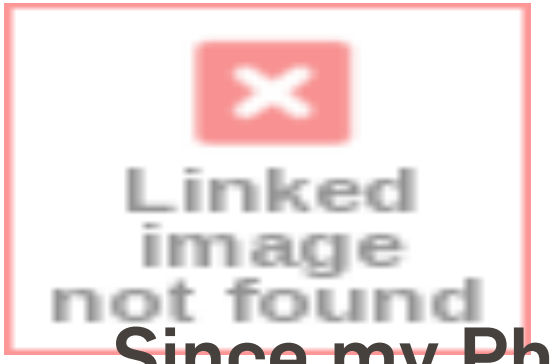
[Lemaignan et al. **Artificial Cognition for Social Human-Robot Interaction: An Implementation** Artifical Intelligence, 2016]

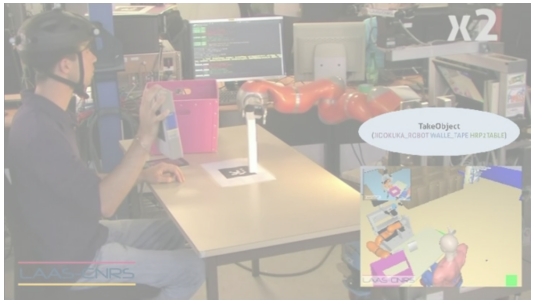

Emotions
Theory
of Mind
Natural Language
Gestures
Joint Action
Attention

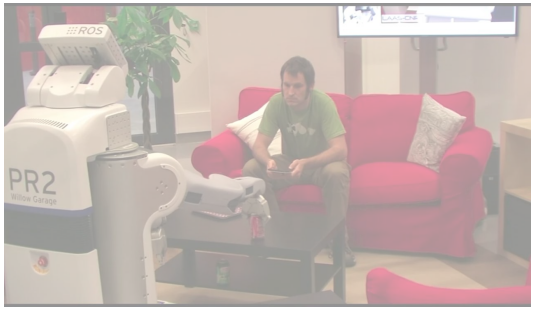

Emotions
Theory of Mind
Gestures
Natural Language
Joint Actions
Attention


Cognitive
Arch.

EXPERIMENTAL WORK

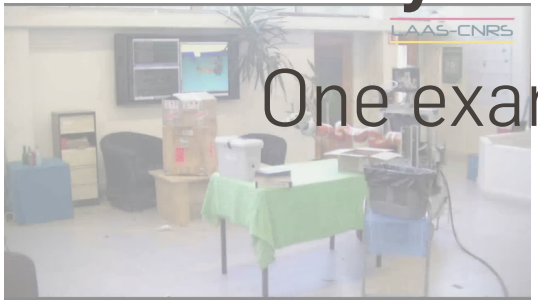


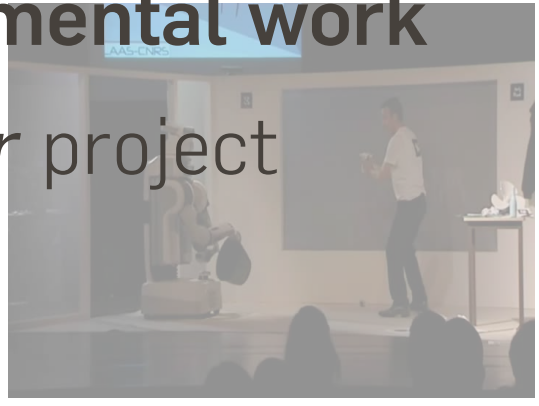


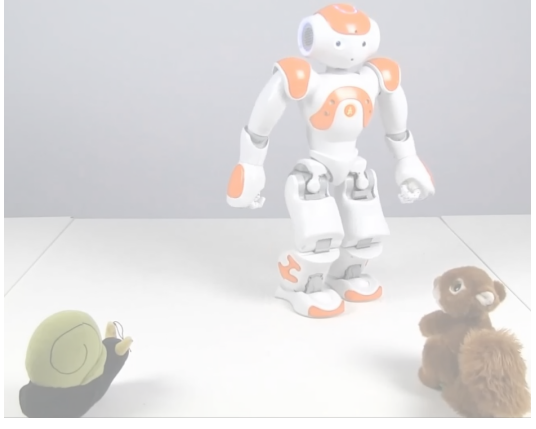


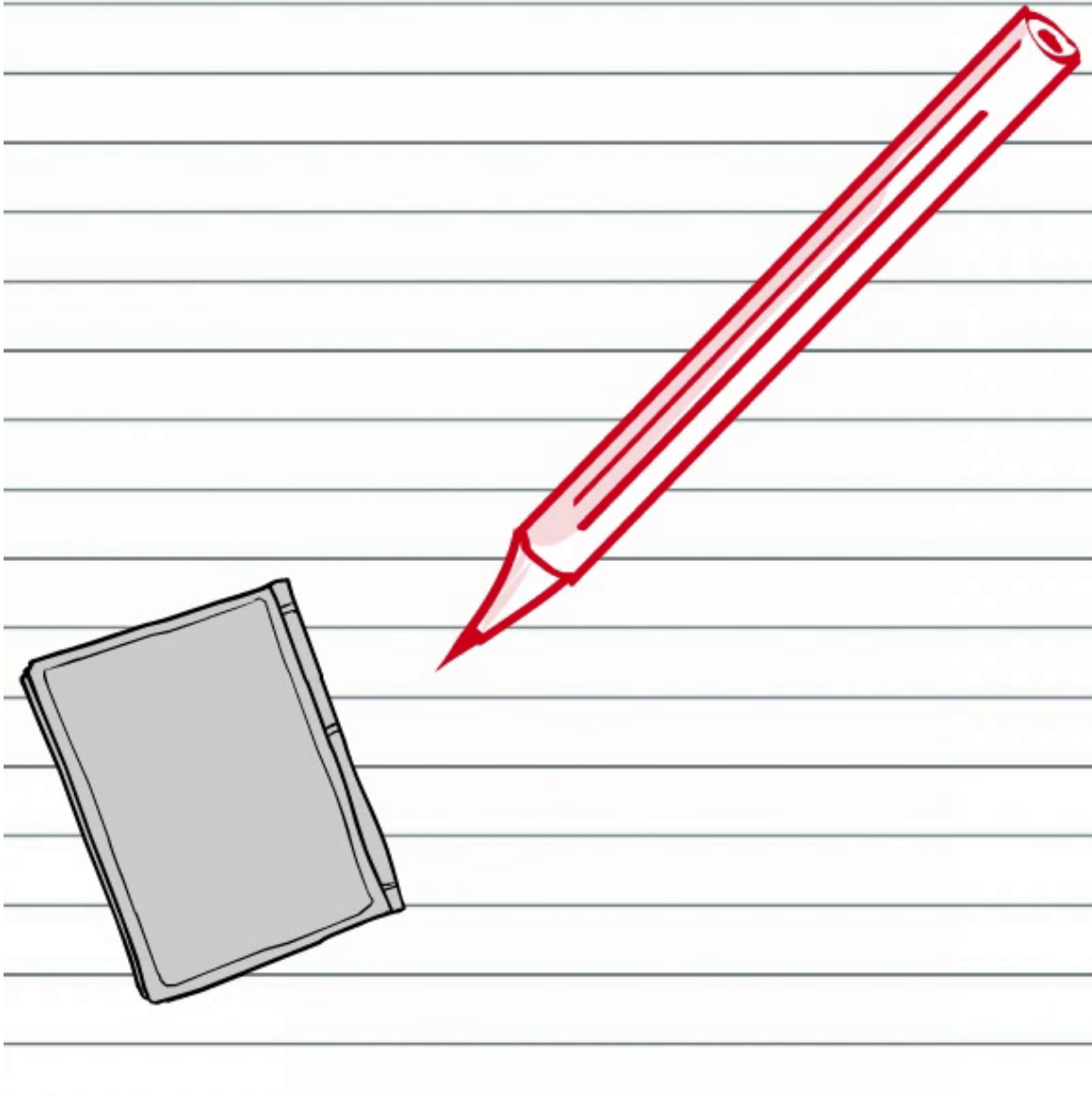
Since my PhD, a lot of experimental work

One example: the *CoWriter* project

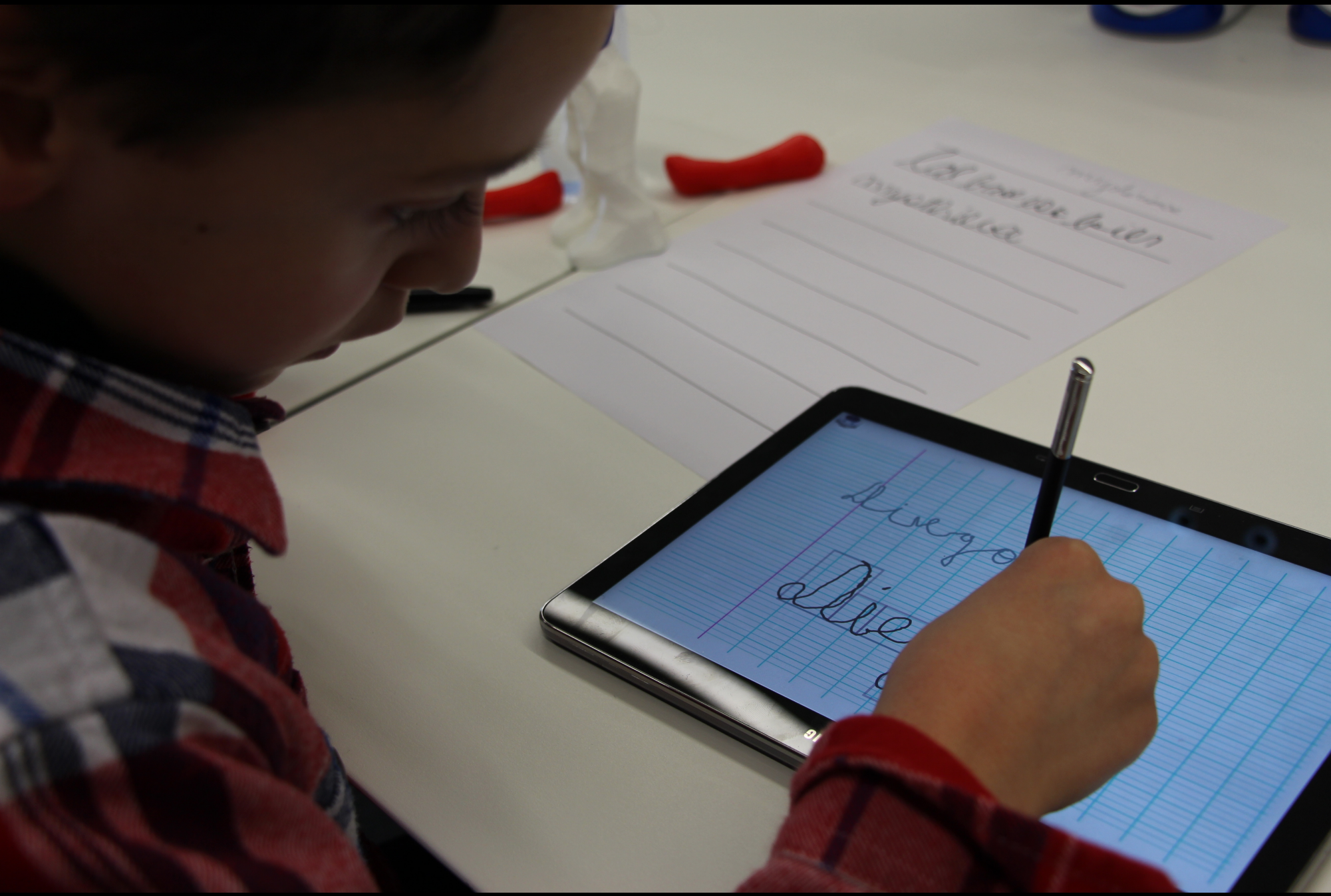






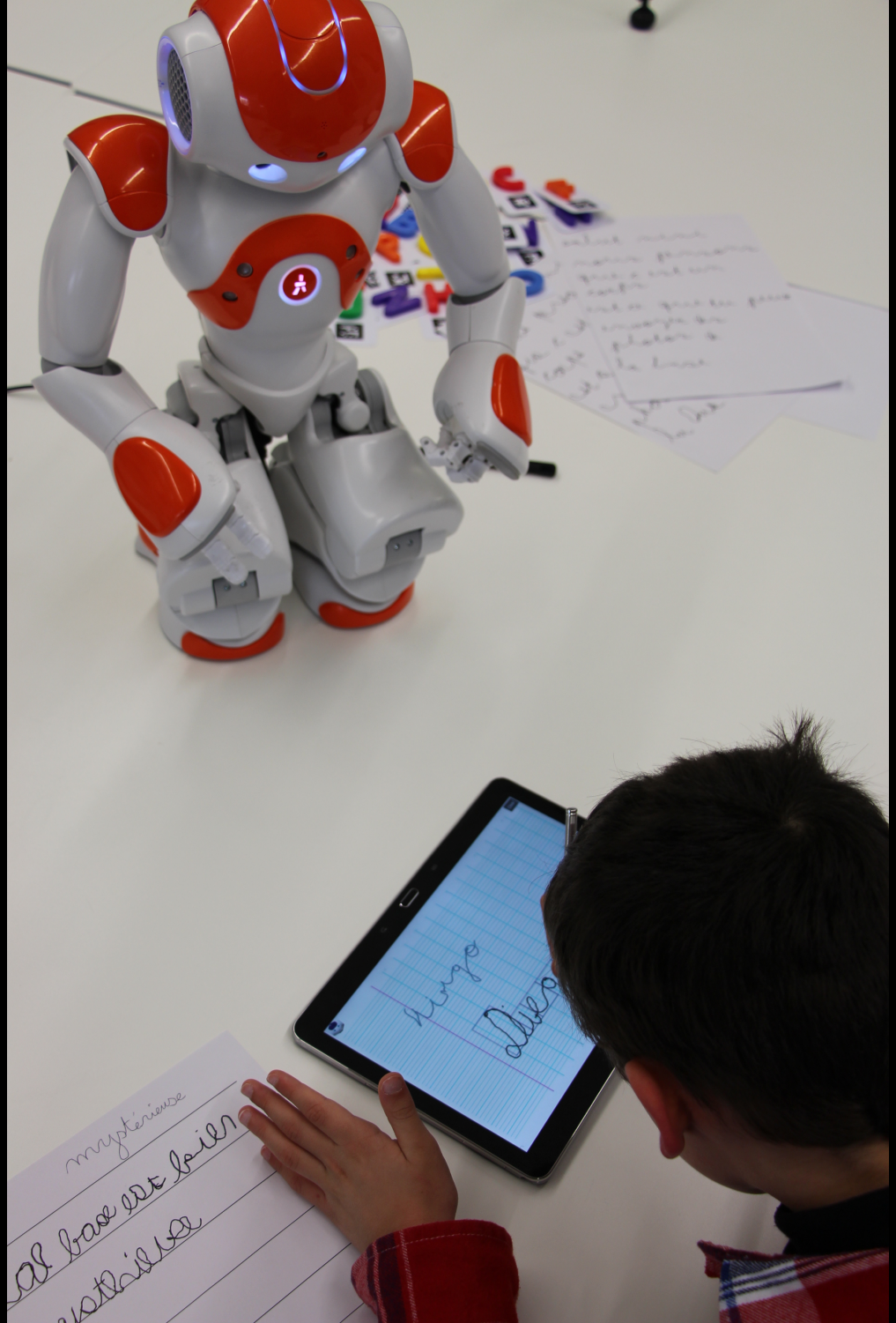

Joe


I'm not good at this
what's the use?
... but, everyone can see



ROBOTS?

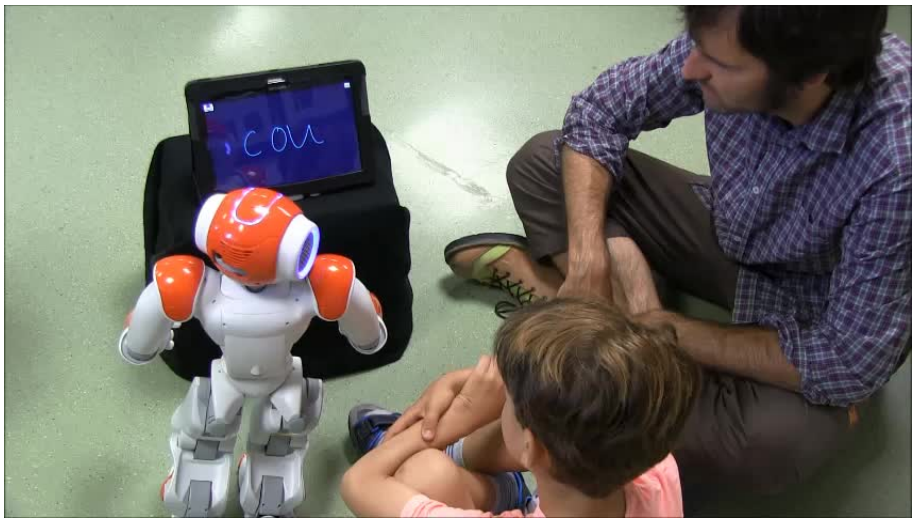
- Robots do not know how to write!
- Learning by Teaching
- (nice side-effect: we can adapt to each child and each disabilities)

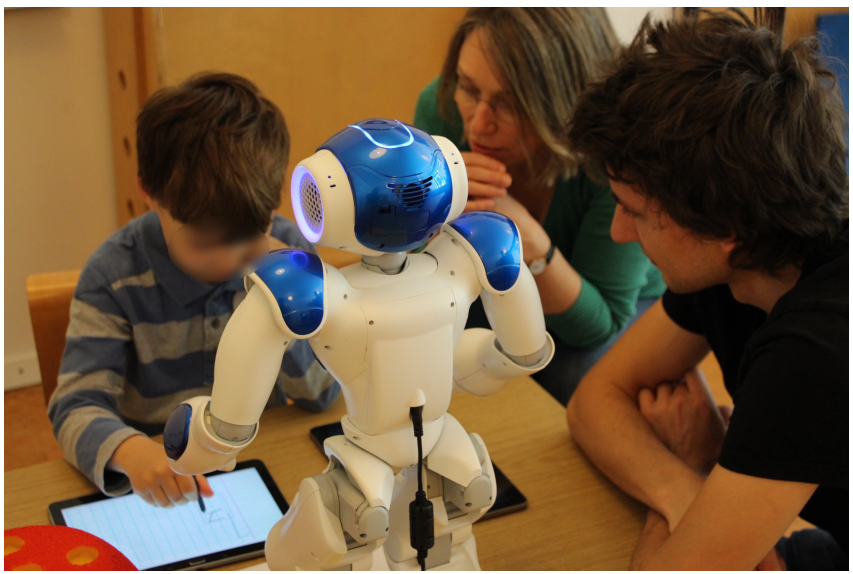

mystérieuse

Mind modelling is **mutual**

We can take advantage of it in HRI at fundamental levels

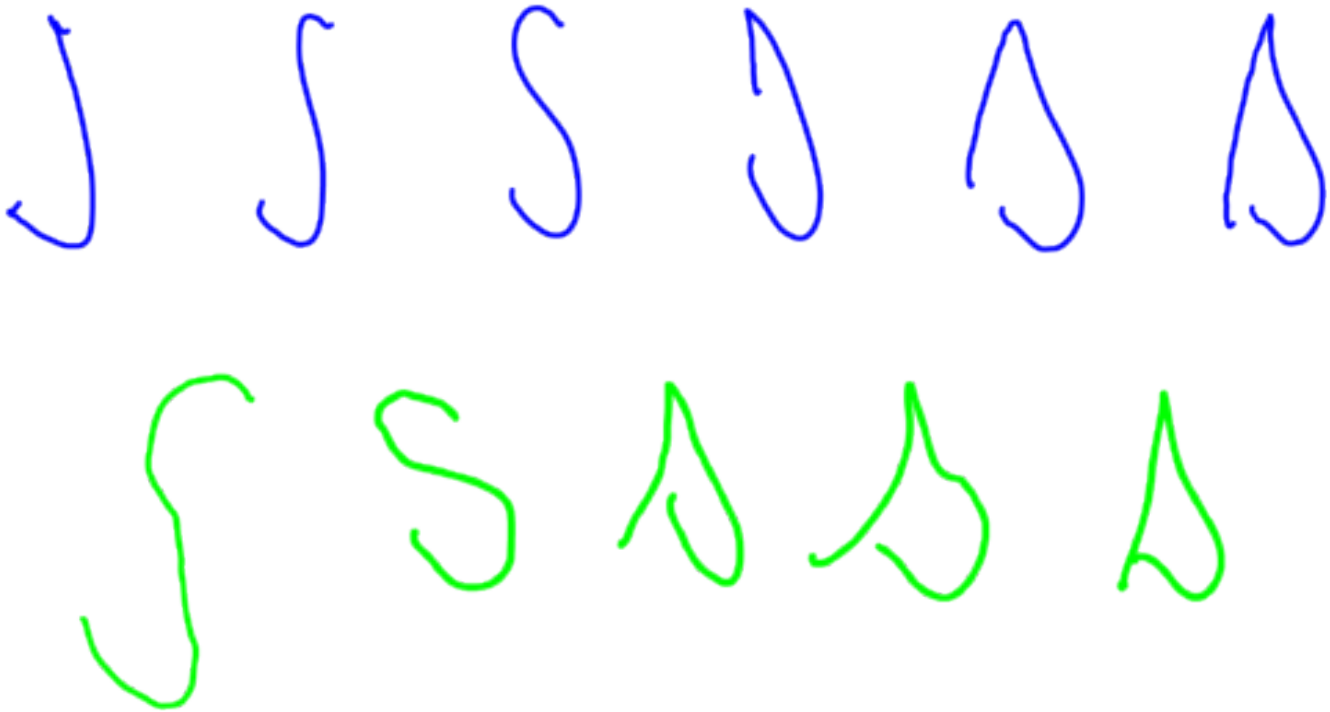
COGNITIVE ENGAGEMENT AND META-COGNITION





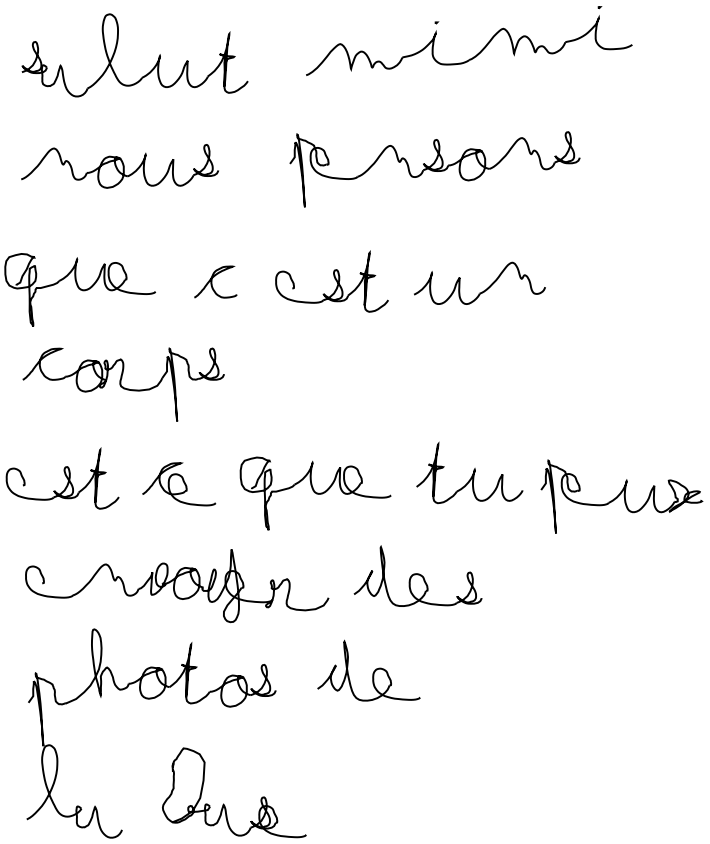
[Lemaignan et al. **Learning by Teaching a Robot: The Case of Handwriting** – Robotics and Automation Magazine 2016]

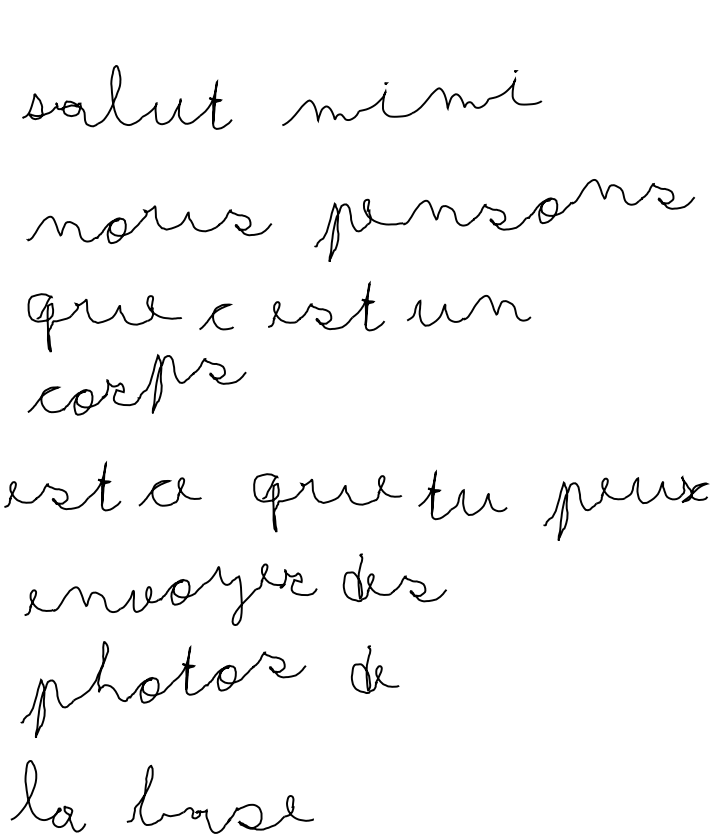
LEARNING FROM DEMONSTRATION



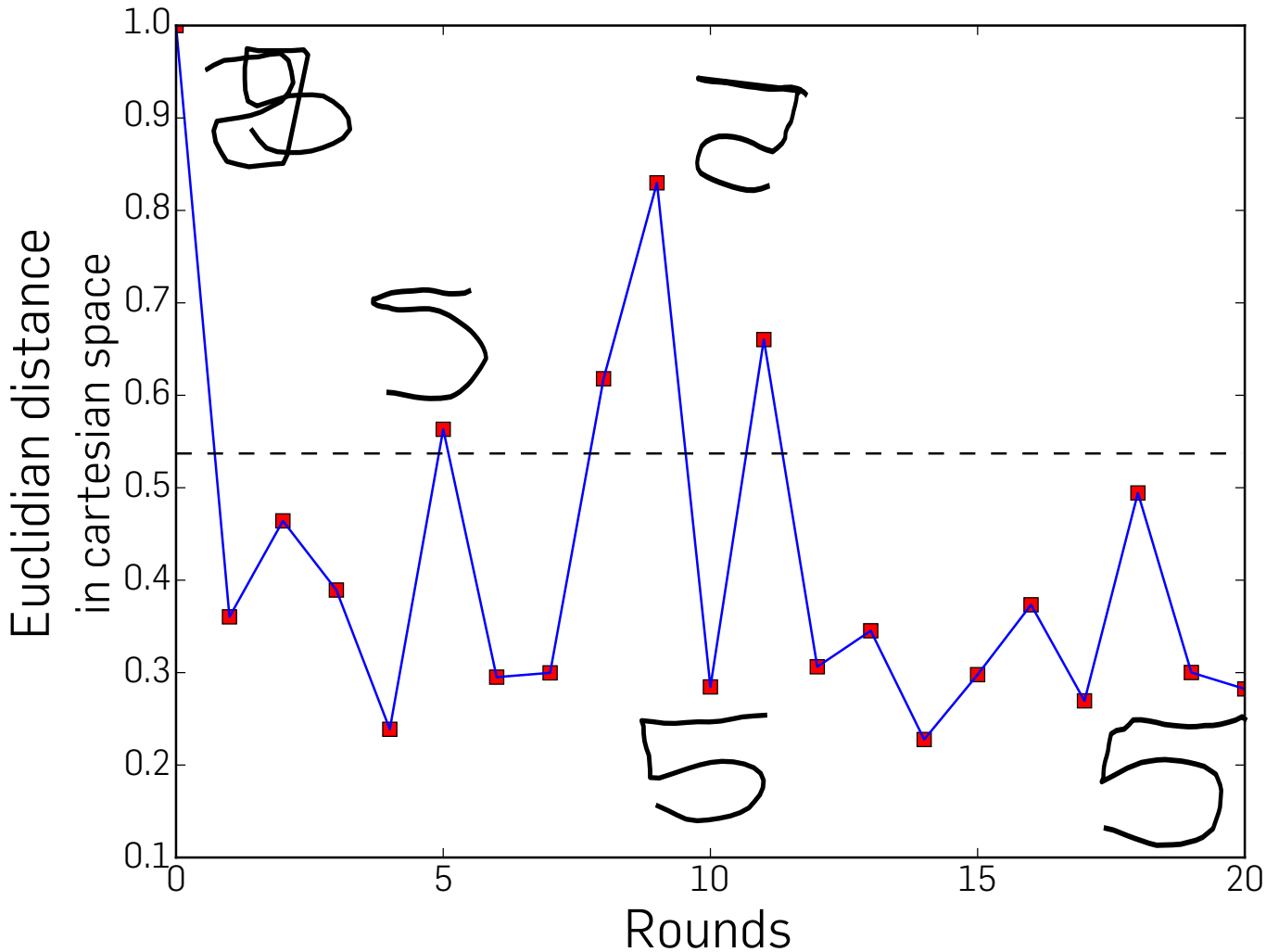
[Hood, Lemaignan et al. **When Children Teach a Robot to Write: An Autonomous Teachable Humanoid [...]** – HRI 2015]

BEFORE – AFTER





LEARNING TO DRAW A 5



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

THE ROBOT AS A SOCIAL AGENT

- **The robot as a cognitive agent is key here**
  - Protégé effect
  - metacognition
- **New role**: not a 'tool to teach robotics', not a facilitator
- (note: a tool for the teacher vs a social agent for the child!)
- Could we replace it by someone else? Not easily

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

THE NEXT STEPS

SCAFFOLDING THE RESEARCH

What are the next challenges for artificial social cognition?

- typical socio-cognitive tasks are either toy scenarios (i.e. do not mirror real-world siutation) or simple, constrainted tasks that do not reflect the complexity & dynamics of real-world interactions
- yet, research ***in the wild*** is difficult to conduct rigorously and to replicate
- $\Rightarrow$ I want to set **free play** as our next horizon: ***rich set of cognitive and social dynamics***; importance of motivation/drive; forces us to deal with uncertain and unexpected situations

THEORETICAL FRAMEWORK: STAGES OF PLAY

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

1. **Solitary (independent) play**: Playing separately from others, with no reference to what others are doing.
2. **Onlooker play**: Watching others play. May engage in conversation but not engaged in doing. True focus on the children at play.
3. **Parallel play** (adjacent play, social coaction): Playing with similar objects, clearly beside others but not with them (near but not with others.)
4. **Associative play**: Playing with others without organization of play activity. Initiating or responding to interaction with peers.
5. **Cooperative play**: Coordinating one's behavior with that of a peer. Everyone has a role, with the emergence of a sense of belonging to a group. Beginning of "team work."

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

WHAT SOCIAL BEHAVIOURS DO I AIM FOR?

- **behavioural alignment**: each time we play together, our interaction flow improves, becomes more natural
- **'natural' (i.e. emergent) turn-taking**: I 'just' know when it is my turn to act
- **'natural' protodeclarative pointing**: I 'just' know when I really need to draw your attention on something
- ability to pass **false-belief tasks**, including non-perceptual, non-physical, abstract ones,
- **replication of Parten's hierarchy** $\leftarrow$ ambitious!

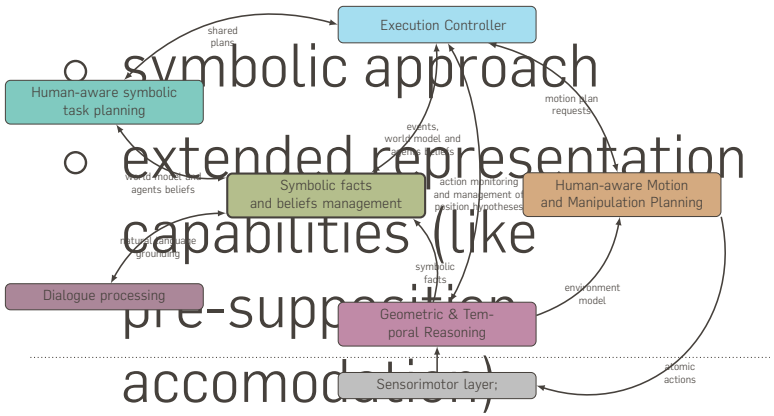
$\Rightarrow$ a research programme for the next 5-10 years

MY RESEARCH PLAN

Two tracks

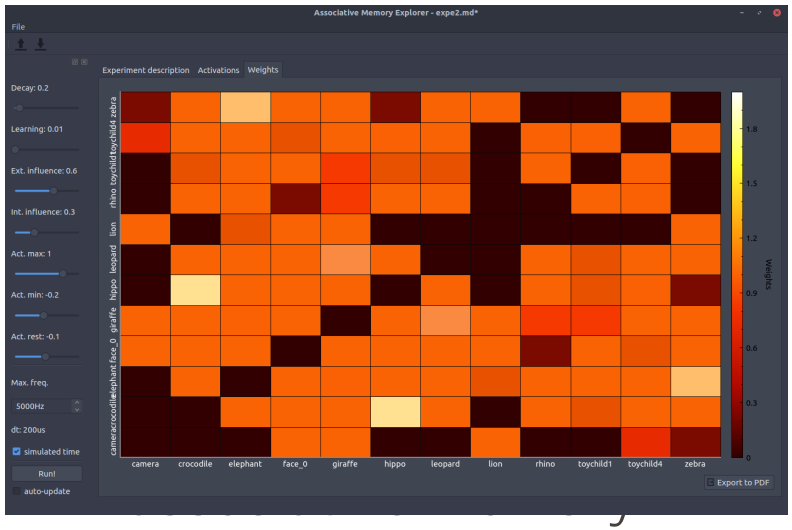
Practical

- symbolic approach
- extended representation capabilities (like pre-supposition, accomodation)



The “Edinburgh Architecture for Socio-Cognitive Robots”

Basic research



Emergent Artificial Social Cognition

hybridation

DEVELOPING CHILD-ROBOT INTERACTION AT EDINBURGH

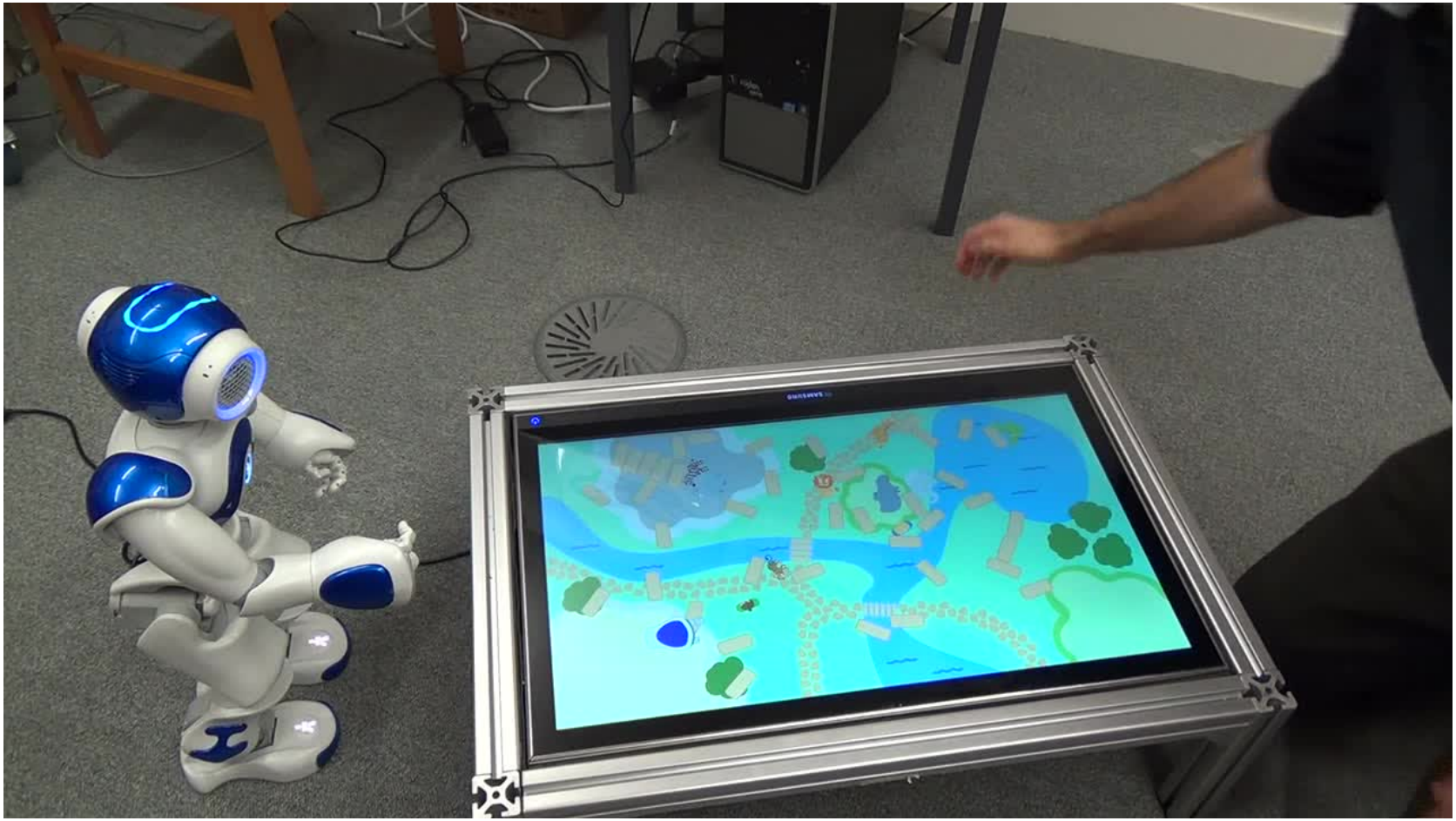
Create a **KidsLab**, within the Robotarium

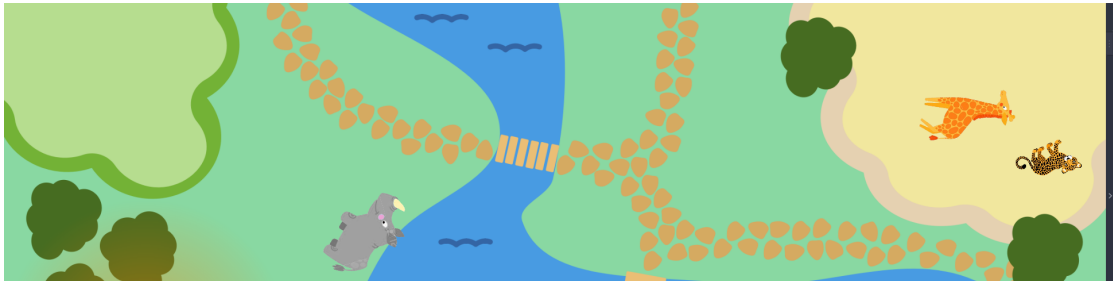
- experimental space suited for 4 to 13 years old
- fully equipped (recording, tracking, touch tables)
- in connection with UoE Wee Sciences lab and the RabLab

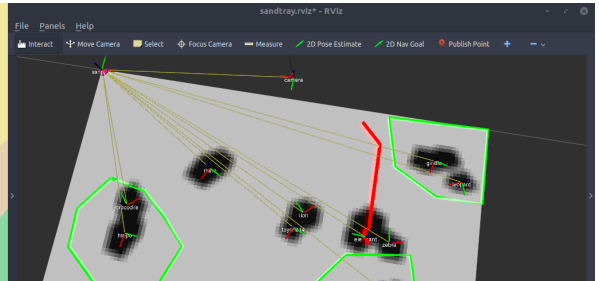
Establish collaboration with the (new) **Royal Hospital for Sick Children**

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

A FRAMEWORK FOR FREE PLAY







FUNDING

Short term

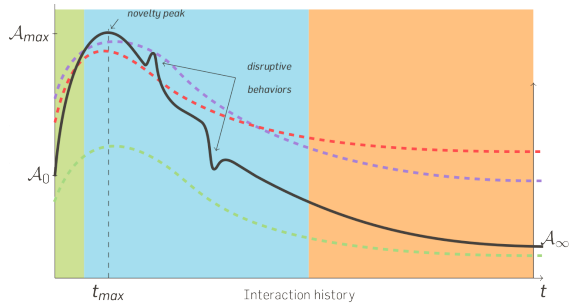
- **EPSRC First Grant**
  up to £125,000
  max 2 years, no closing date to submit
  covers equipment, but no PhD
  fund the **KidsLab** project
- **EPSRC Early Career Fellowship**
  max 5 years, no closing date to submit
  covers equipment and staff
  project on **Emergence of artificial social cognition**
  possible partnerships: developmental robotics (e.g. INRIA Flowers), cognitive neurosciences (e.g. Graziano lab), learning technologies (e.g. EPFL CHILI Lab)

Mid-term (within 2-3 years)

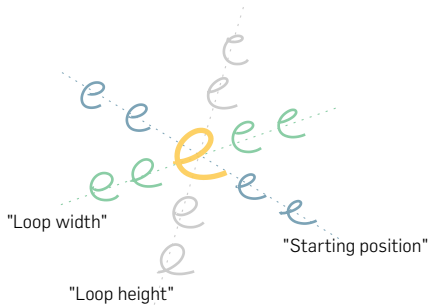
- **Co-Investigator on a H2020 project**
  close contacts with IST Lisbon, LAAS-CNRS, Plymouth, EPFL, INRIA Bordeaux

WHAT ELSE?

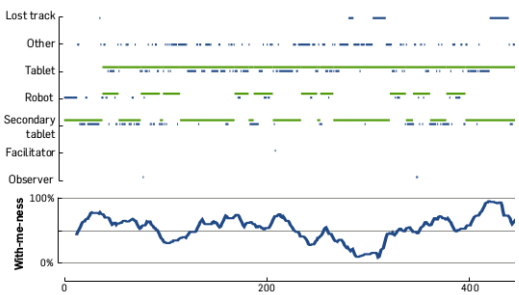
Social HRI on the field

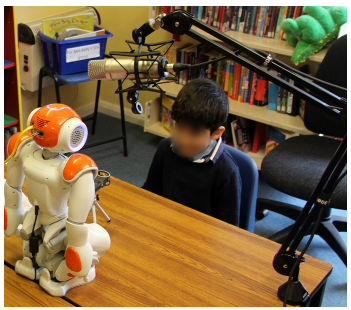


Robots for Learning



Methodology





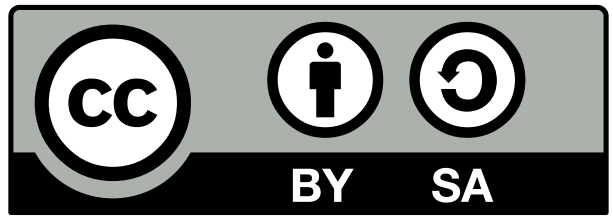
Expectations
How do you imagine a robot? *What could it look like?* *Have you ever seen a robot before?*
Impression
When you first saw R, what did you think? *Is R a robot? How do you know?* *Did you expect R would come over to you when you call it?* *What happened when you put the domino in the box?*
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)?


Not a lonely work!

Thank you!

Séverin Lemaignan
severin.lemaignan@plymouth.ac.uk



You can download the sources of this presentation here:
https://github.com/severin-lemaignan/presentation-cognitive-robotics

SUPPLEMENTARY MATERIAL

5. Performing in Human Environments

6. pyRobots Example

7. Dialogue Grounding

8. Child-Robot Interaction for Learning

9. Child-Robot Interaction: the Practical Side

10. Reframing the research

11. Sketching a model

12. Experimental investigation

13. Theories on theory of mind

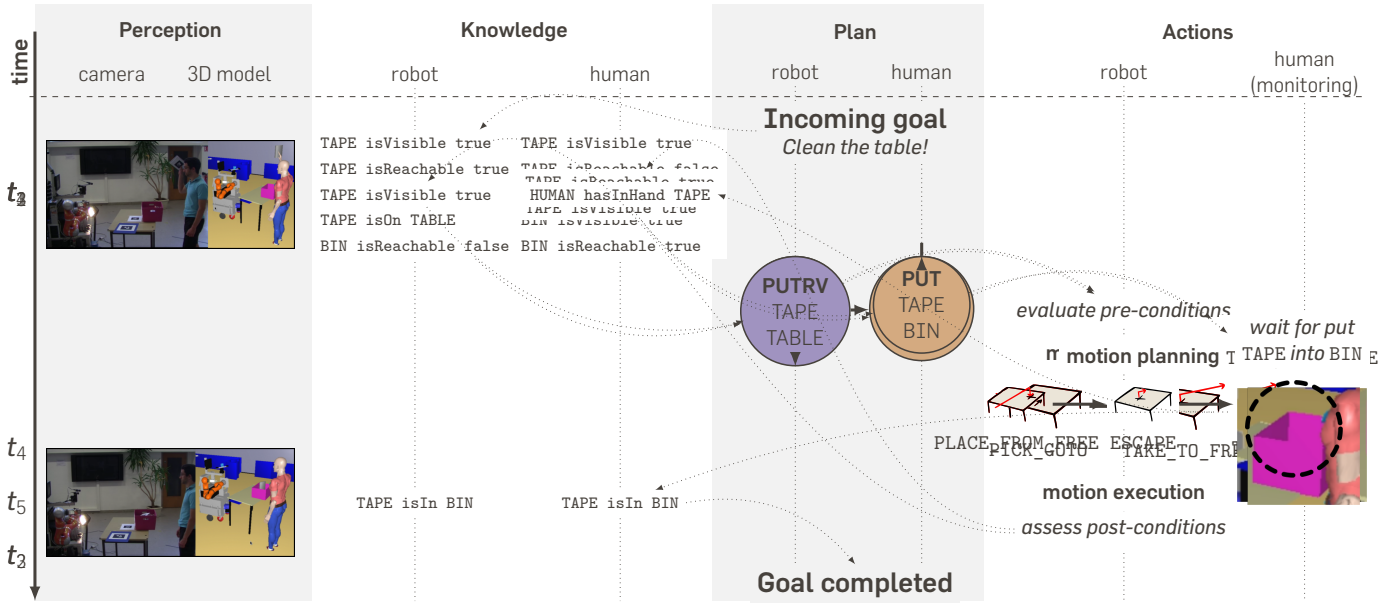
14. Dynamics of Interaction

15. Cognitive Architectures

PERFORMING IN HUMAN ENVIRONMENTS

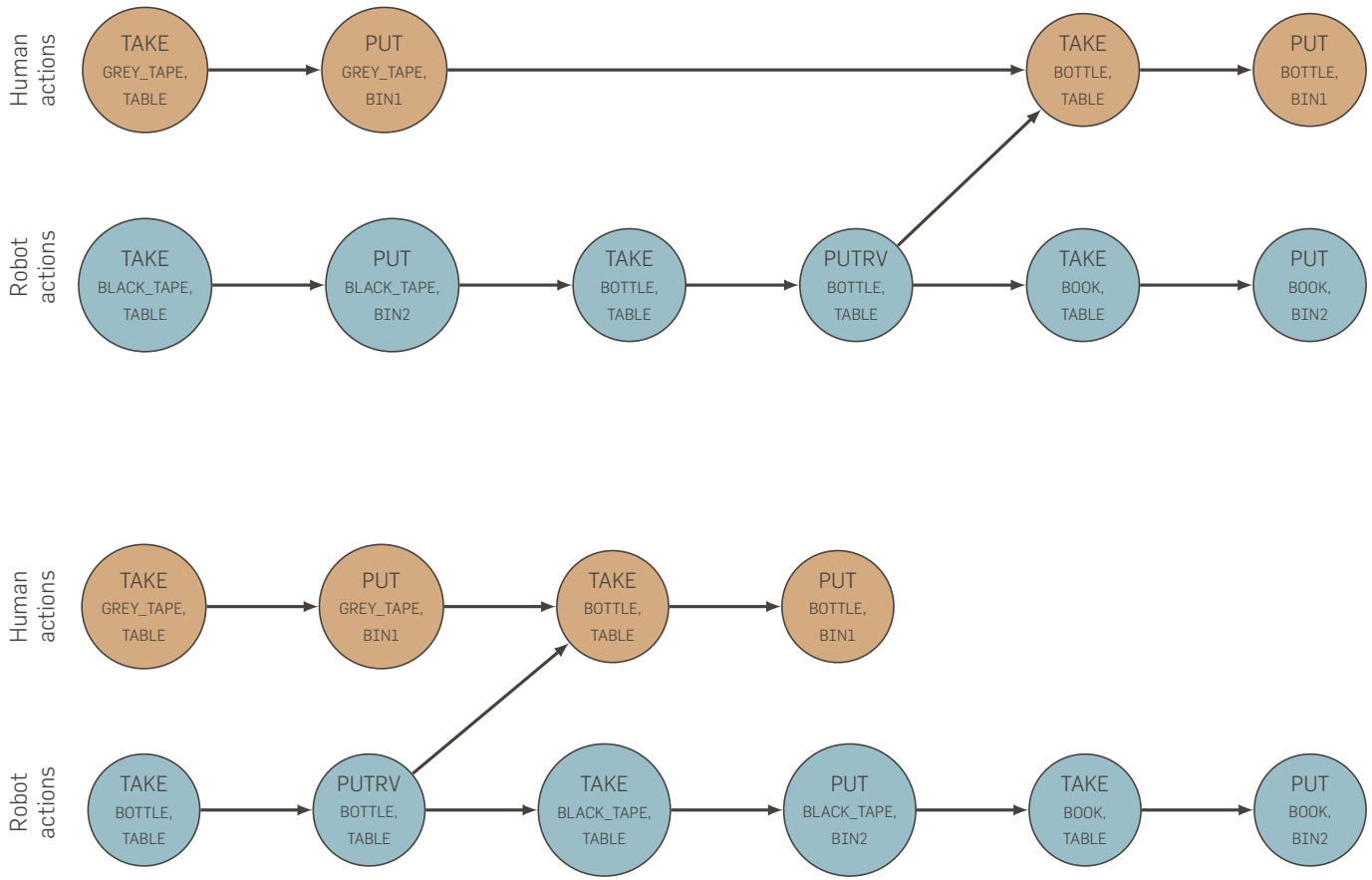

LAAS-CNRS

"CLEANING THE TABLE"...

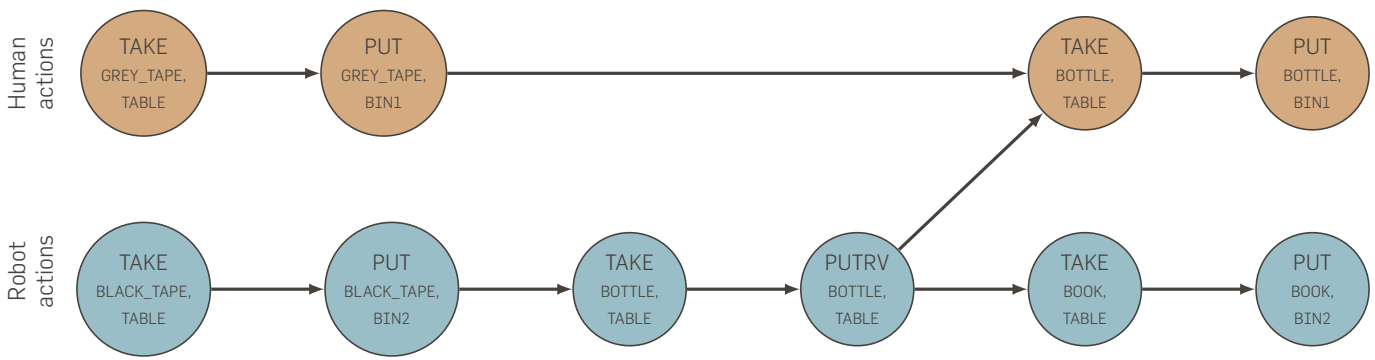


[Lemaignan et al., **Artificial Cognition for Social Human-Robot Interaction: An Implementation**, Artifical Intelligence, 2016]

PLANNING FOR THE HUMAN



PLANNING FOR THE HUMAN




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
```

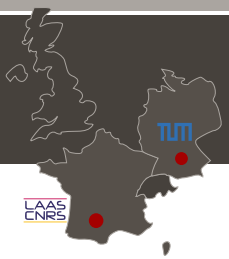
PYROBOTS

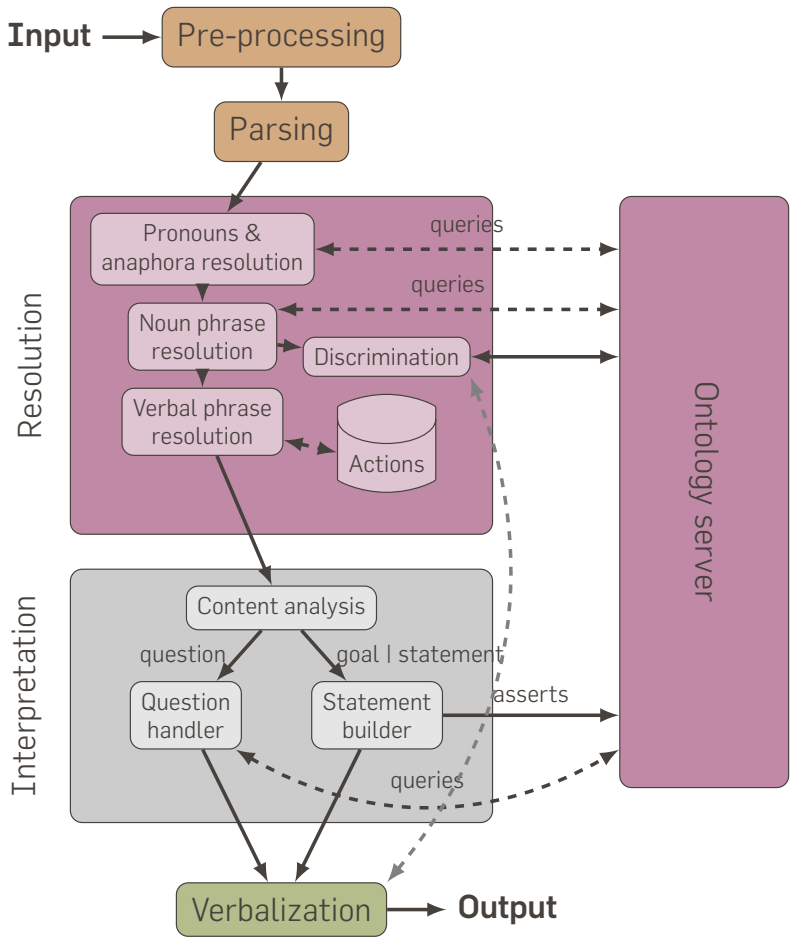
```python
from robots import GenericRobot
from robots.concurrency import action, ActionCancelled
from robots.resources import Resource, lock

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
```

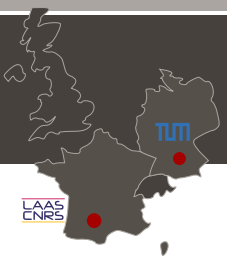
DIALOGUE GROUNDING

DIALOGUE GROUNDING





DIALOGUE GROUNDING



"Give me the book on the table"

↓

`me` → `baby_1`

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

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

↓

```
baby_1 desires action1,
action1 type Give,
action1 performedBy myself,
action1 actsOnObject harry_potter,
action1 receivedBy baby_1
```

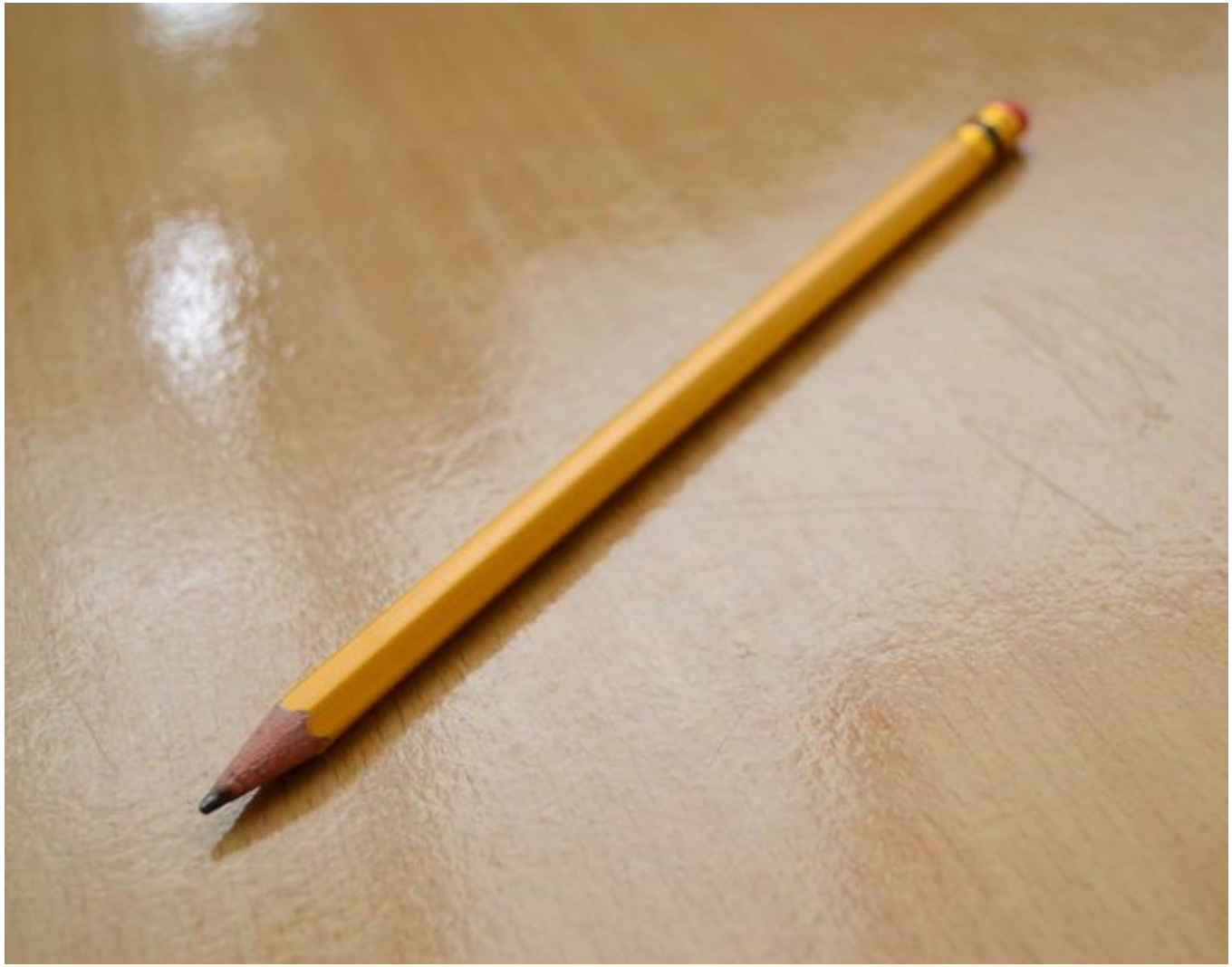
[Lemaignan et al. **Grounding the Interaction: Anchoring Situated Discourse in Everyday HRI** – Intl. J. of Social Robotics 2011]

CRI FOR LEARNING

SOCIAL OR NOT SOCIAL?

NON-SOCIAL INTERACTION

What is the most effective learning tool in a classroom?



Pens and paper are pervasive...

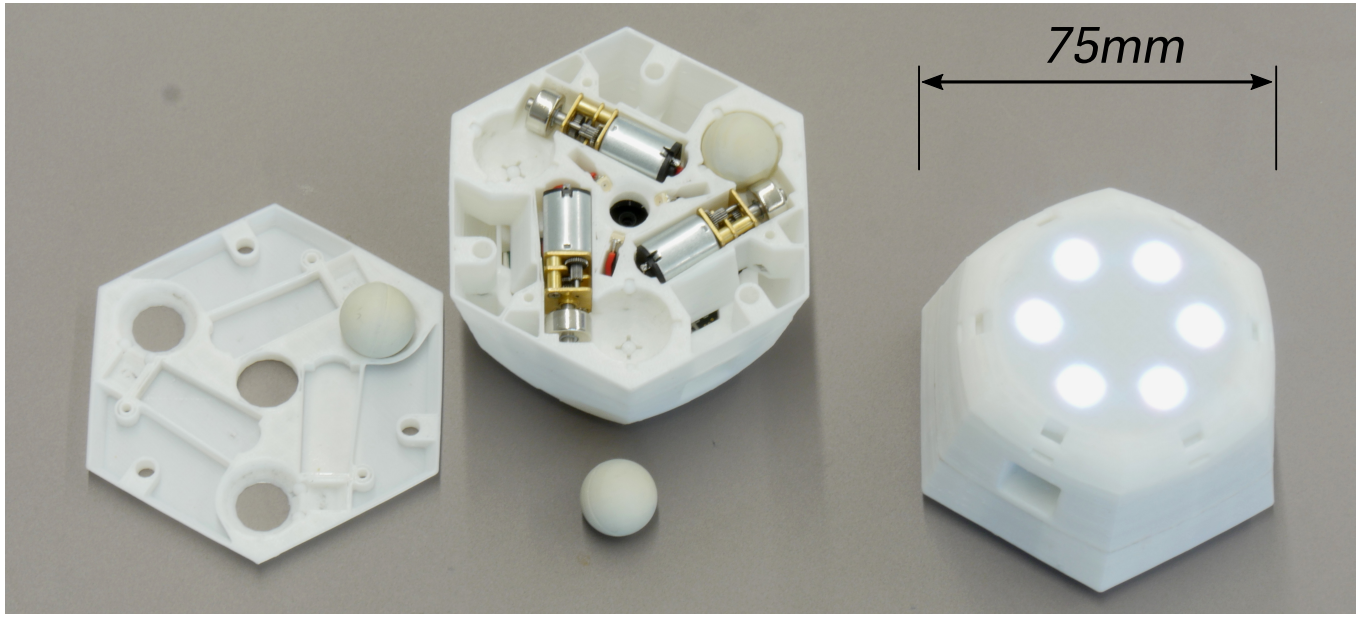
...so should the robots

CELLULO: DESIGN PRINCIPLES

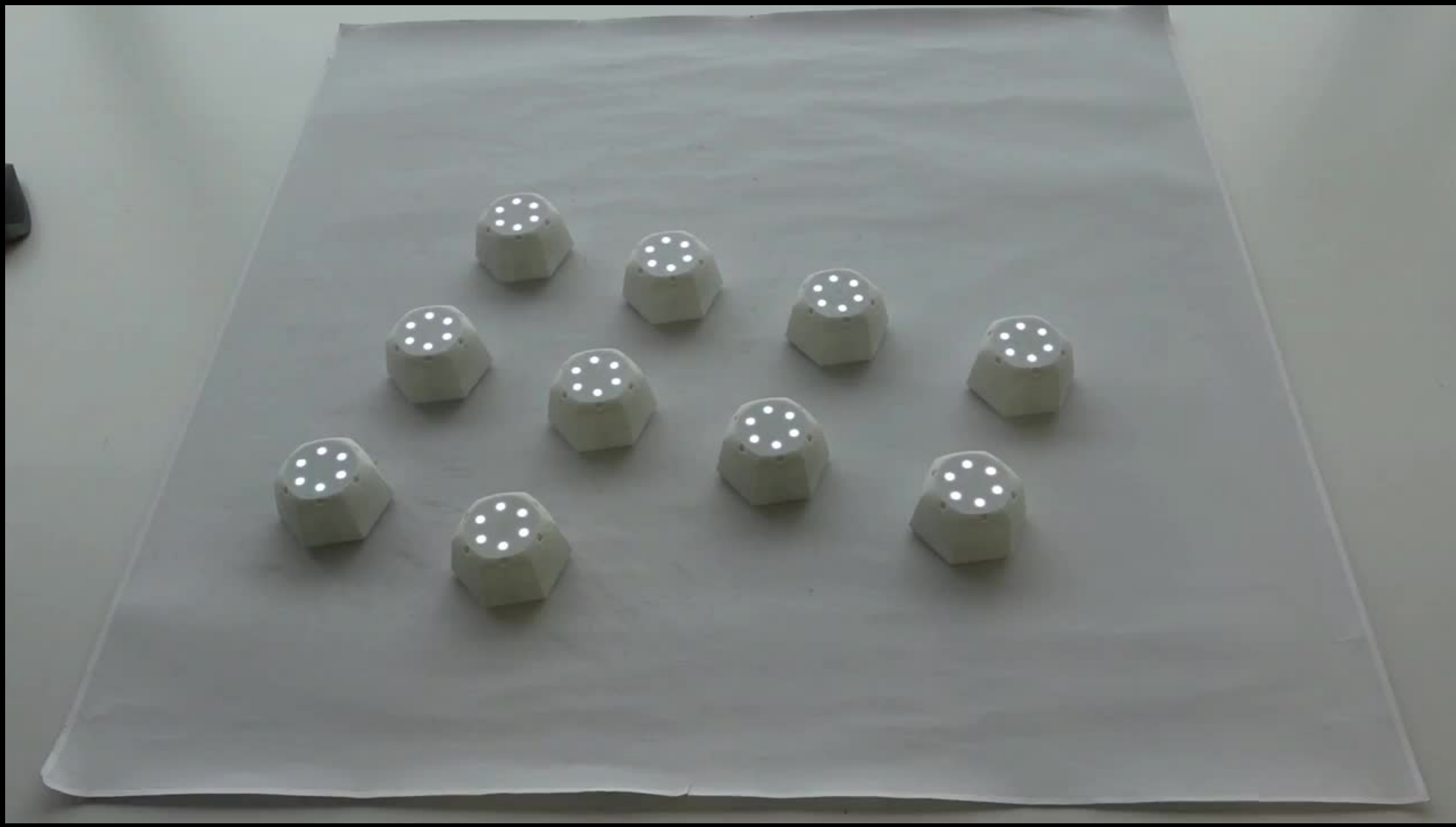
- **ubiquitous**: a pervasive yet unremarkable tool that blend into the daily learning routine; has to be trustworthy (i.e. reliable), readily replaceable (i.e. cheap, no affective bonding), intuitive (i.e. few simple affordances)
- **versatile**: applicable to a broad range of learning scenarios; the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



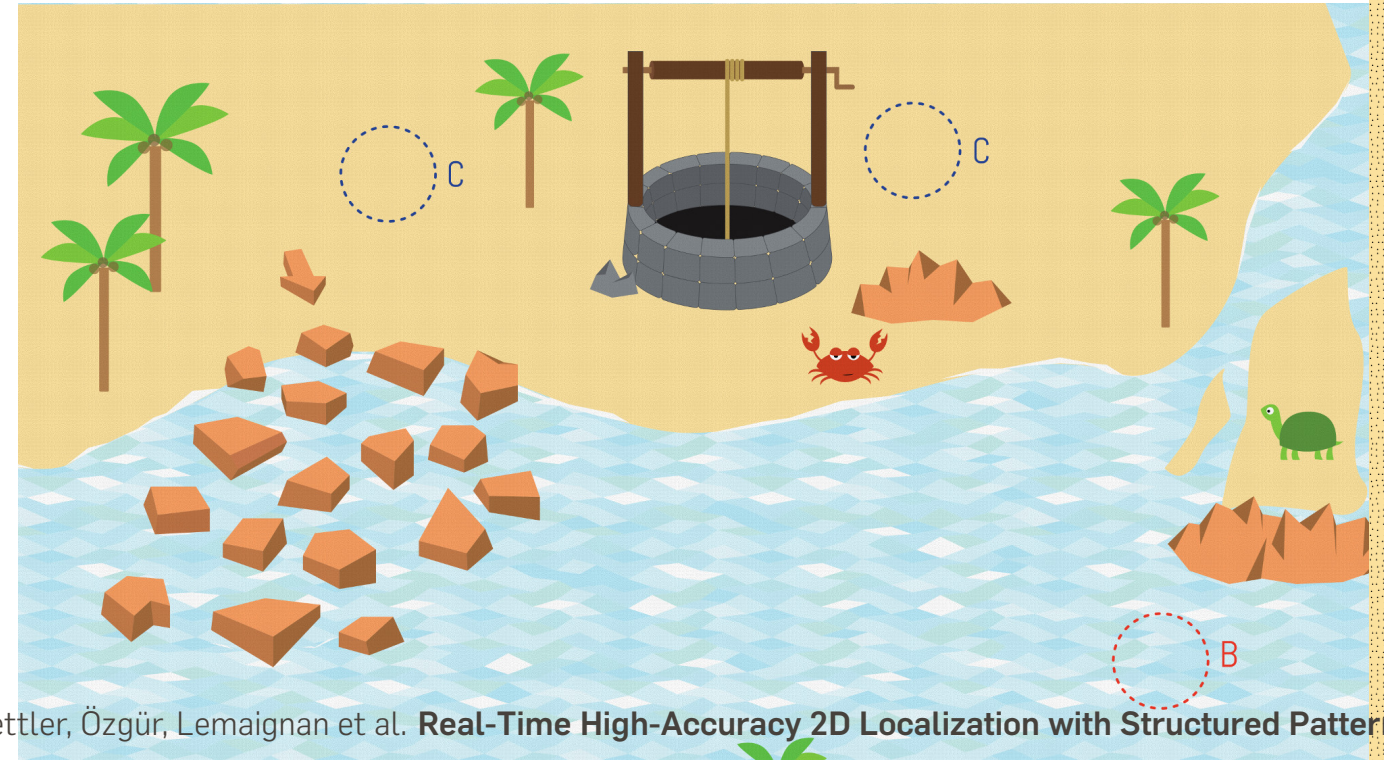
- Holomonic motion
- Sub-mm absolute localisation (no external hardware)
- Haptic feedback + tactile RGB LED buttons
- Bluetooth
- Affordable (prototype: €125)



INTERACTION WITH THE PAPER

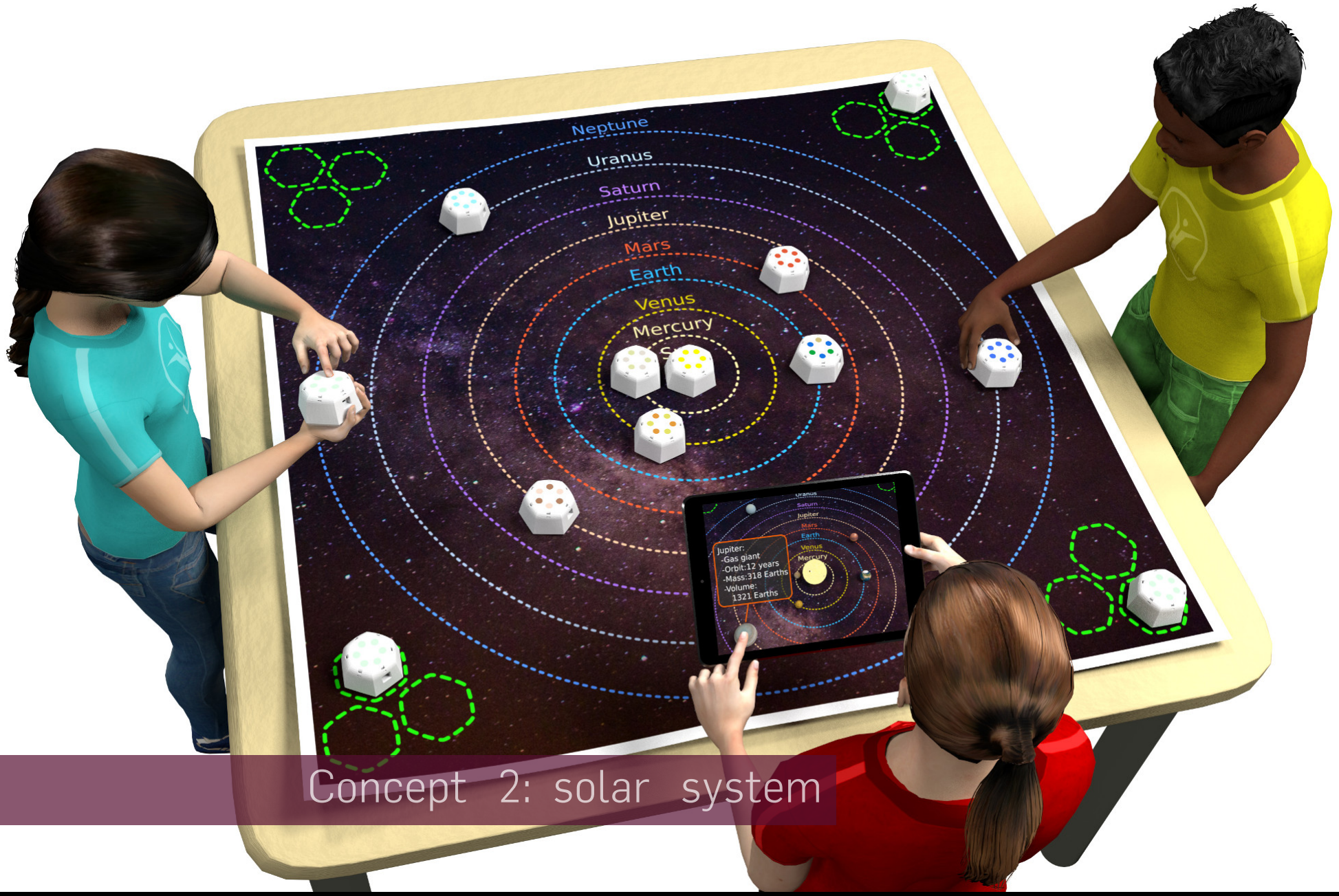
Critically, Cellulo is meant as an **interaction between (classroom-friendly) paper and the robots**.

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



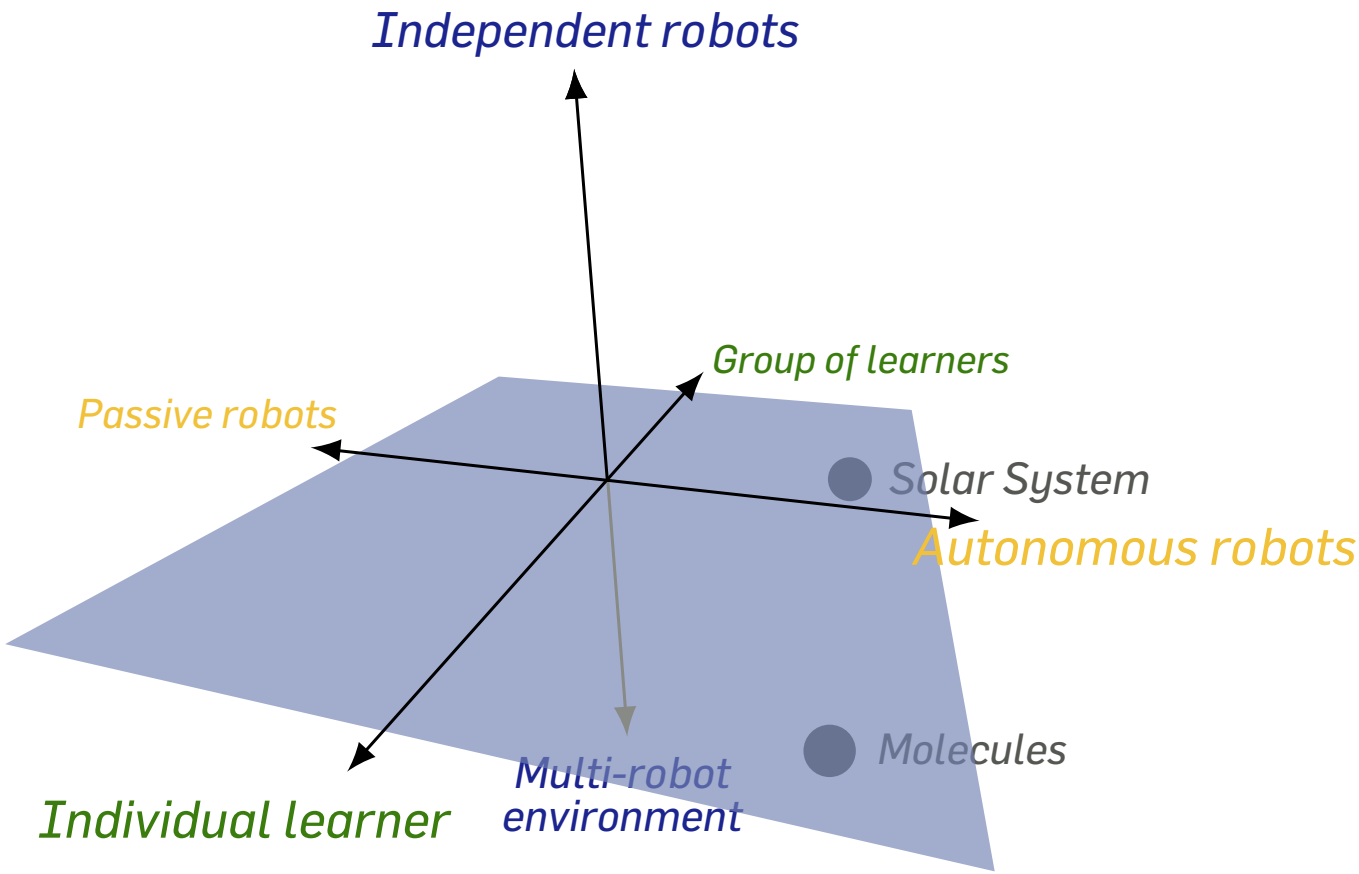
[Hostettler, Özgür, Lemaignan et al. **Real-Time High-Accuracy 2D Localization with Structured Patterns** – ICRA 2016]


$O_2 + 2H_2 = 2H_2O$
Reactor
Reactor 1
Chemistry Lab
Concept 1: chemistry



Concept 2: solar system

INTERACTION DESIGN SPACE



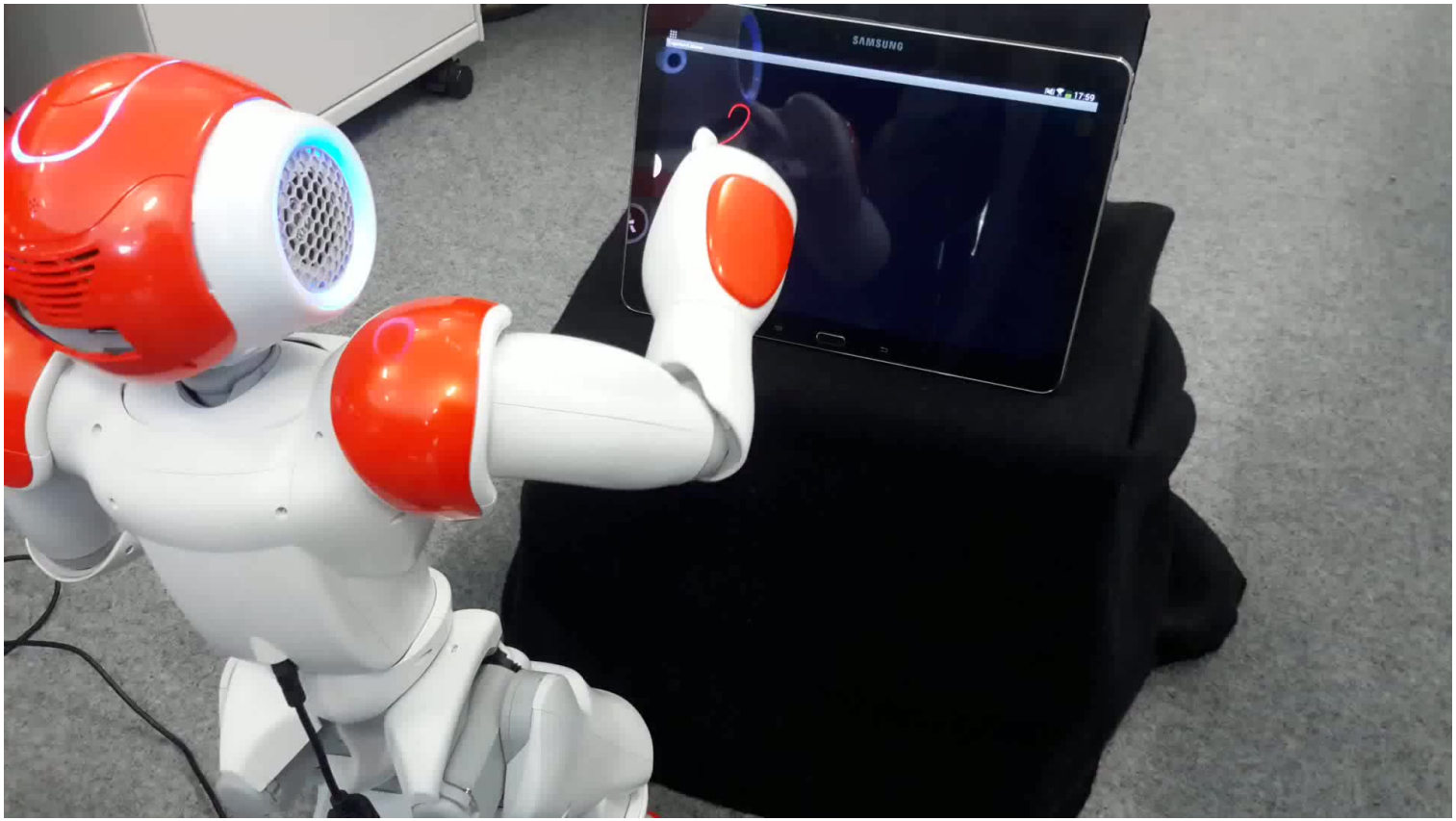
[Özgür, Lemaignan et al. **Cellulo: Versatile Handheld Robots for Education** – HRI 2017]

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

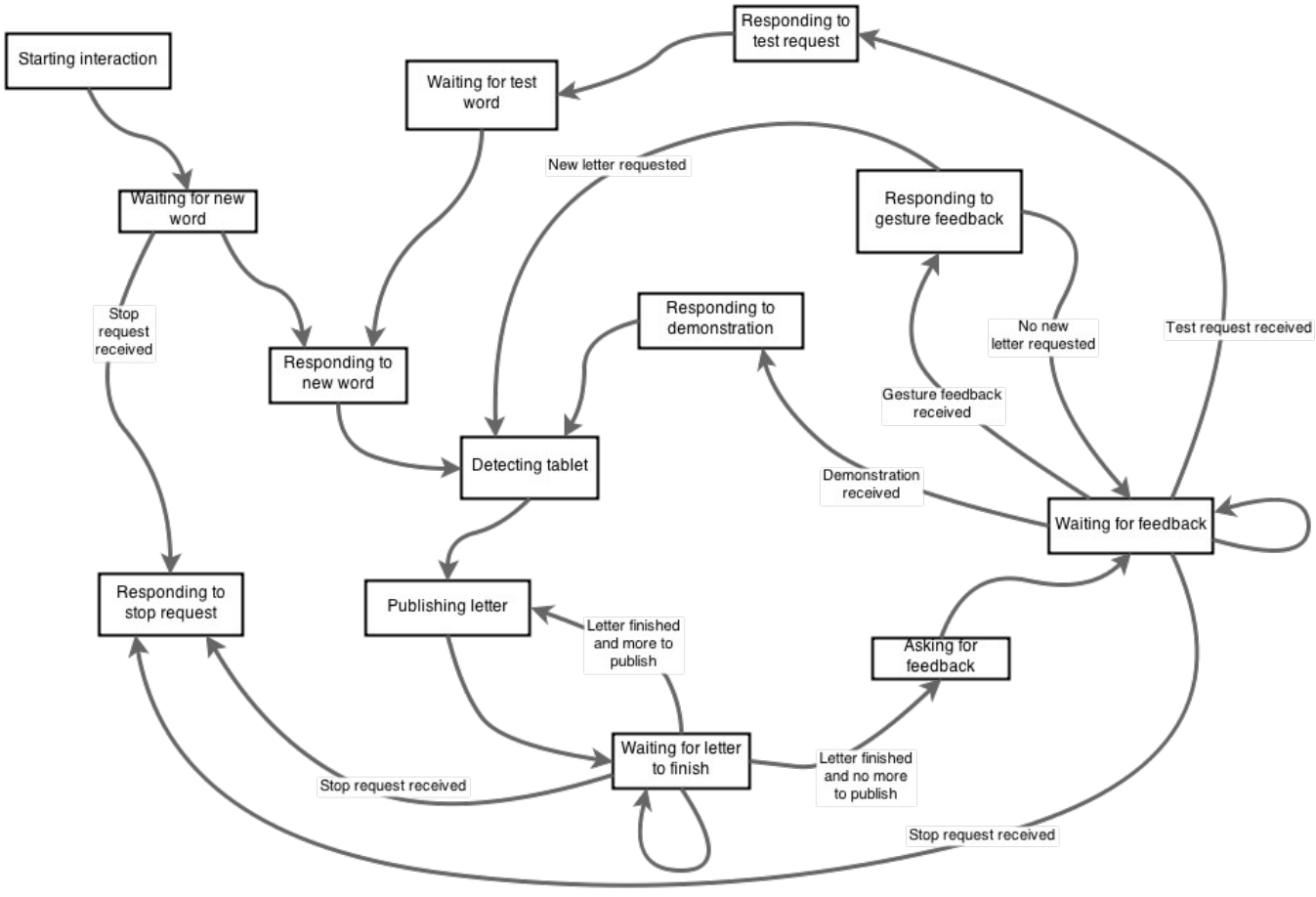
TECHNICAL CHALLENGES

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

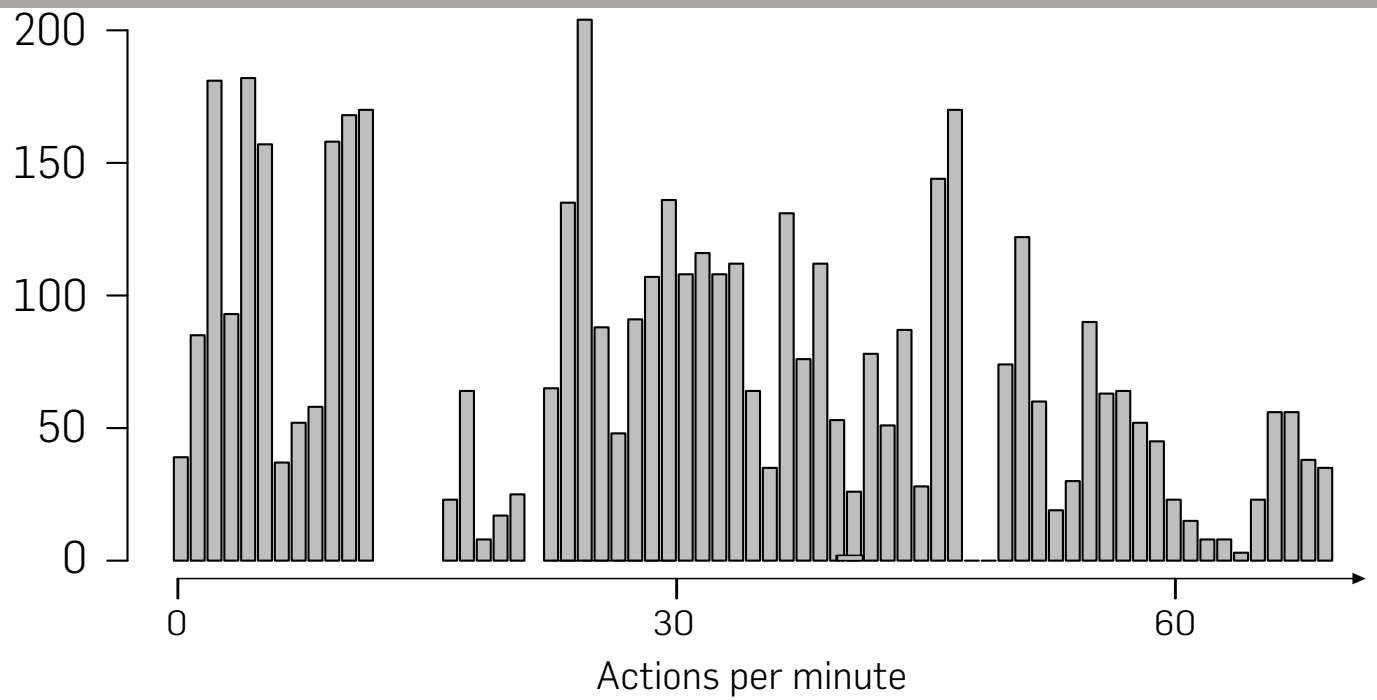
COWRITER IMPLEMENTATION



[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

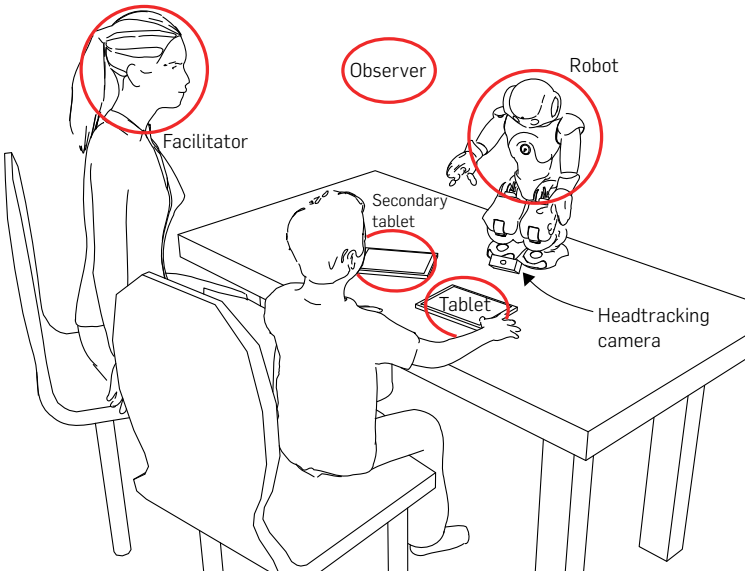
“With-me-ness”: real-time estimation of surface engagement

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

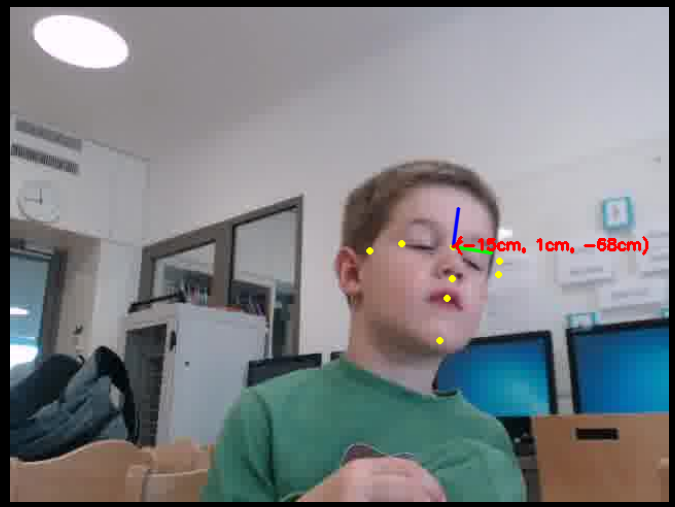
EXPECTED FOCUS

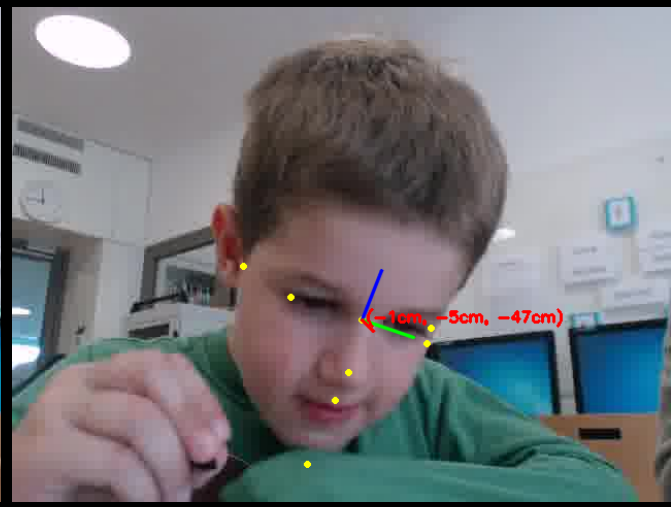
Example for the CoWriter task:

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

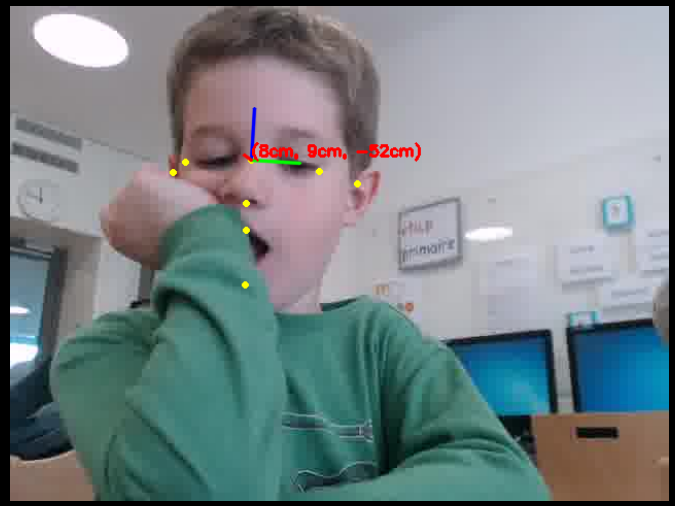


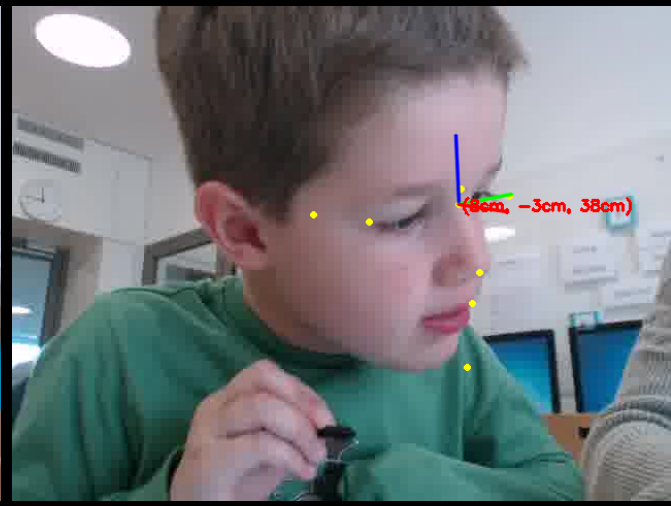

(−6cm, −7cm, −76cm)


(−15cm, 1cm, −68cm)

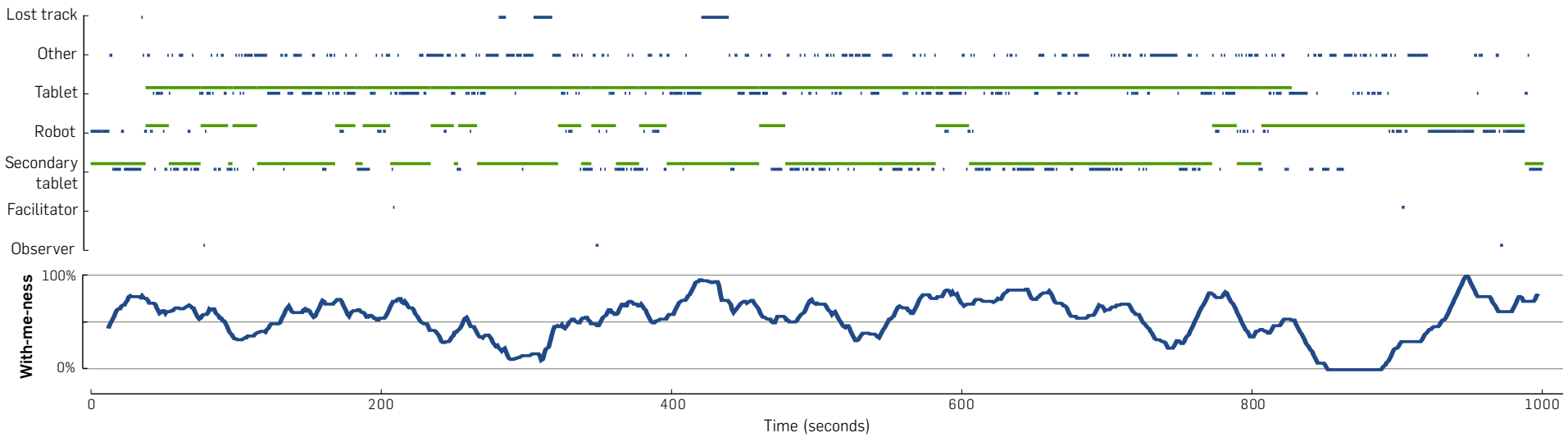

(−1cm, −5cm, −47cm)


(0cm, 8cm, −49cm)


(8cm, 9cm, −52cm)


(8cm, −3cm, 38cm)

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**
- $\Rightarrow$ **relative** metric!

◂ Supplementary material

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

CONSTRUCTS FOR COGNITIVE PERCEPTION ANALYSIS

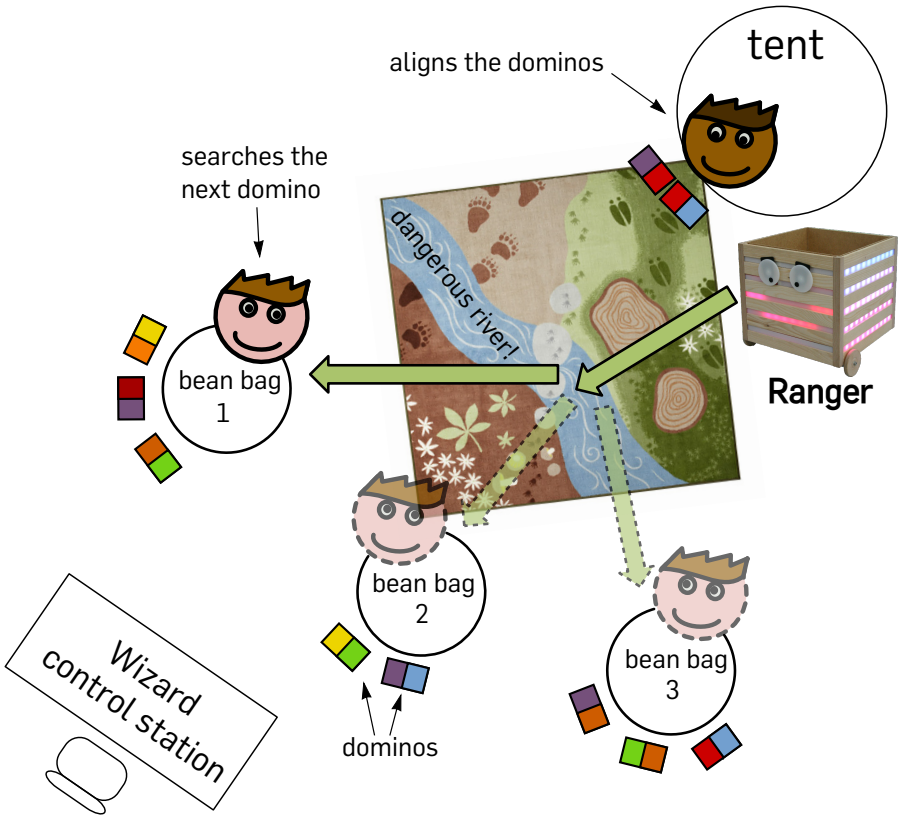
Expectations
How do you imagine a robot? *What could it look like?* *Have you ever seen a robot before?*
Impression
When you first saw R, what did you think? *Is R a robot? How do you know?* *Did you expect R would come over to you when you call it?* *What happened when you put the domino in the box?*
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

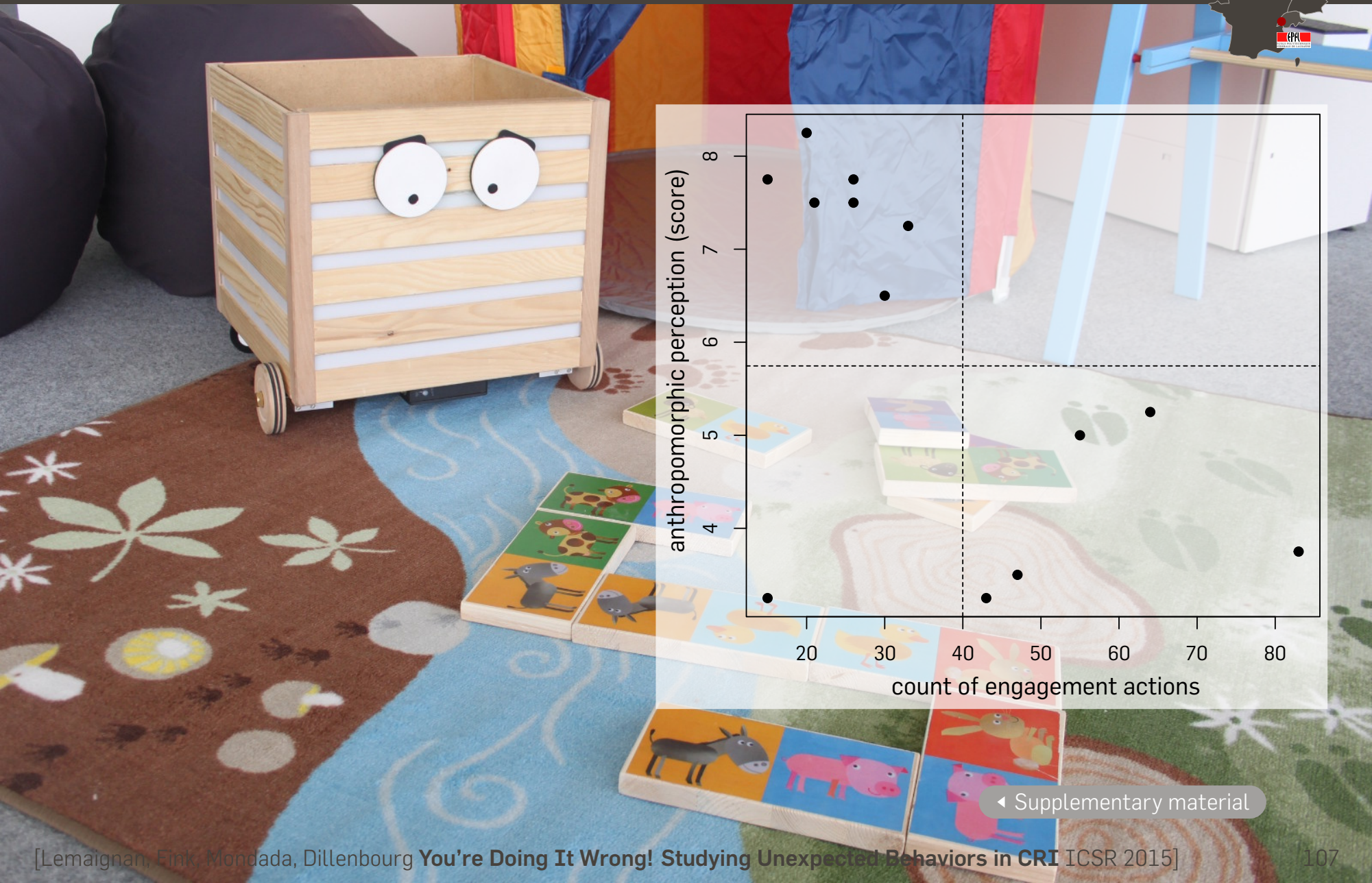
BEHAVIOUR VS PERCEPTION?

Any relation between the behavioural and perceptual measurements?



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

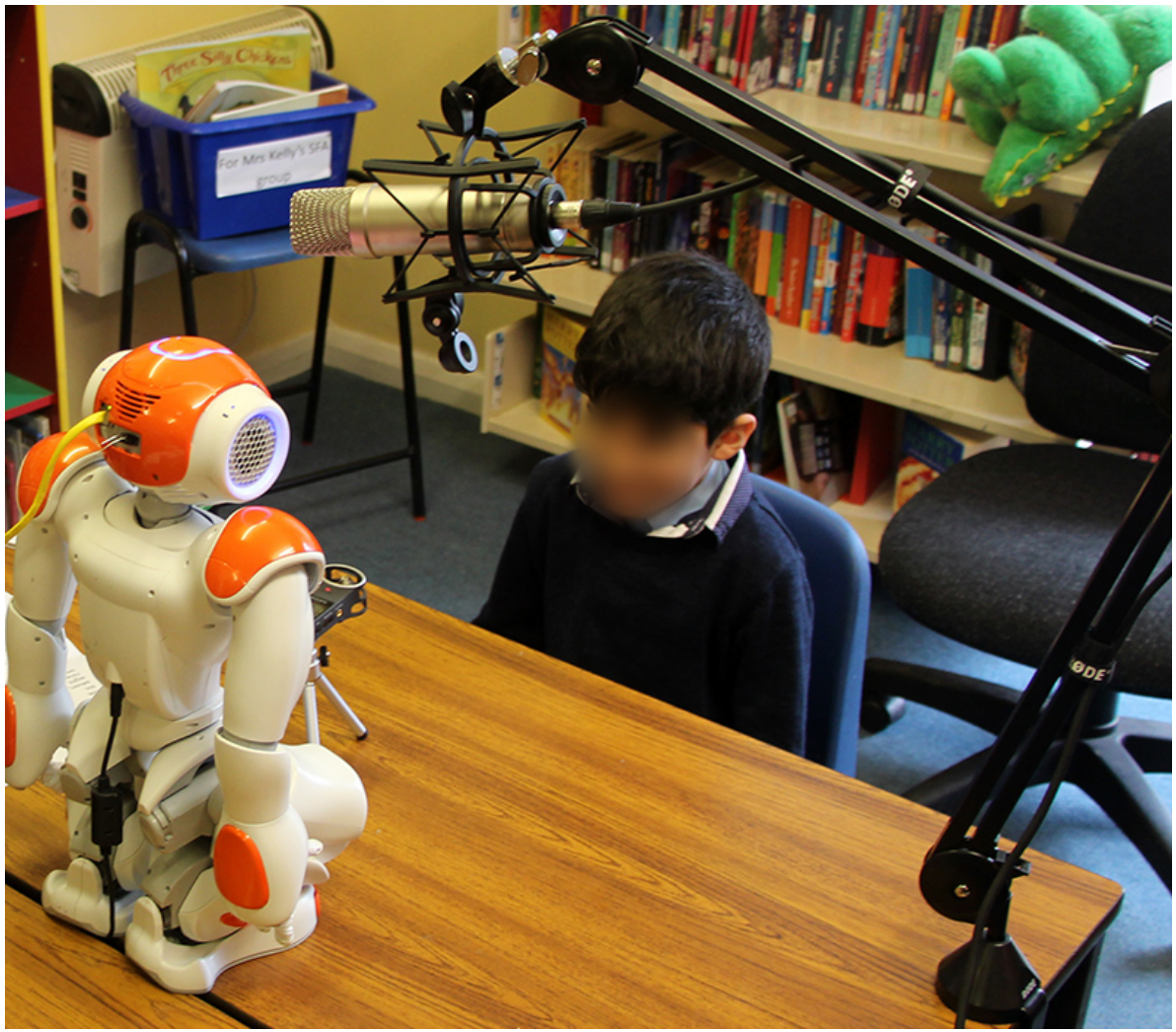
ANTHROPOMORPHISM != ENGAGEMENT



◂ Supplementary material

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

AUTOMATIC SPEECH RECOGNITION WITH CHILDREN



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

Google Bing Sphinx Nuance

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

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**

...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?

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"*

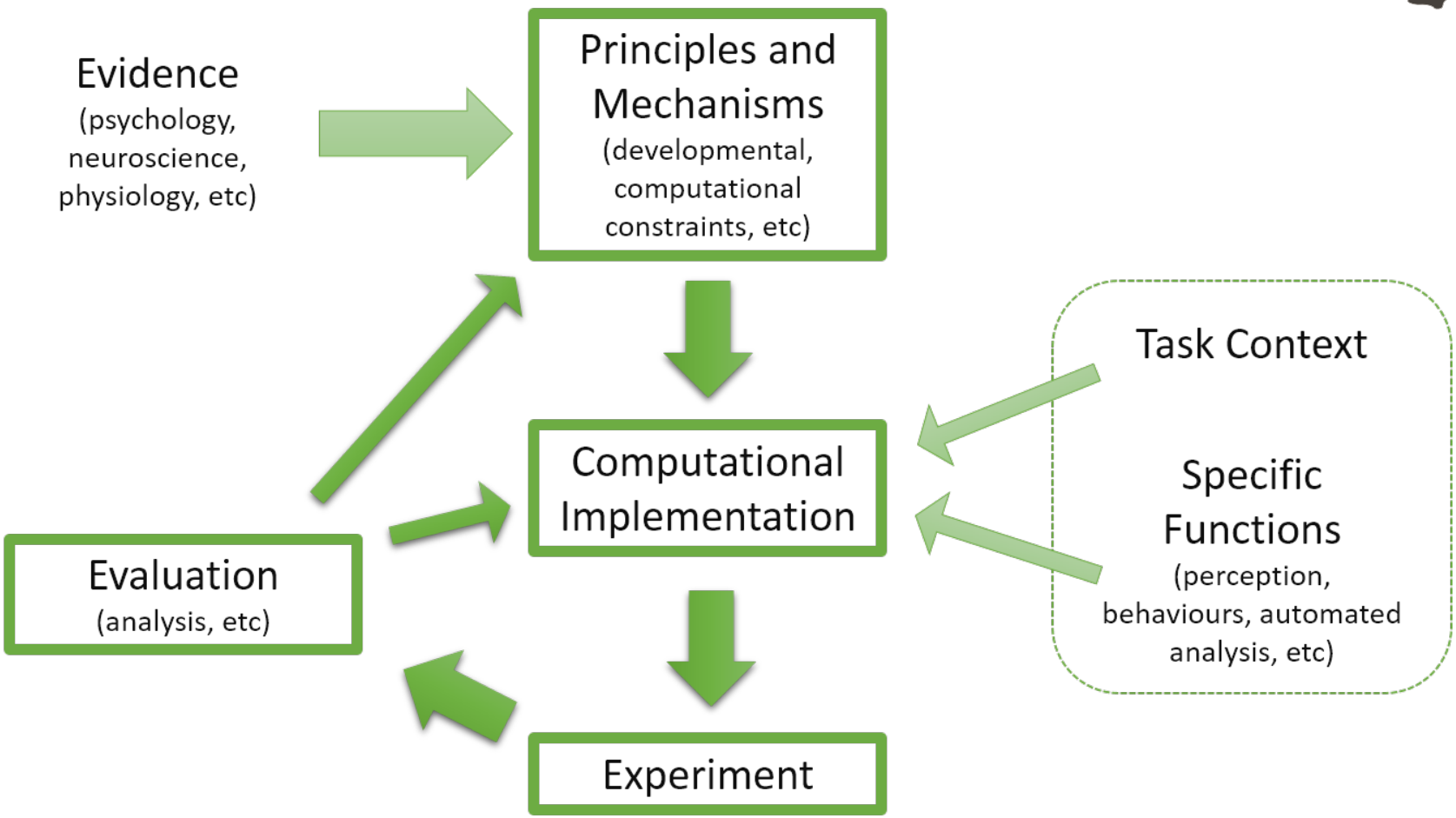
A model's value is gained from

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

...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]

COGNITIVE ARCHITECTURE AS A METHODOLOGY



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

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')

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".

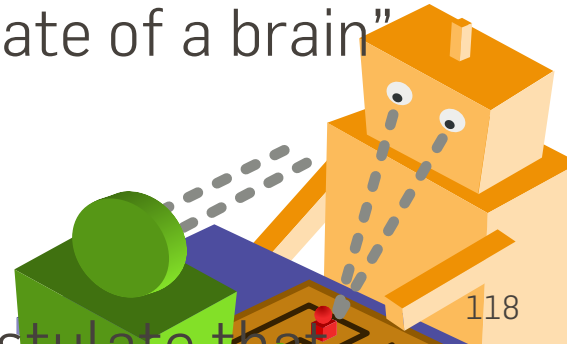
[see discussion in Vernon, **Artificial Cognitive Systems: A Primer** 2014]

MODELING OTHERS' MIND?

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

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]

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**
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

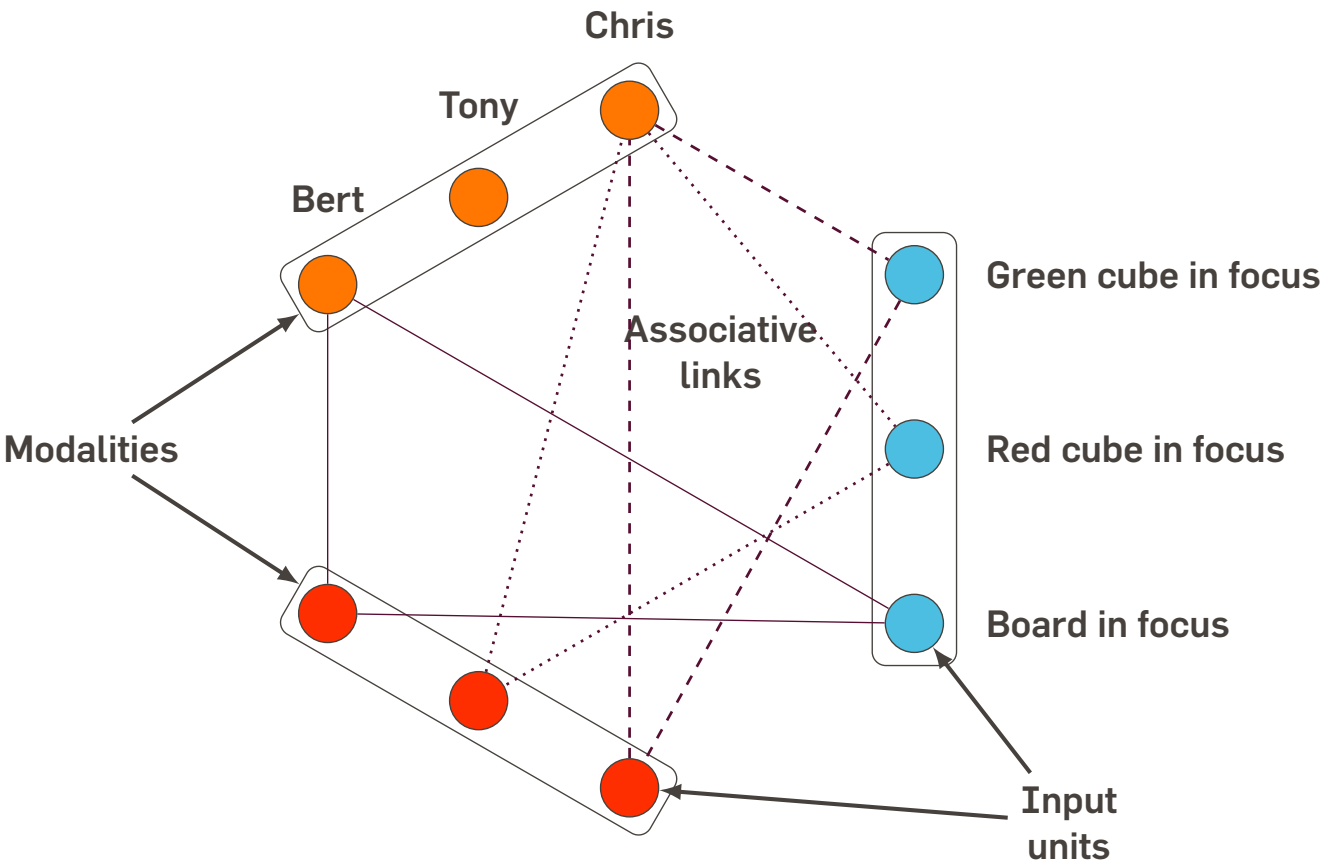
- *phenomenal experience* modelled in a connectionist (sub-symbolic) fashion (associative memory network)
- *access conciousness* in a cognitivist (symbolic) fashion (typically, an ontology or epistemic logics)

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

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]

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

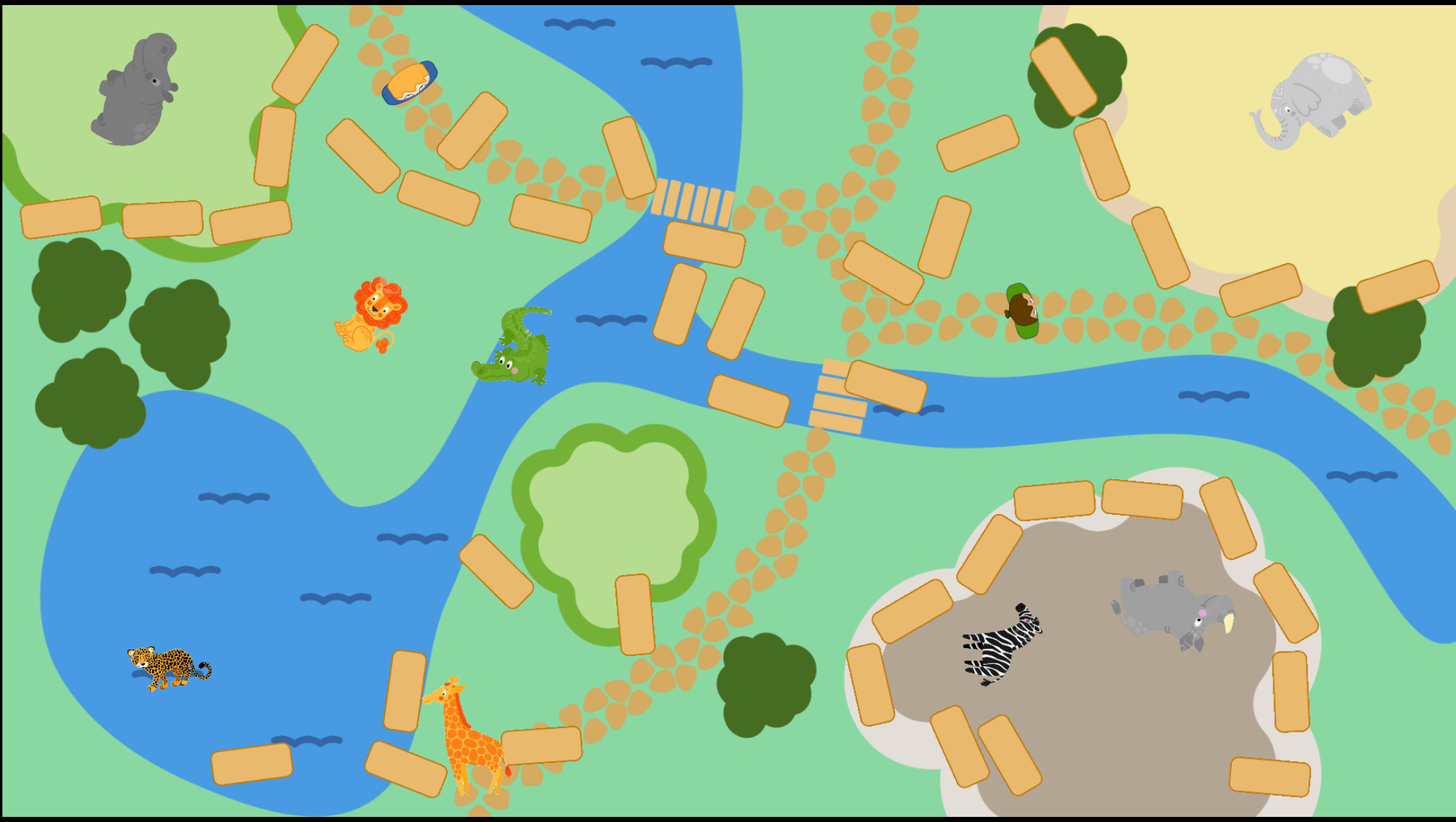
Hypothesis 3': representations of one's and others' minds, along with a situation (physical and temporal environment), are ***necessary*** and ***sufficient*** conditions for social behaviours to emerge

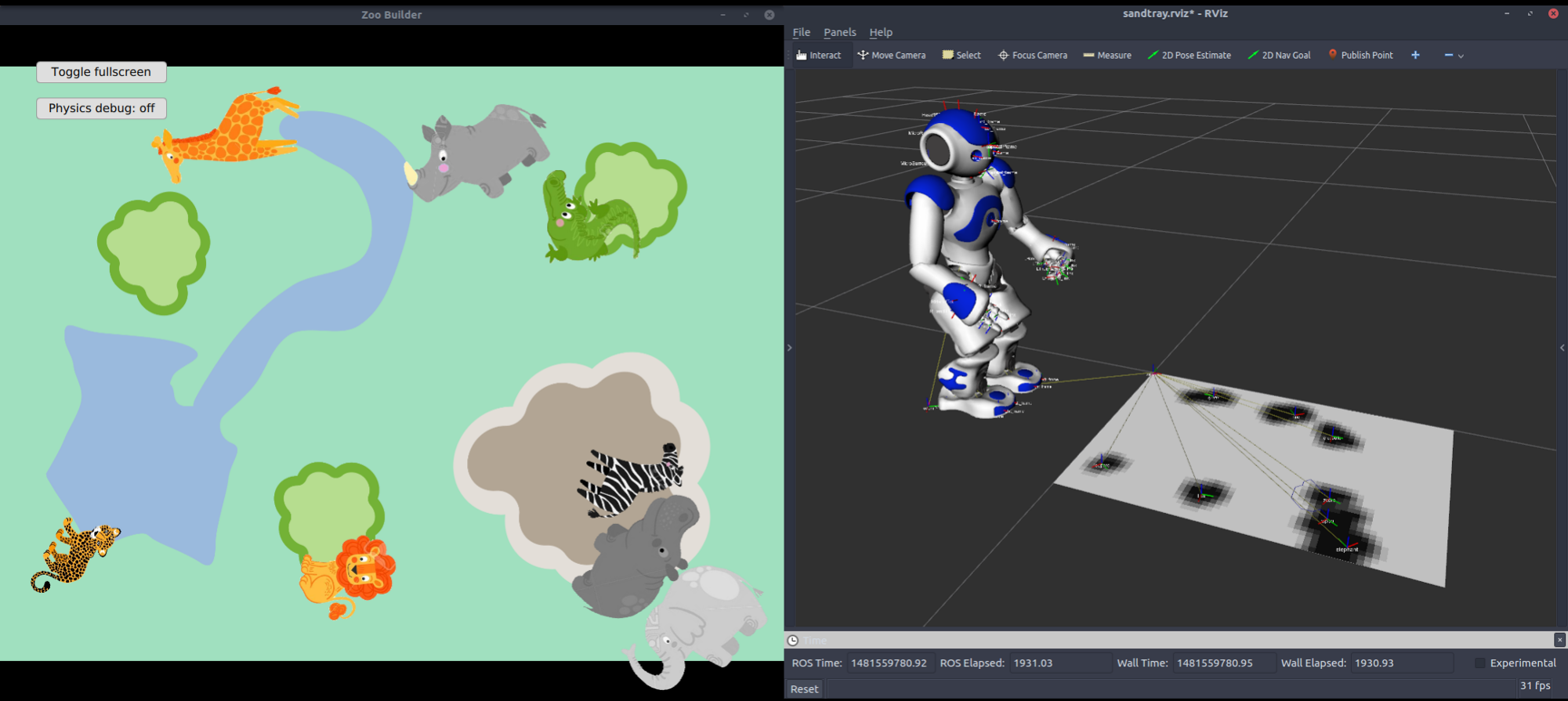
How? ...recruit me to discover ;-)

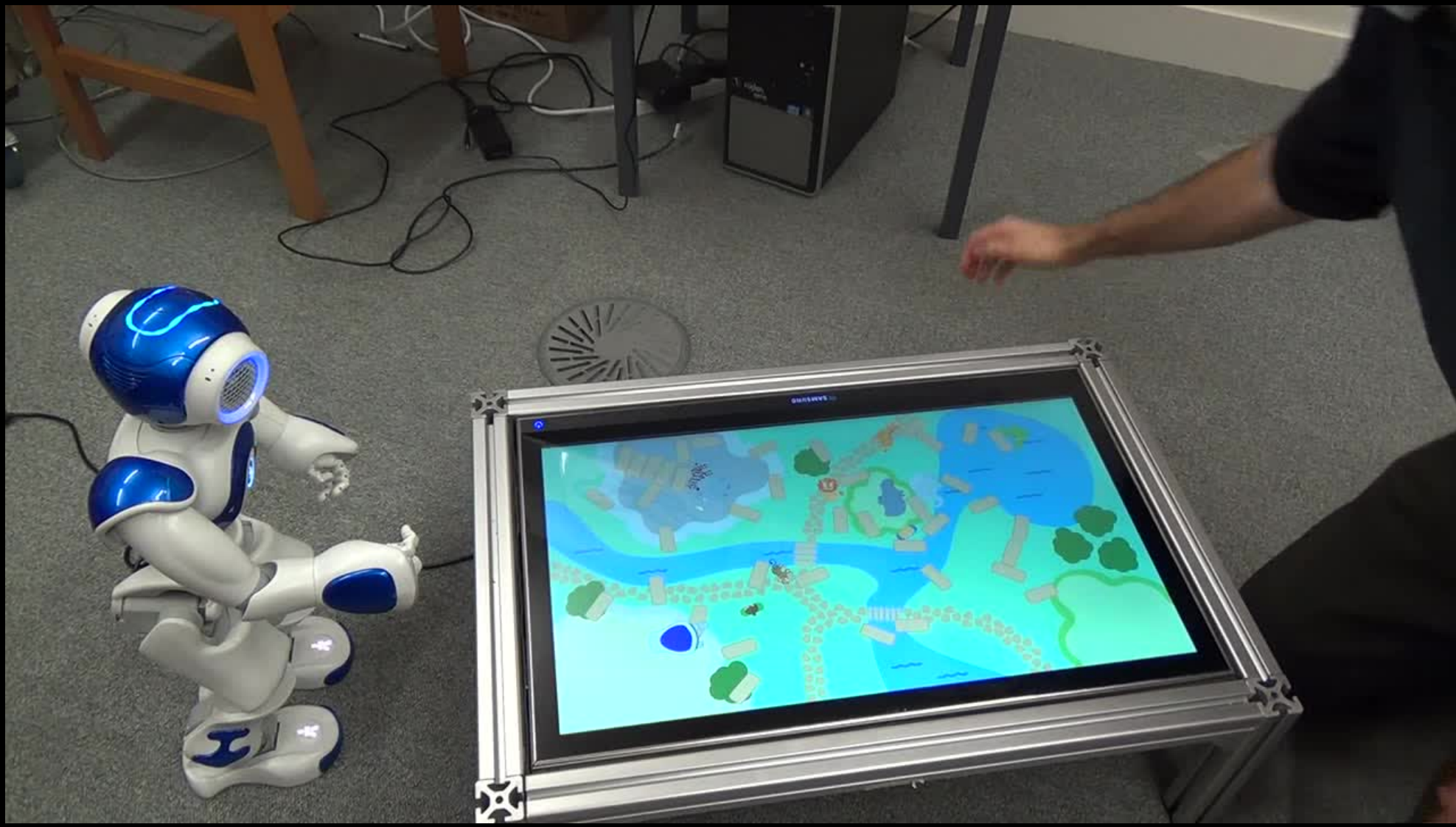
EXPERIMENTAL INVESTIGATION

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




Zoo Builder
Toggle fullscreen
Physics debug: off
sandtray.rviz* - RViz
File
Panels
Help
Interact
Move Camera
Select
Focus Camera
Measure
2D Pose Estimate
2D Nav Goal
Publish Point
Time
ROS Time: 1481559780.92
ROS Elapsed: 1931.03
Wall Time: 1481559780.95
Wall Elapsed: 1930.93
Experimental
Reset
31 fps



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?

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

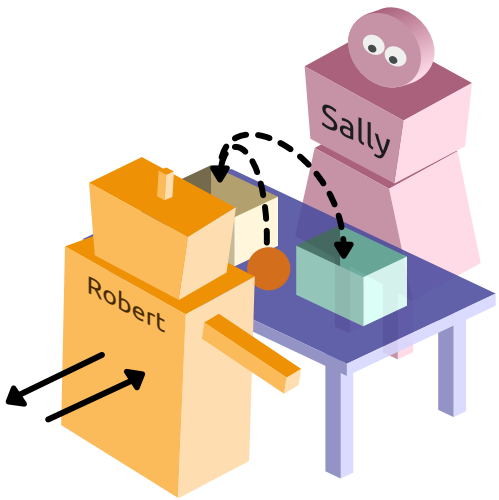
...but not many cognitive architectures model social interactions! (on BICA website, about 0 actually!)

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

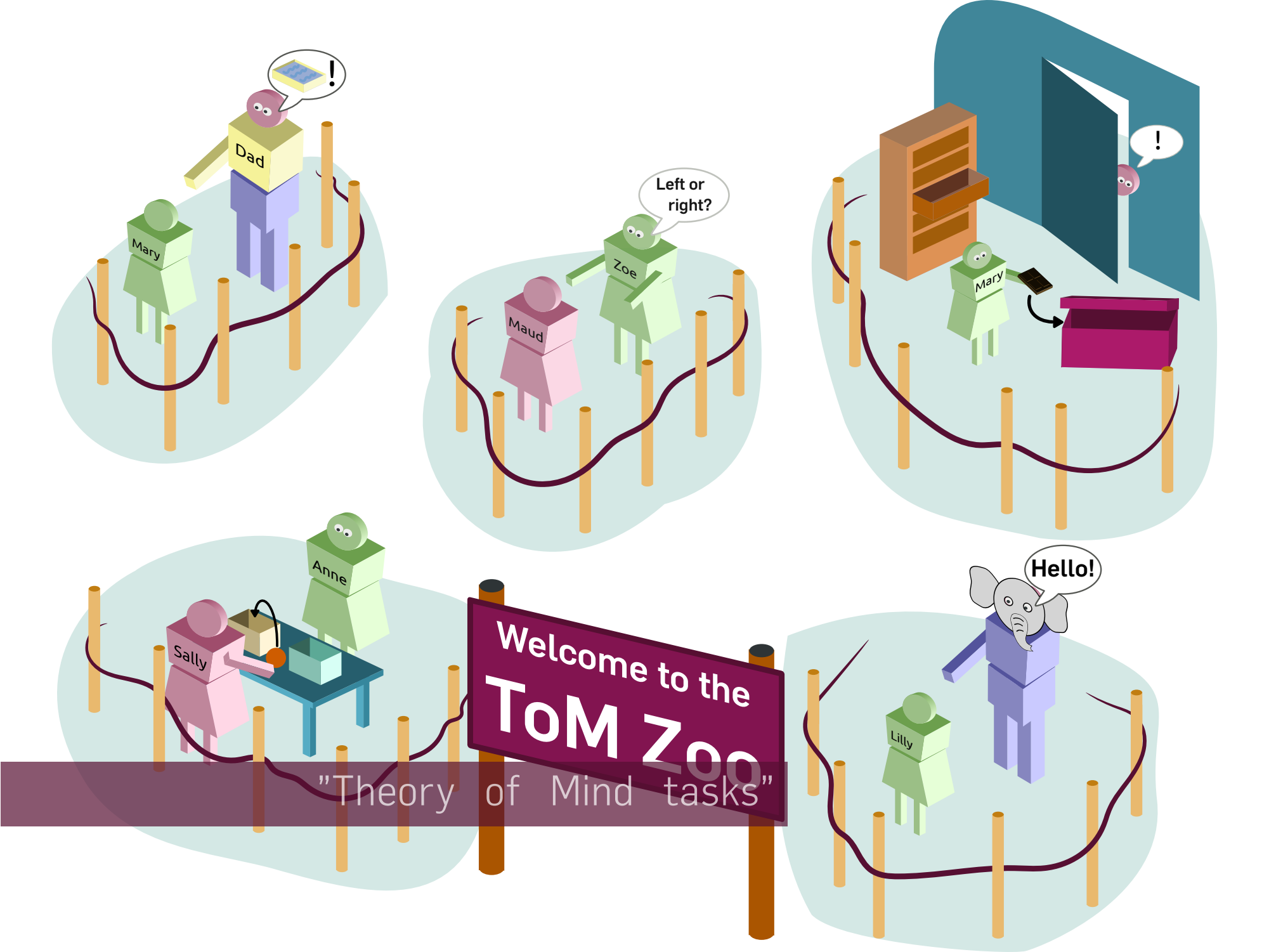
- gaze direction
- facial expression
- body language
- prior knowledge of person
- location of salient objects

Probably not the end of it, though!

THEORIES ON THEORY OF MIND

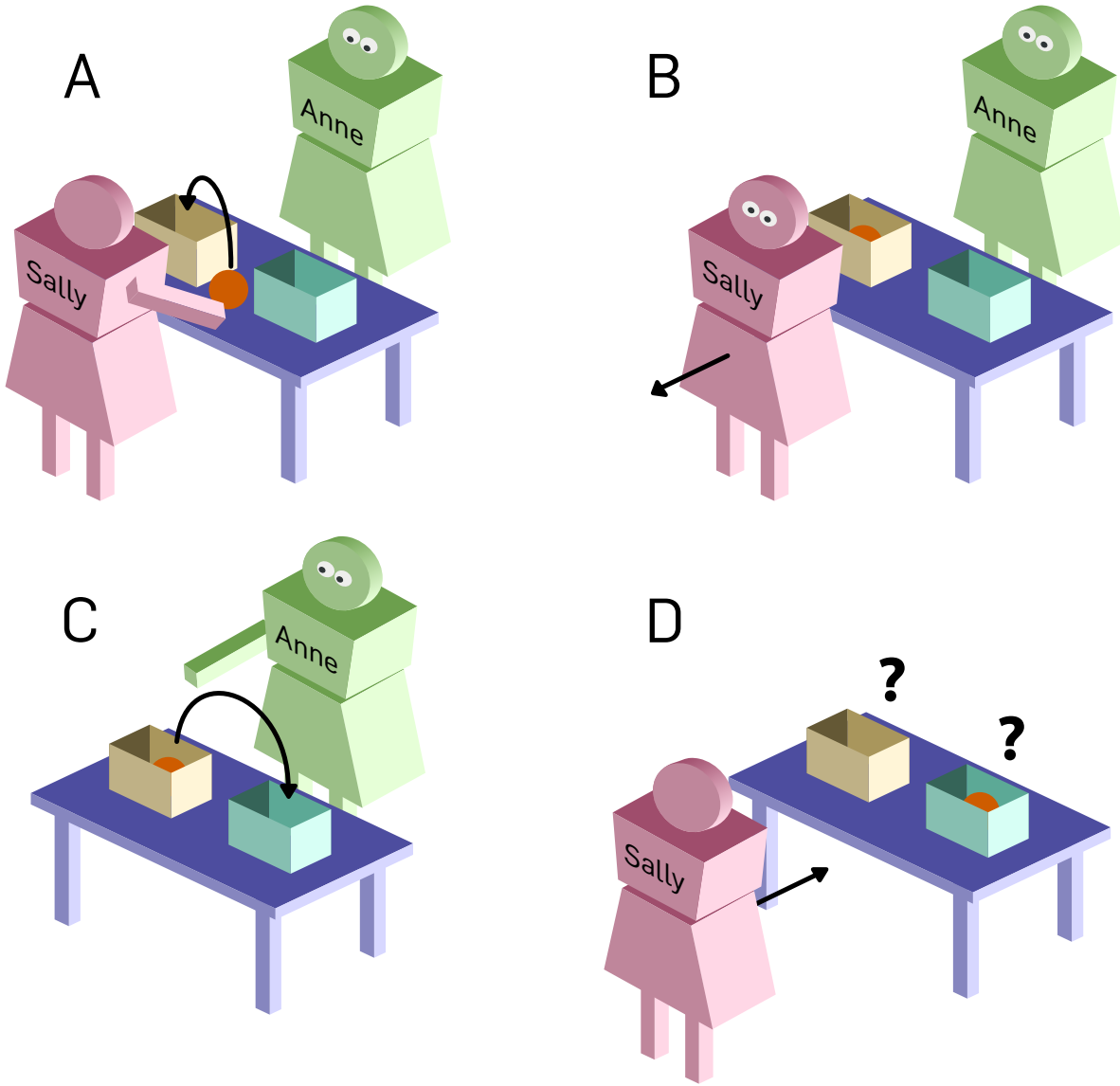


$$\begin{array}{ccc} \mathcal{M}^{\circ}(Sally, R, X) & \xleftrightarrow{\Delta_4} & \mathcal{M}^{\circ}(R, R, X) \\ \Delta_5 \updownarrow & & \updownarrow \\ \mathcal{M}^{\circ}(Sally, R, Y) & \longleftrightarrow & \mathcal{M}^{\circ}(R, R, Y) \end{array}$$



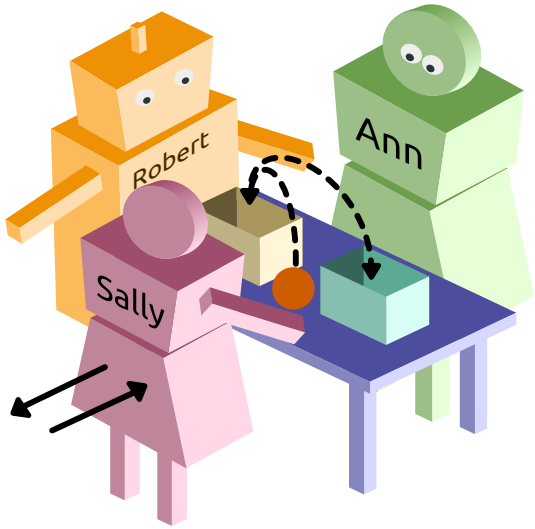
"Theory of Mind tasks"

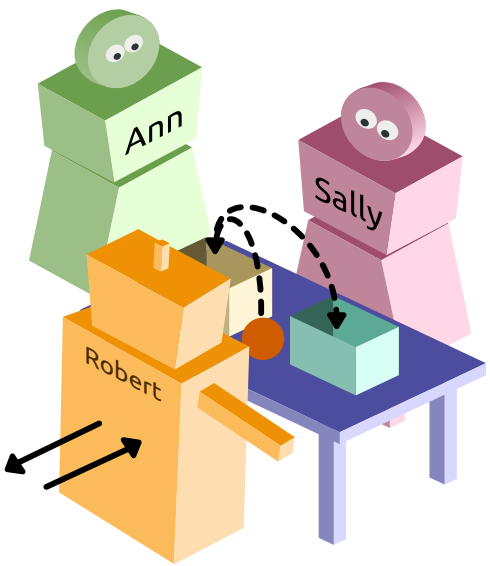
1ST ORDER TOM: THE FALSE-BELIEF EXPERIMENT

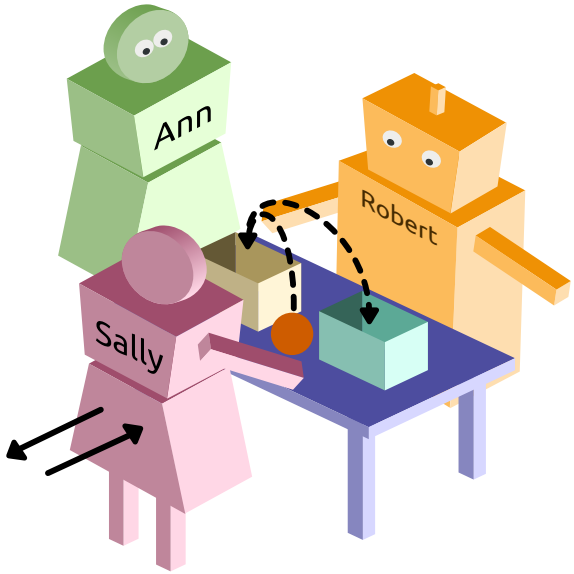


[Wimmer and Perner, **Beliefs about beliefs: Representation and constraining function [...]**, Cognition, 1983]
[Lemaignan, Dillenbourg **Mutual Modelling in Robotics: Inspirations for the Next Steps** – HRI 2015]

THE FALSE-BELIEF EXPERIMENT, RELOADED

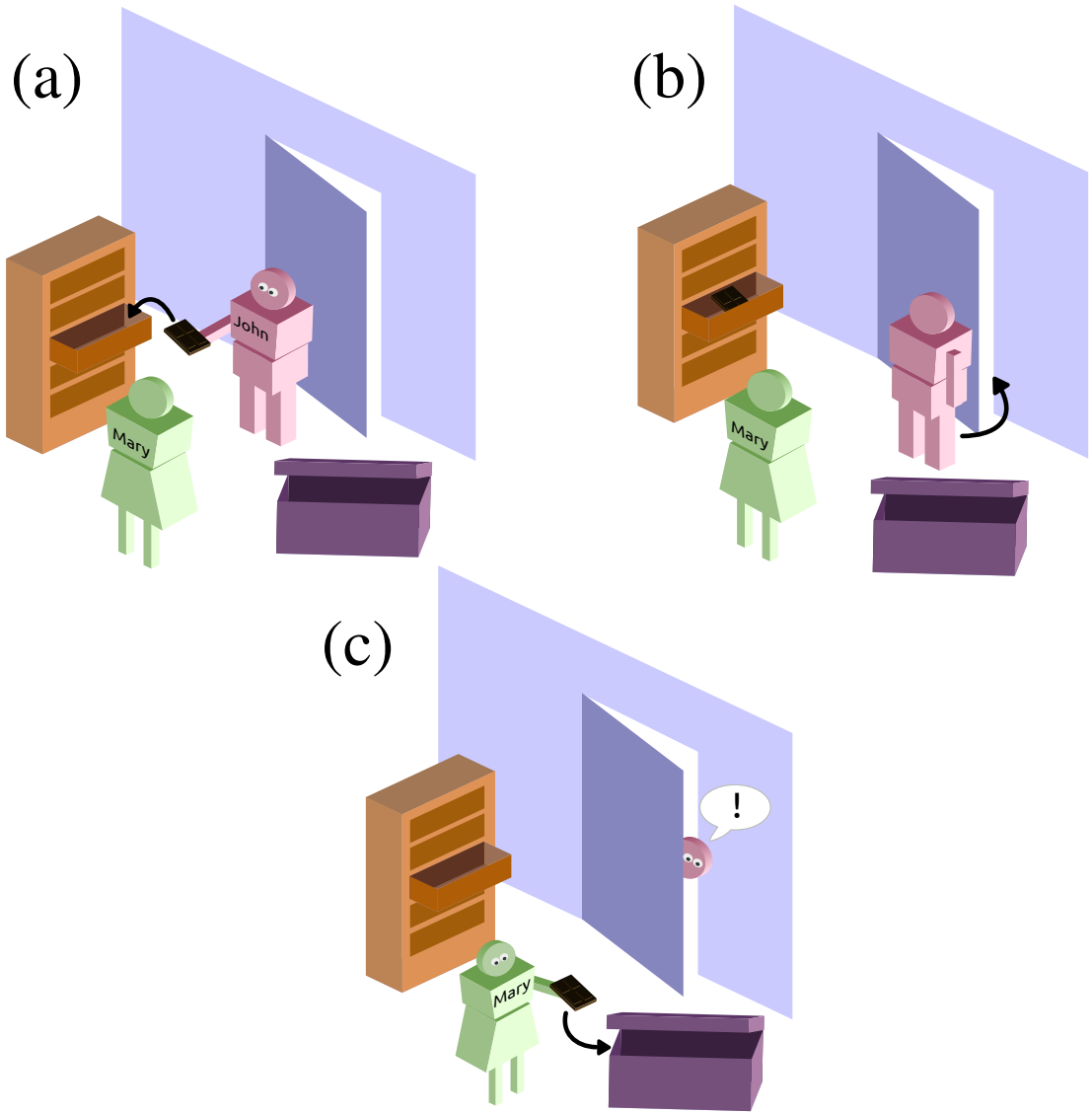






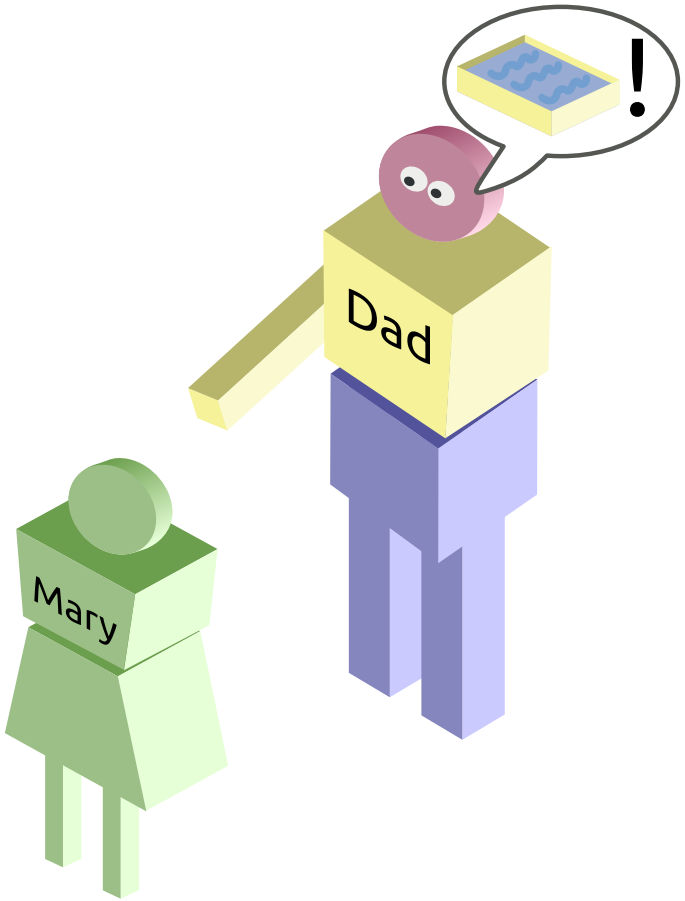
[Dillenbourg, Lemaignan et al., **The Symmetry of Partner Modelling**, Intl. J. of Computer Supported Collaborative Learning 2016]

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]

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]

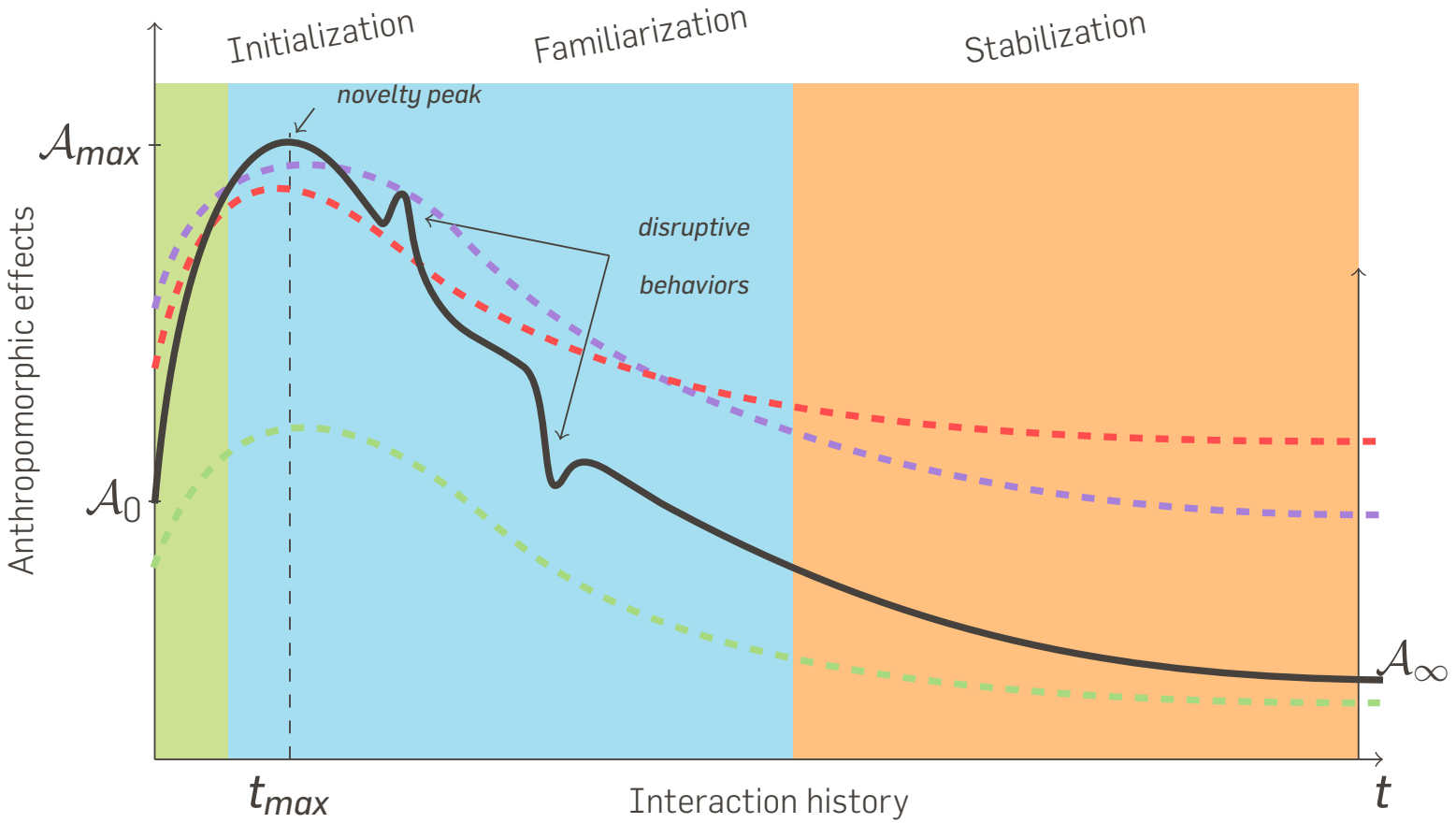
AUTISTIC ASSETS AND DEFICITS OBSERVED IN REAL LIFE

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

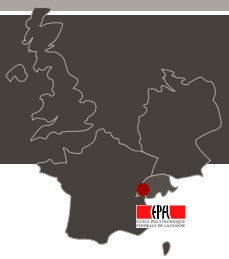
HOW DO WE PERCEIVE ROBOT OVER TIME?

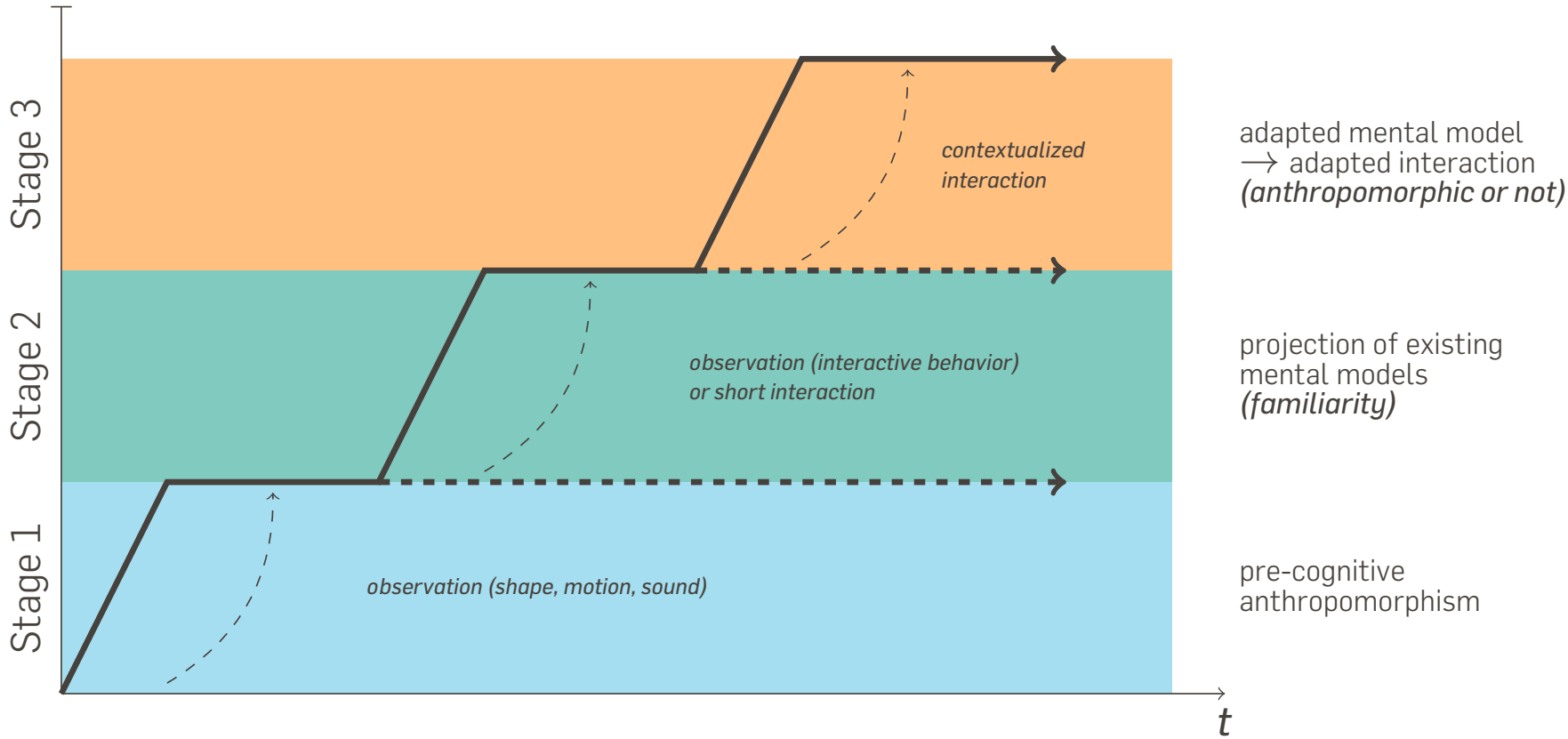


◂ Supplementary material

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

COGNITIVE INTERPRETATION?





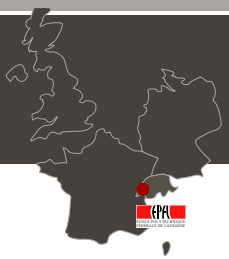
adapted mental model
⟶ adapted interaction
(anthropomorphic or not)

projection of existing mental models
(familiarity)

pre-cognitive anthropomorphism

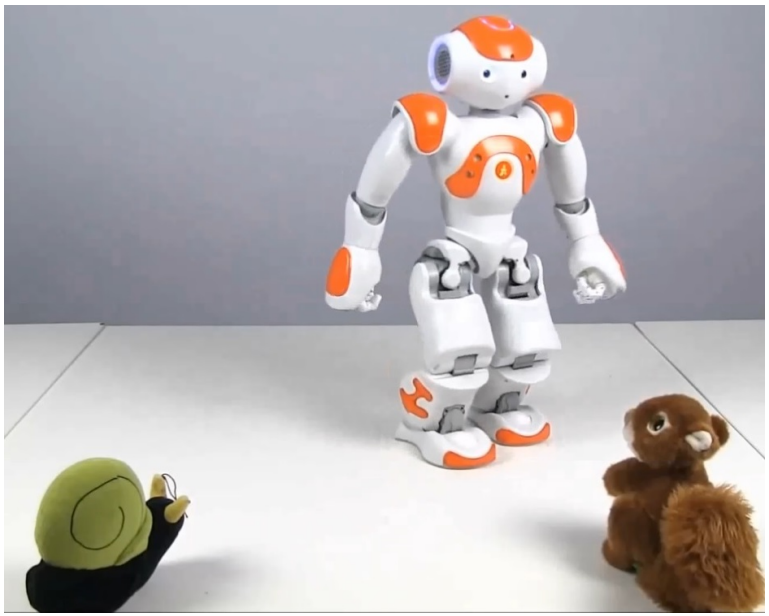
[Lemaignan, Fink, Dillenbourg, Braboszcz **The Cognitive Correlates of Anthropomorphism**, HRI 2014]

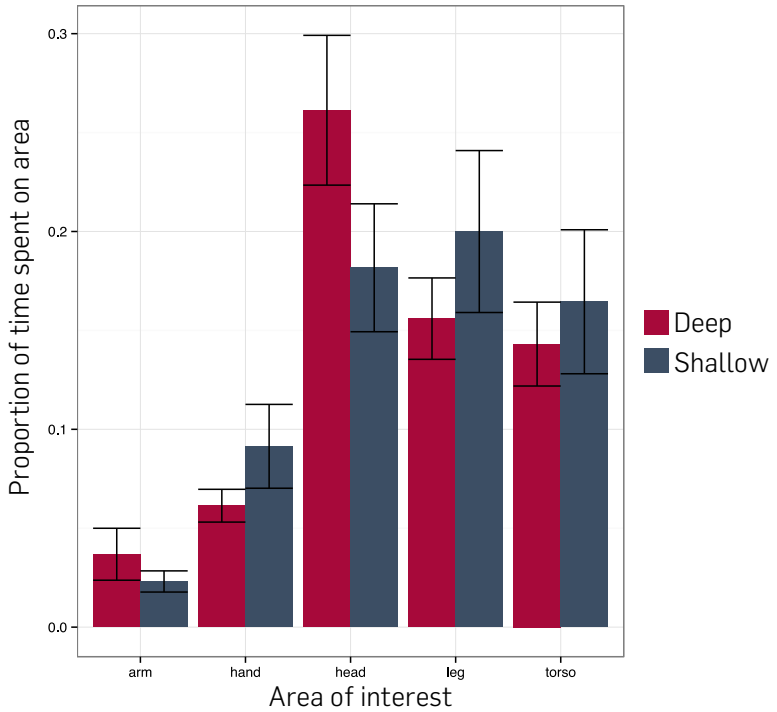
UNEXPECTED BEHAVIOURS



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

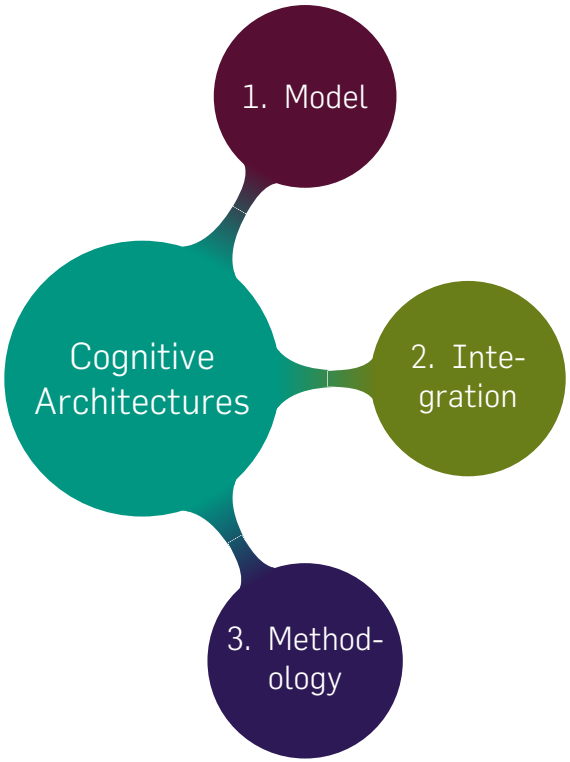
COGNITIVE CONTEXT AND ANTHROPOMORPHISM





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]

"MODELING OF HUMAN COGNITION"...