1) Human-humanoid interaction 2) Learning for damage recovery

Team LARSEN, INRIA

Serena Ivaldi

serena.ivaldi@inria.fr

Jean-Baptiste Mouret

jeanbaptiste.mouret@inria.fr

larsen

informatiques mathématiques Inria

Our lab in Nancy

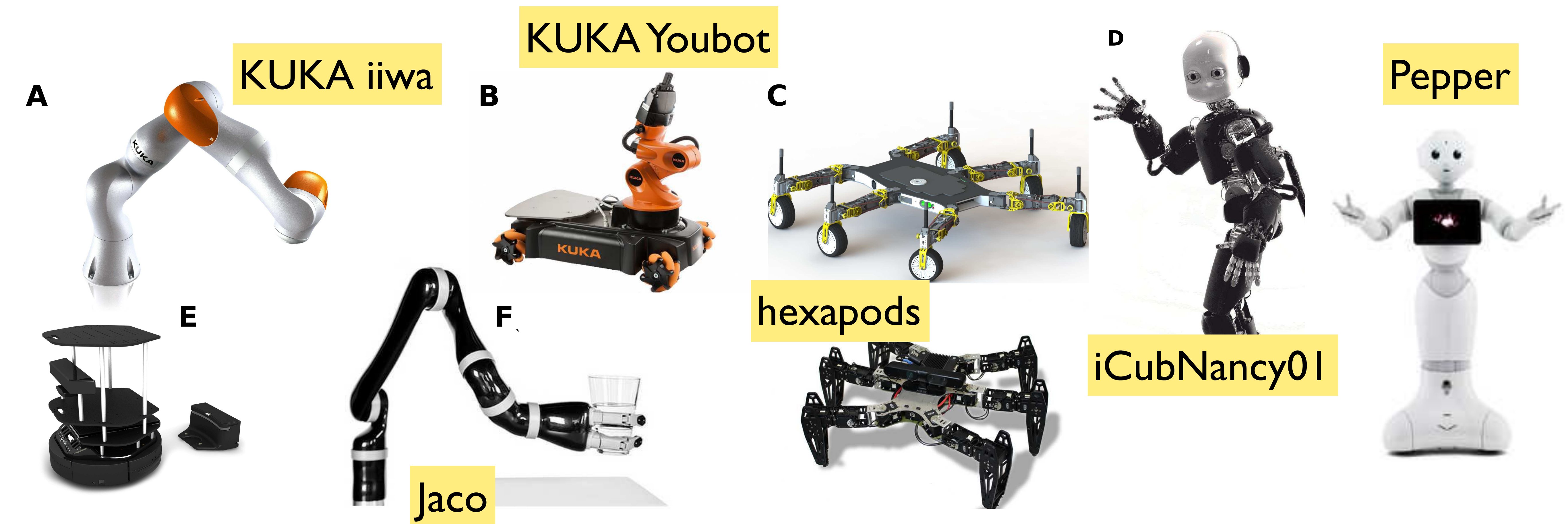
since January 2015

EU FP7	CODYCO	(2013-2017)
FR CPER	SCIARAT	(2015-2020)
EU ERC	Resibots	(2015-2020)
EU H2020	ANDY	(2017-2020)

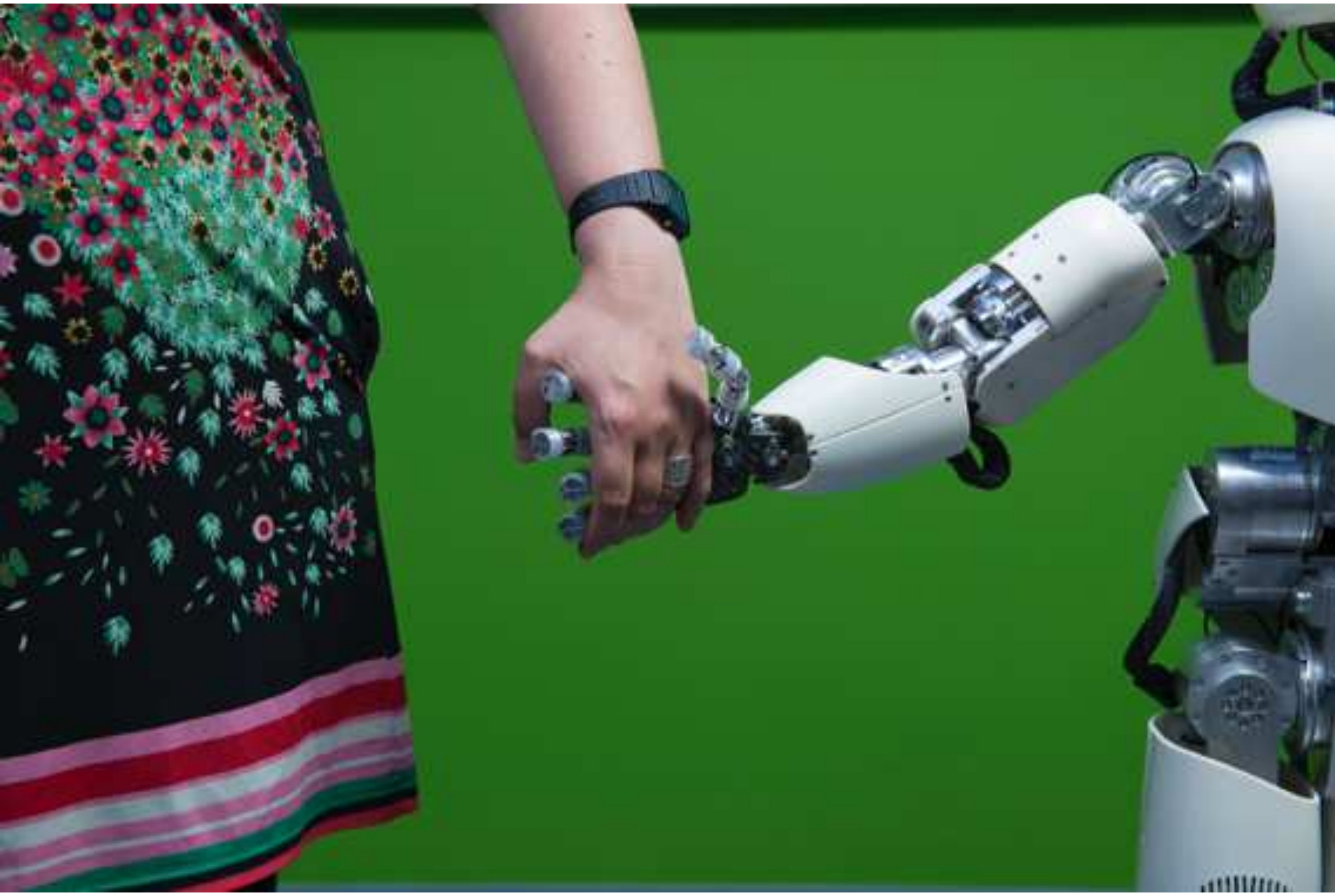
European Commission

Horizon 2020
European Union funding
for Research & Innovation

Come visit us!



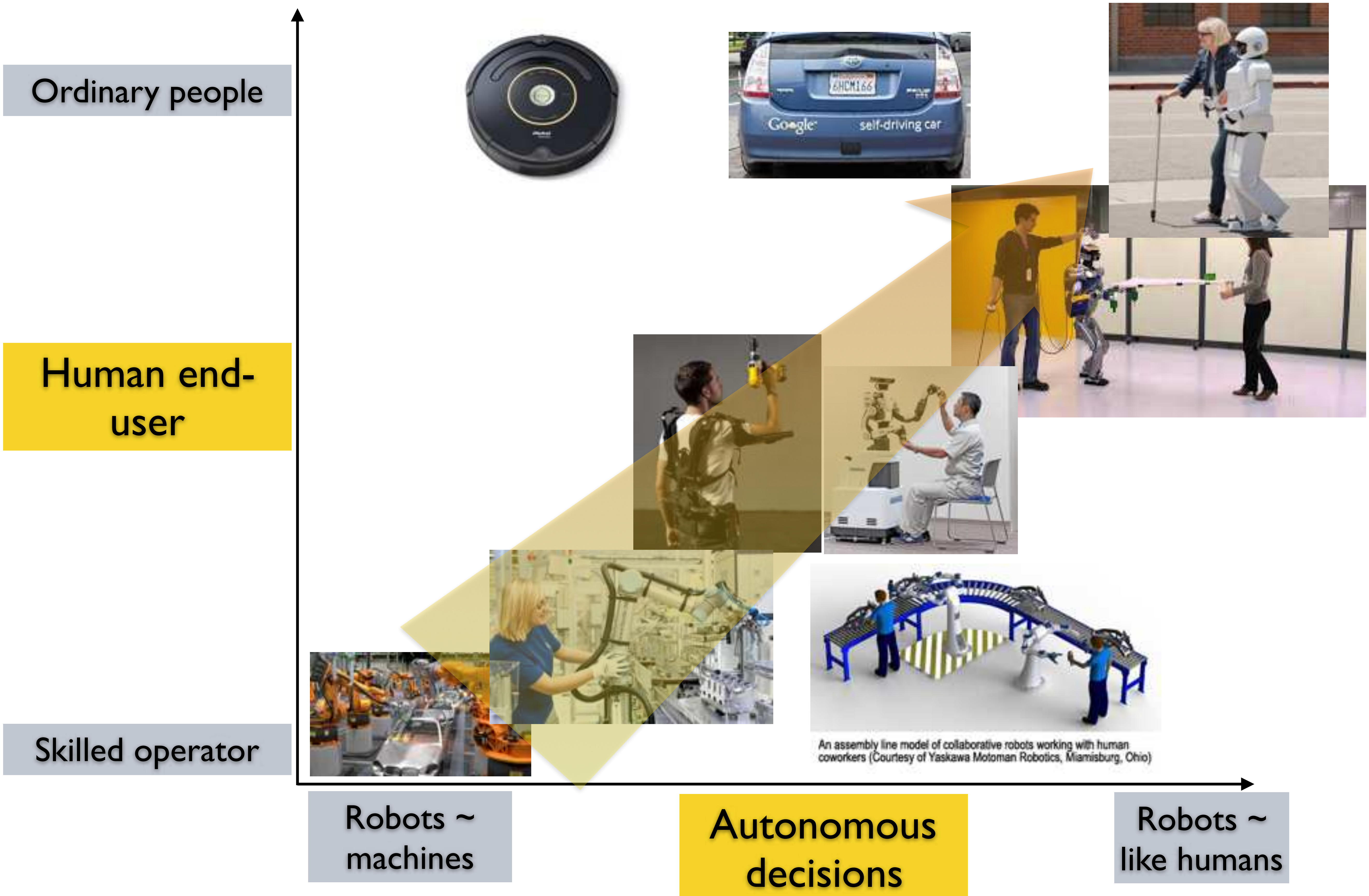
Part I: human-humanoid interaction



Serena Ivaldi

serena.ivaldi@inria.fr

Why HRI: more and more collaboration



Why HHI: more and more humanoids

More human-like shape
More sensors
More complexity
More tasks
More versatility
More interaction with non-experts
…

Anthropomorphism

Robots ~ like humans

N tasks
high expectations

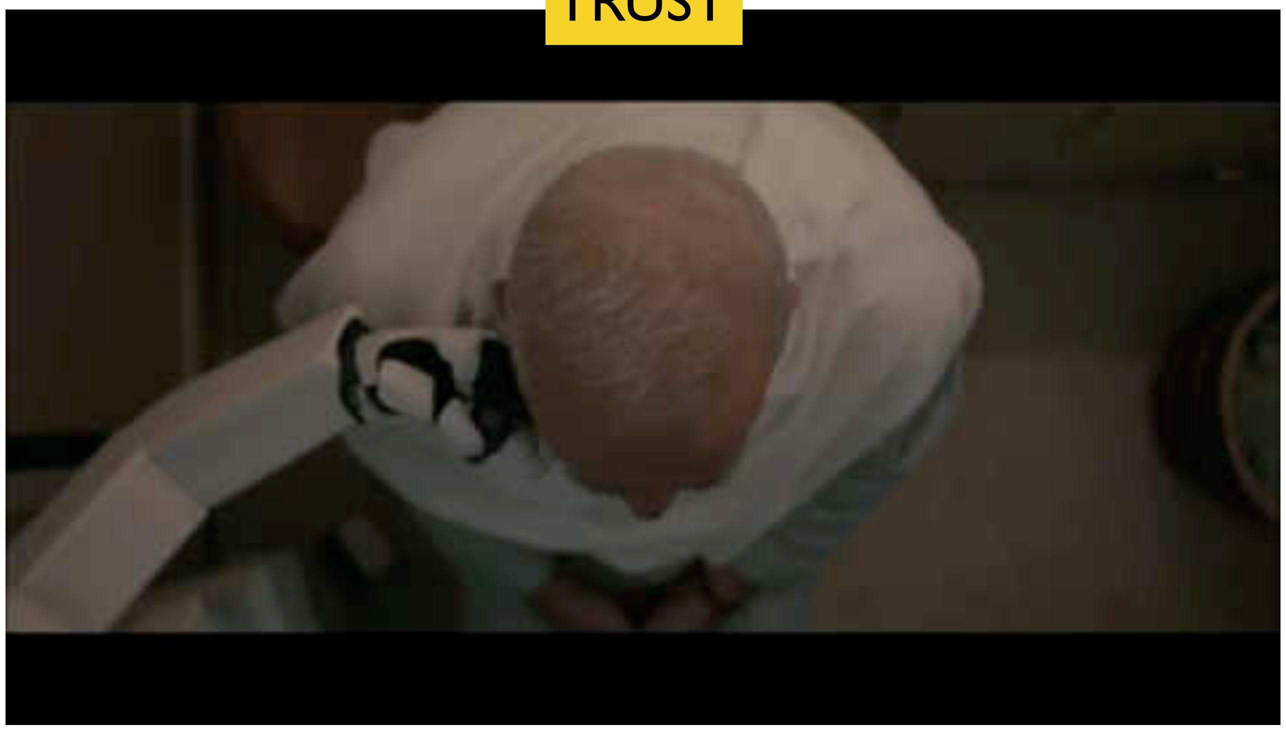
Robots ~ machines

1 task
low expectations

Problems (some)

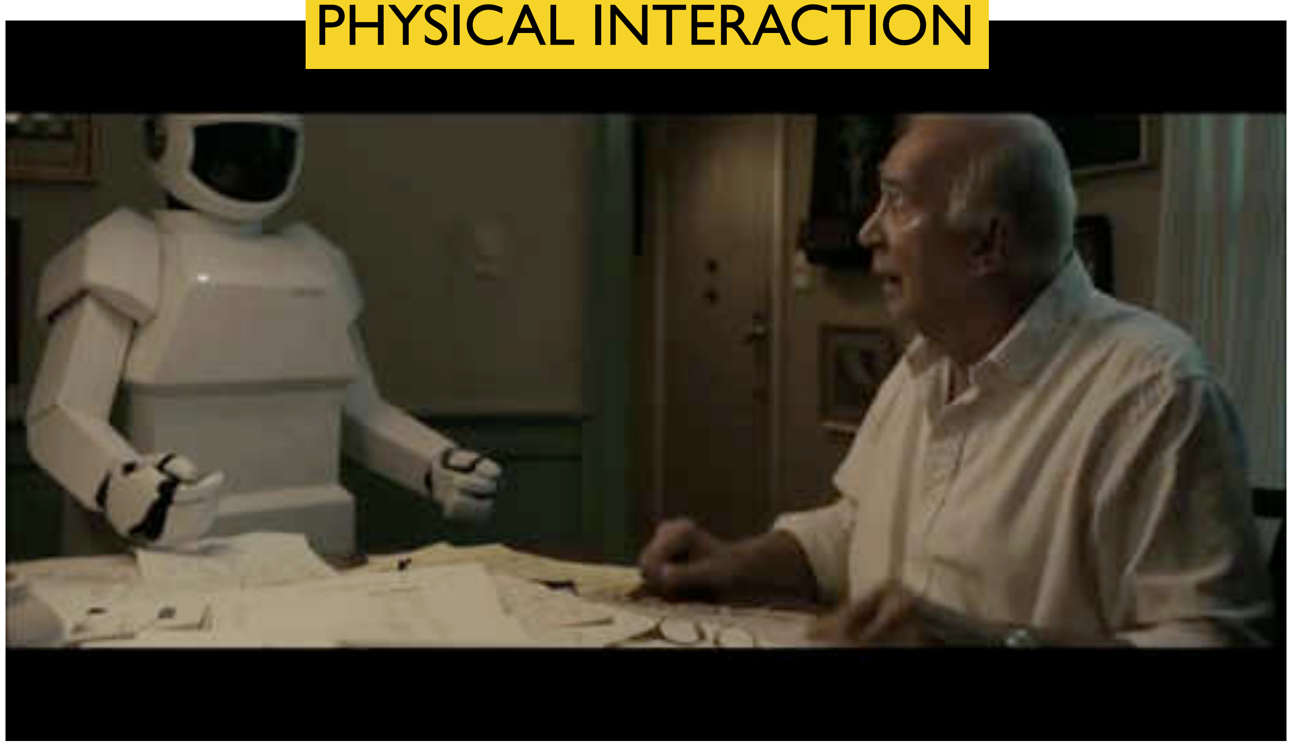
ACCEPTANCE

TRUST



SOCIAL INTERACTION

PHYSICAL INTERACTION



From the movie “Robot and Frank” (2012)

Problems (some)

ACCEPTANCE

UTAUT model, Venkatesh et al (2003)

TRUST

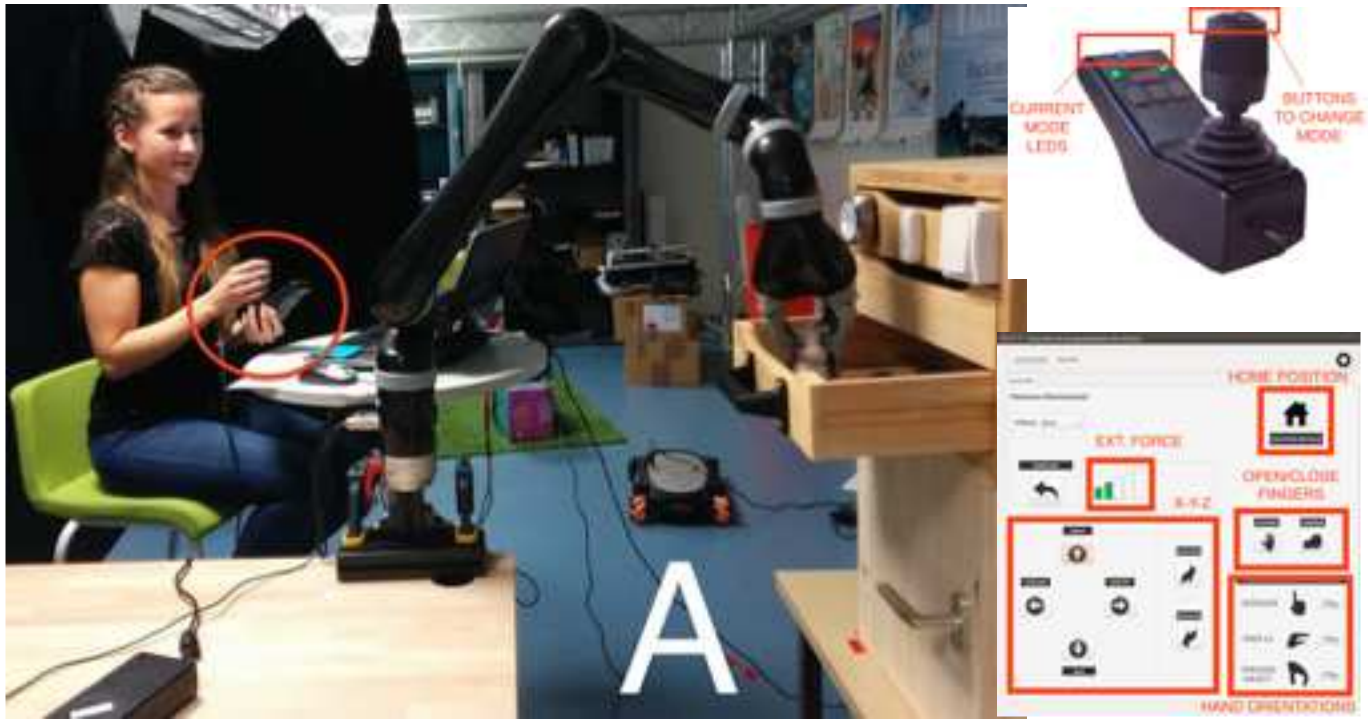
Trust in automation model, Schaefer et al. (2016)

- These questions may sound atypical (~psychology) or far from AI & robotics
  => wrong! The **humans are the final end-users of our AI technology**

- Classical models of technology acceptance and trust not adequate for the robotics case
  => lack of quantitative data supporting models
  => **need to do experiments**

Experiments (some)

ACCEPTANCE



Marichal et al (2016), Malaisé et al (2016) Int. Conf. Soc. Robotics

- The control interface is part of the robot
- Must be easy to use by non-experts
- Performance in using an interface is not the primary criteria for adoption
- Expected improvement, learning and playfulness play a key role.

TRUST

Gaudiello et al (2016) Computers in Human Behavior

- General distrust towards robots.
- People trust more the robot for its functional savvy than its social savvy.
- Very frequently, people disagree with the robot even if they think it's right.

HRI methodology

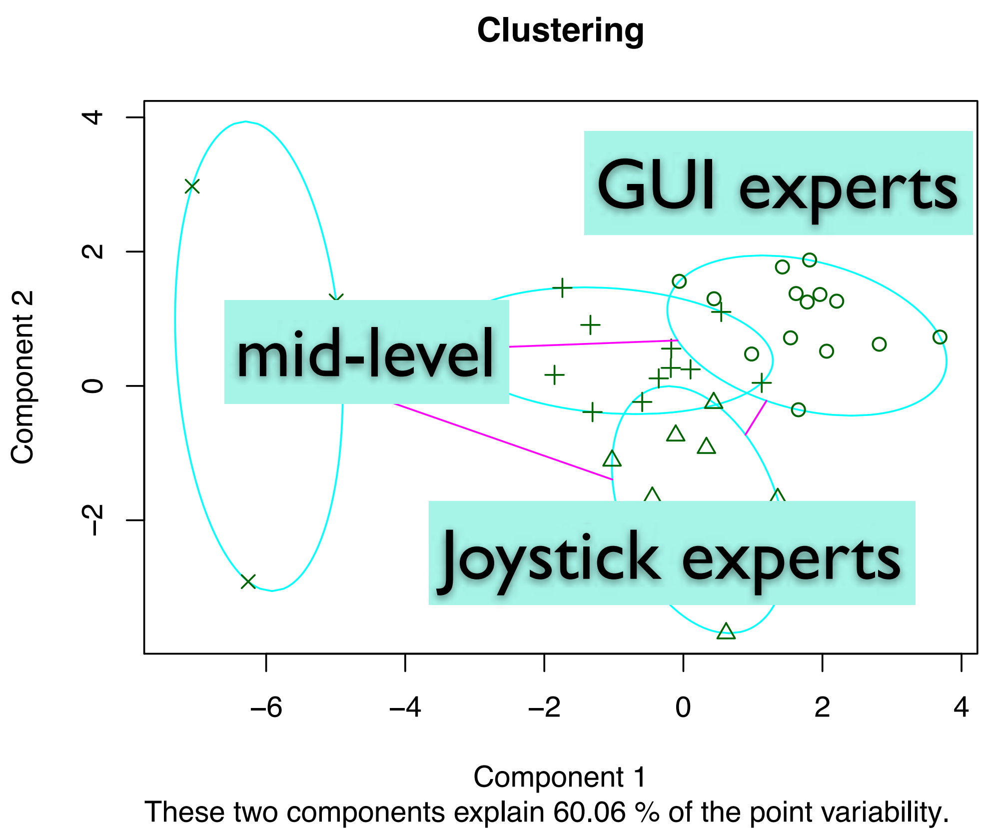
Subjective / Objective Qualitative / Quantitative Measures,

Statistics => Many subjects => Many hours with the robot

Measure	Joystick	GUI	Wilcoxon
Objective			
Duration Task 1 (s)	130.00	110.50	**V=567.00 p<0.05**
Duration Task 2 (s)	129.00	61.50	**V=741.00 p<0.001**
Duration Task 3 (s)	102.50	63.50	**V=756.50 p<0.001**
Duration Task 4 (s)	146.50	72.50	**V=707.50 p<0.001**
Duration all tasks (s)	579.50	331.50	**V=790.00 p<0.001**
n. precision errors	3.00	1.00	**V=443.00 p<0.001**
n. mapping errors	11.50	1.00	**V=820.00 p<0.001**
n. pauses	69.50	71.50	V=176.00 p=0.73
Median Pause Duration (s)	2.80	1.72	**V=10.00 p<0.001**
Max Pause Duration (s)	12.51	7.86	V=119.00 p=0.16
% Inactivity	42.56	43.76	V=182.00 p=0.88
Subjective			
Perceived Ease of Use	12.00	25.00	**V=3.00 p<0.001**
User Satisfaction	19.00	22.00	**V=64.50 p<0.001**
Facilitating Conditions	10.00	13.00	**V=12.00 p<0.001**

Questionnaires and semi-directed interviews designed by constructs of Acceptance models

Codes	GUI	Joystick
Ease of Use		
Learning	22	16
Information	15	3
Thinking	4	11
Errors	8	11
Usage		
After Training	0	29
Quick use	13	1
Context	9	5
Frequency	3	9
Hard Task	1	5
Public	3	1
Control		
Possibilities	12	14
Accuracy	10	12
Pre-defined Move	7	0
Smoothness	2	5
Speed Control	3	1
Feeling		
Comfort	11	7
Enjoyment	2	6
Satisfaction	4	5
Worried	1	4
Ergonomic		
Intuition	4	15
Installation	6	8
In Hand	0	9
Attention	5	5
Individual Characteristics		
Habits	9	11
Age	4	2
Impairment	0	3



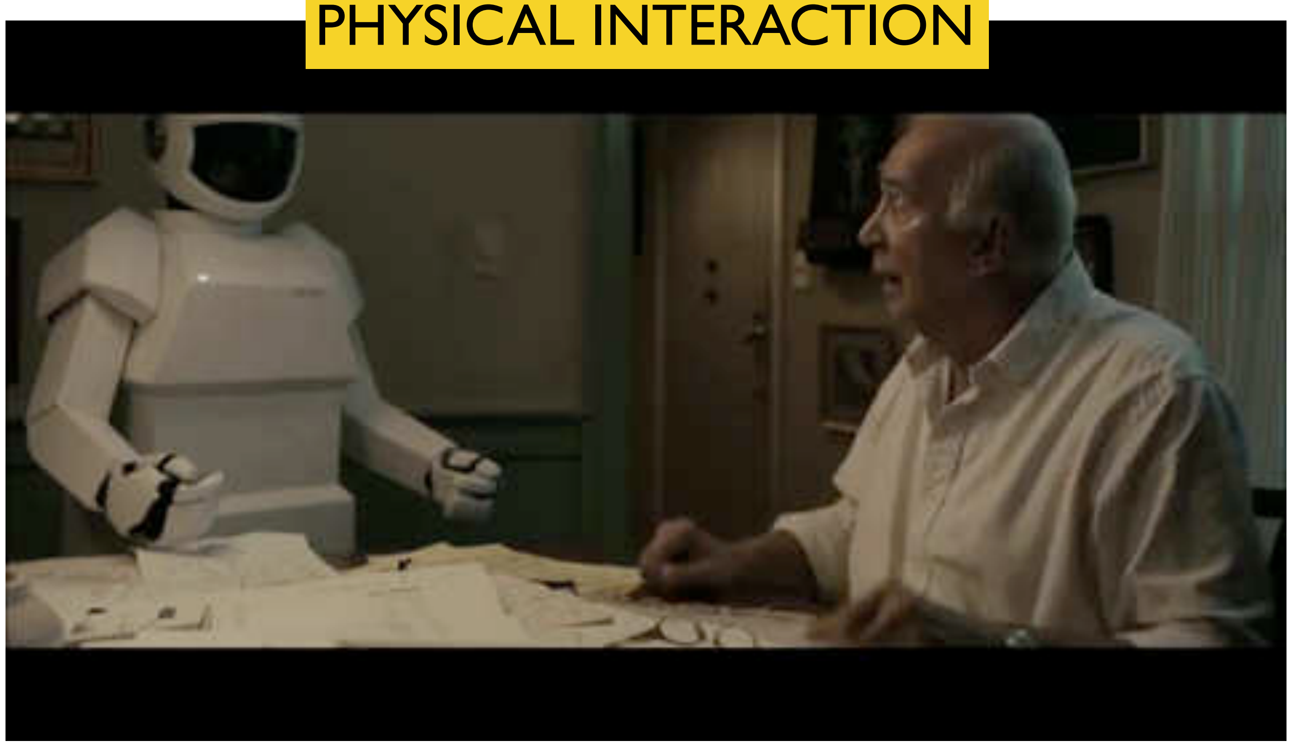
Automatic clustering of performance metrics and individual factors to identify stereotypical group behaviours in interacting with the robot.

Problems (some)

COLLABORATION

SOCIAL INTERACTION

PHYSICAL INTERACTION



From the movie "Robot and Frank" (2012)

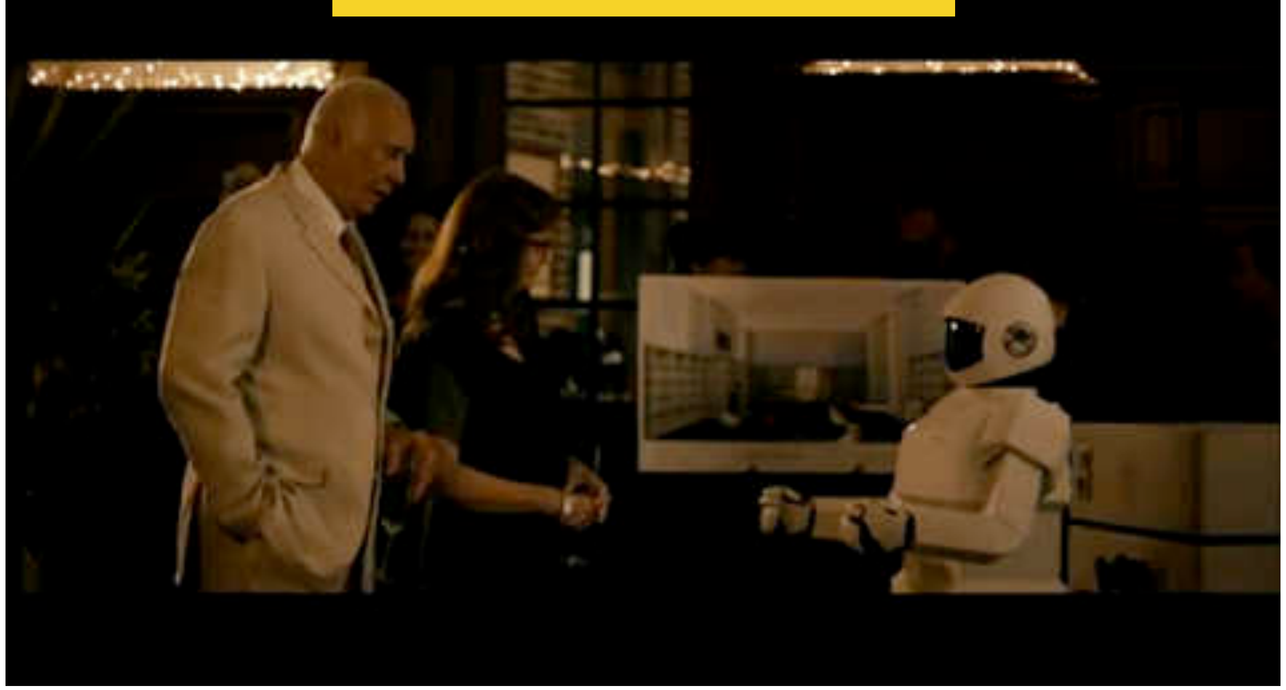
Problems (some)

- Interaction = a problem with uncertainty:
  - **robots do not always have buttons**
  - what can they do? when? what is their goal/task?

- People behave differently => **personality, individual factors**
- Haptic information alone is not sufficient to discriminate intent of motion in physical human-robot collaboration (Dumora et al 2012)

  => **multimodality**

SOCIAL INTERACTION



PHYSICAL INTERACTION

From the movie "Robot and Frank" (2012)

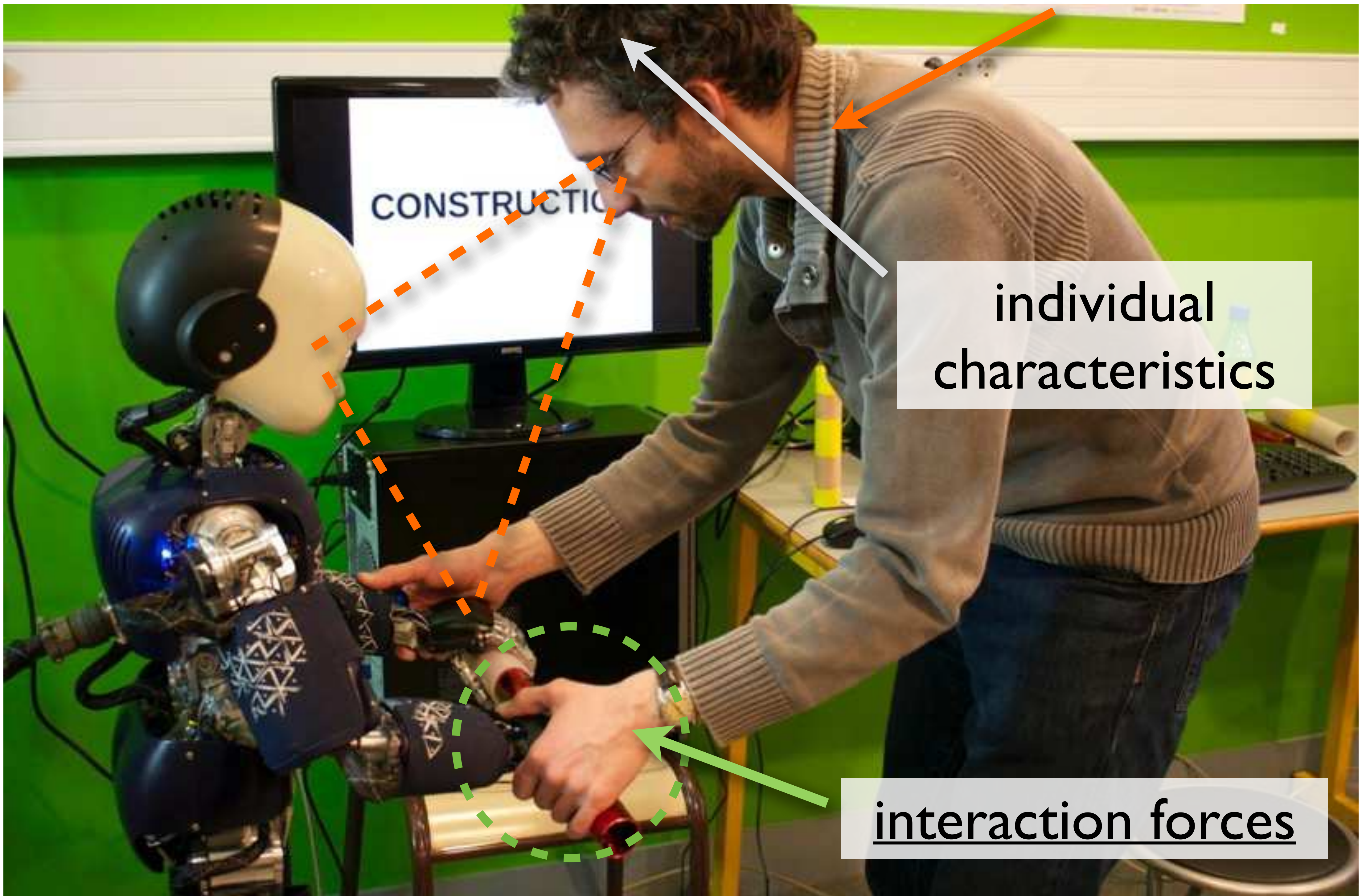
Human-human collaboration

Ordinary people teach iCub how to assembly an object

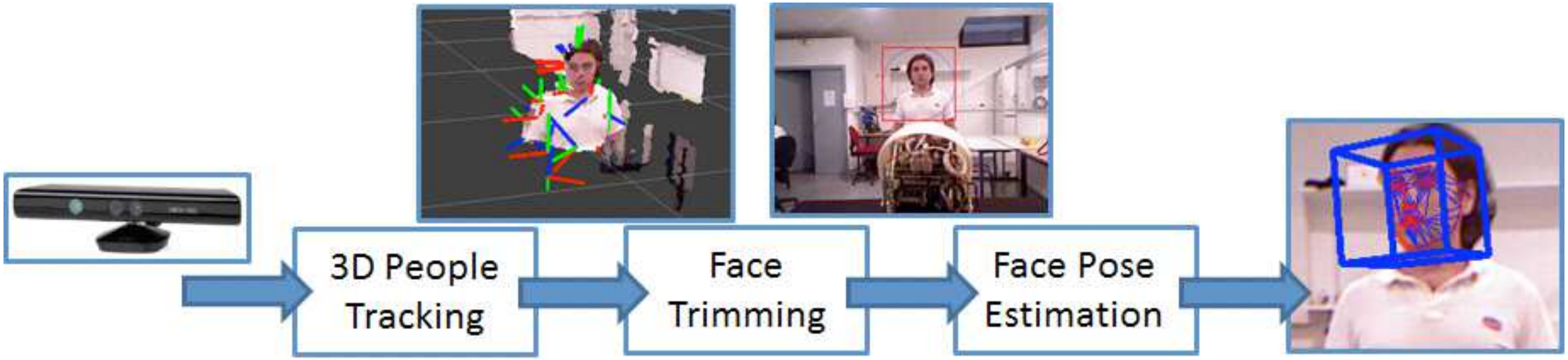
56 participants (19 M, 37 F), aged 36,95±14,32 (min 19, max 65)

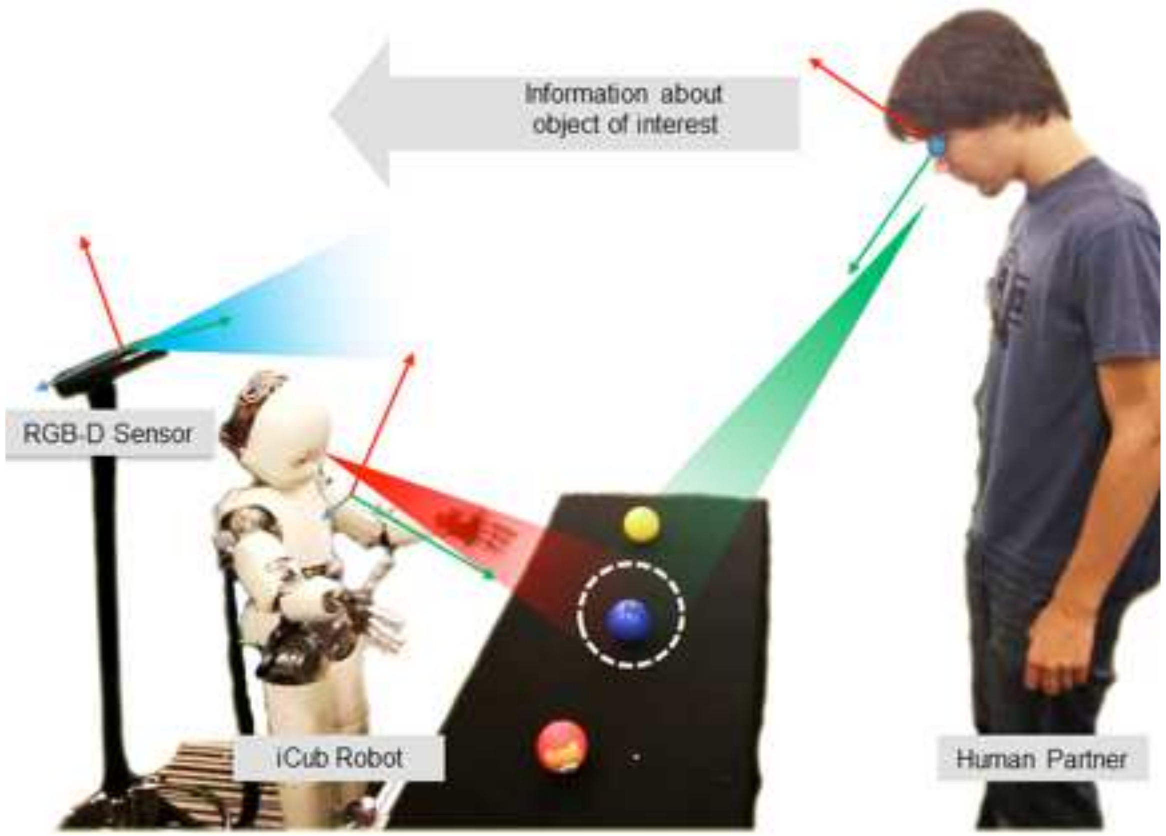
Studying human-robot collaborative assembly

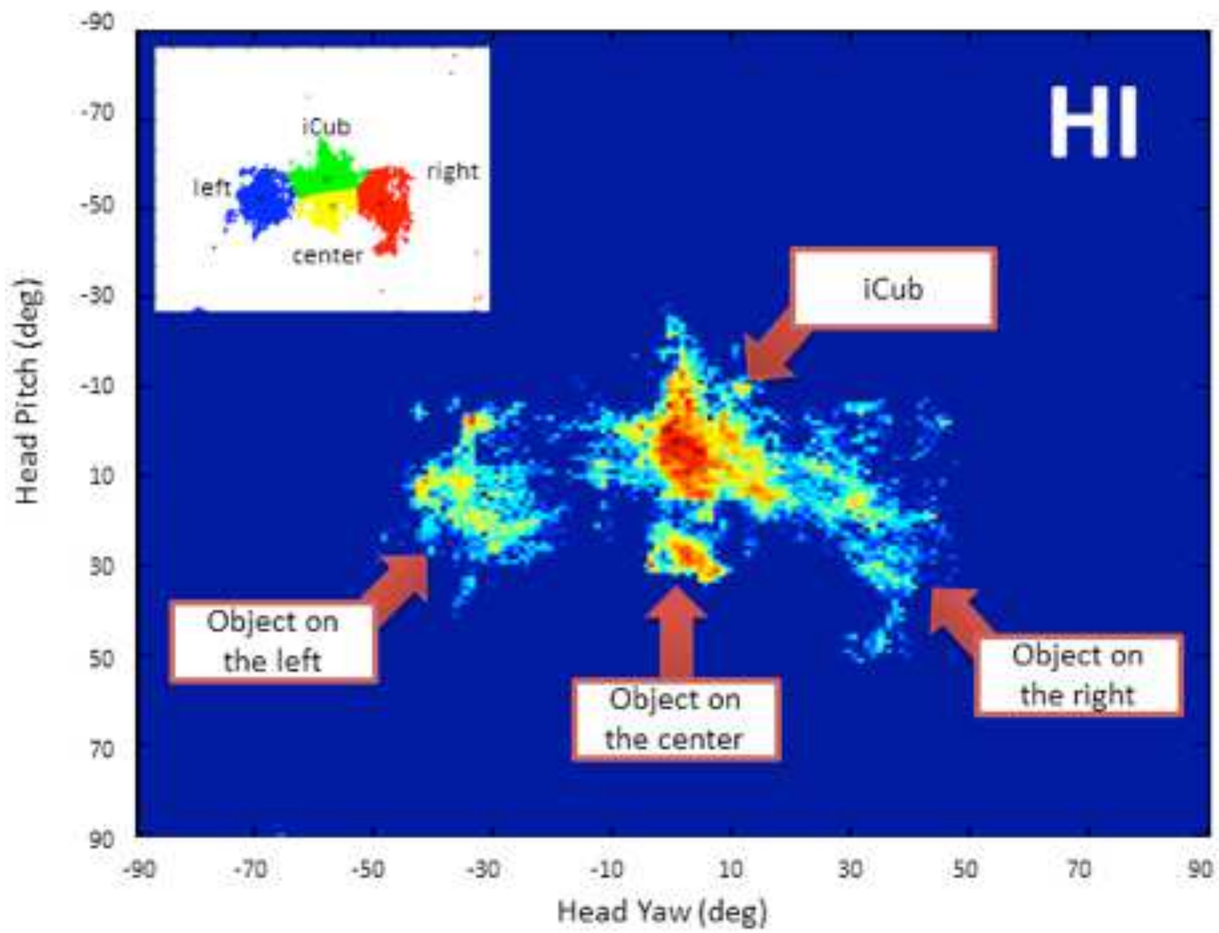
verbal/non-verbal signals



Social signals, e.g., gaze



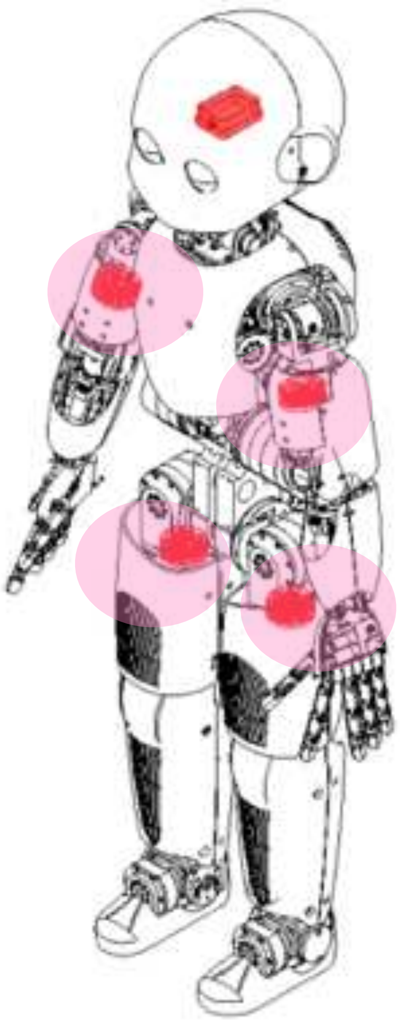




Ivaldi et al, Frontiers in Neurorobotics, 2014

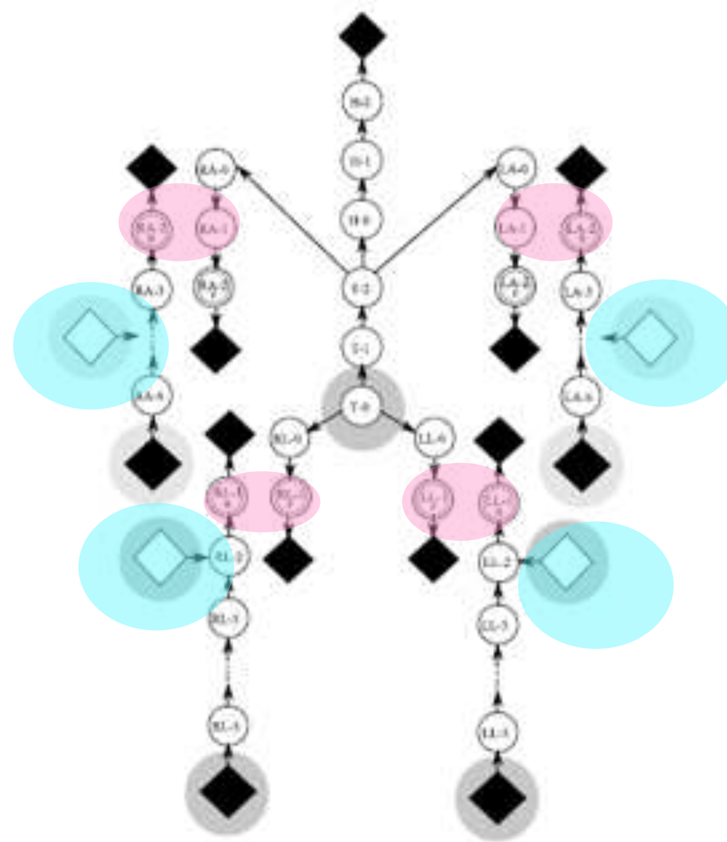
Physical signals, e.g., contact forces

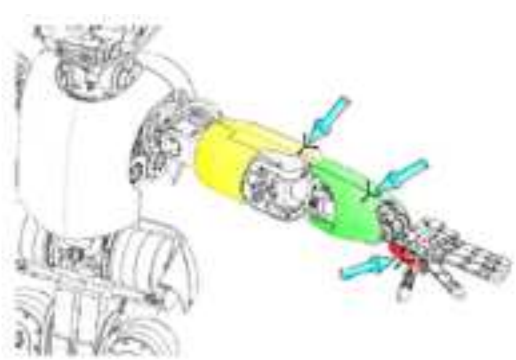
Inertial sensor

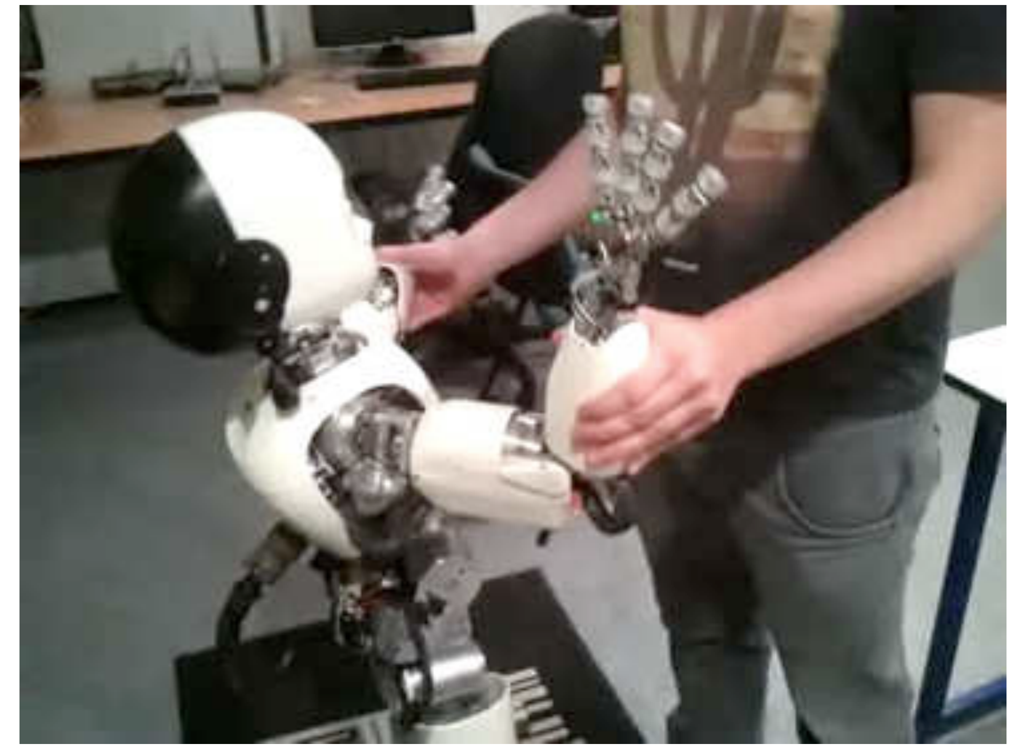


F/T sensor

skin







Ivaldi, et al.
HUMANOIDS 2011

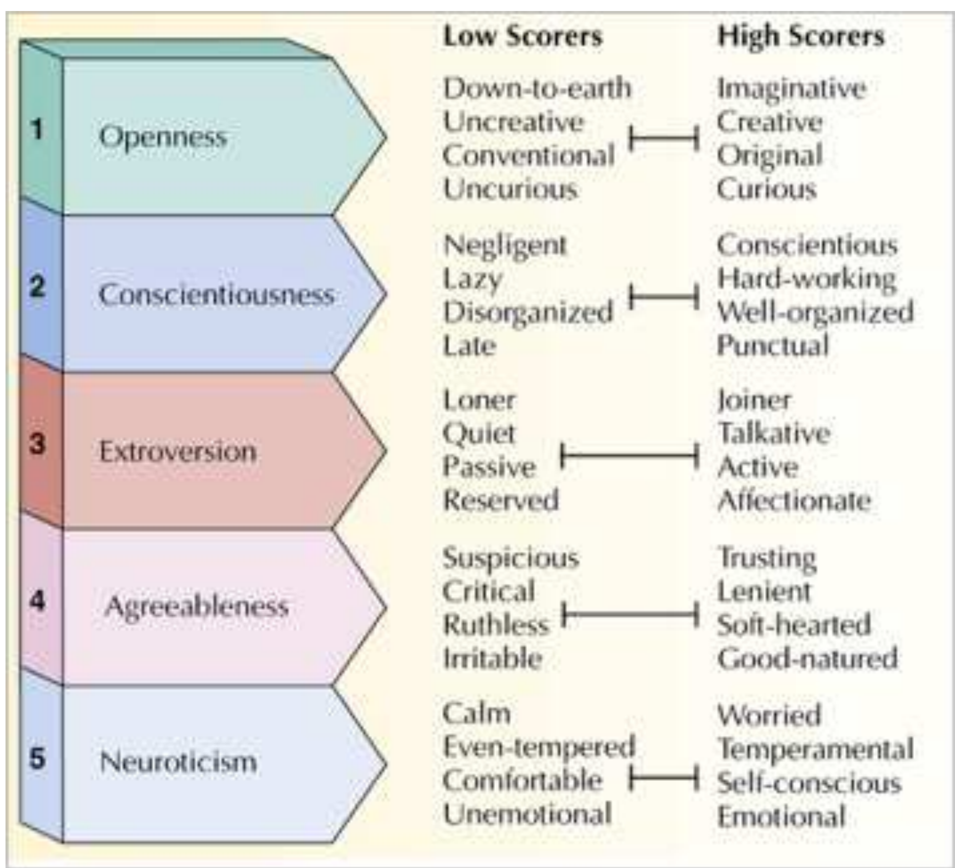
Droniou et al, RAS 2015,
Stulp et al, HUMANOIDS 2013

Individual factors, e.g., extroversion and NARS

Both attitudes and personality traits influence our actions and behaviors, together with other social, contextual and individual factors.

Personality:
behavior patterns, stable in adults

BIG 5 personality



McCrae, R. R., & Costa, P.T. (2003)

Attitudes:
behavior tendencies,
contingent, may change

Negative attitude towards robots
NARS

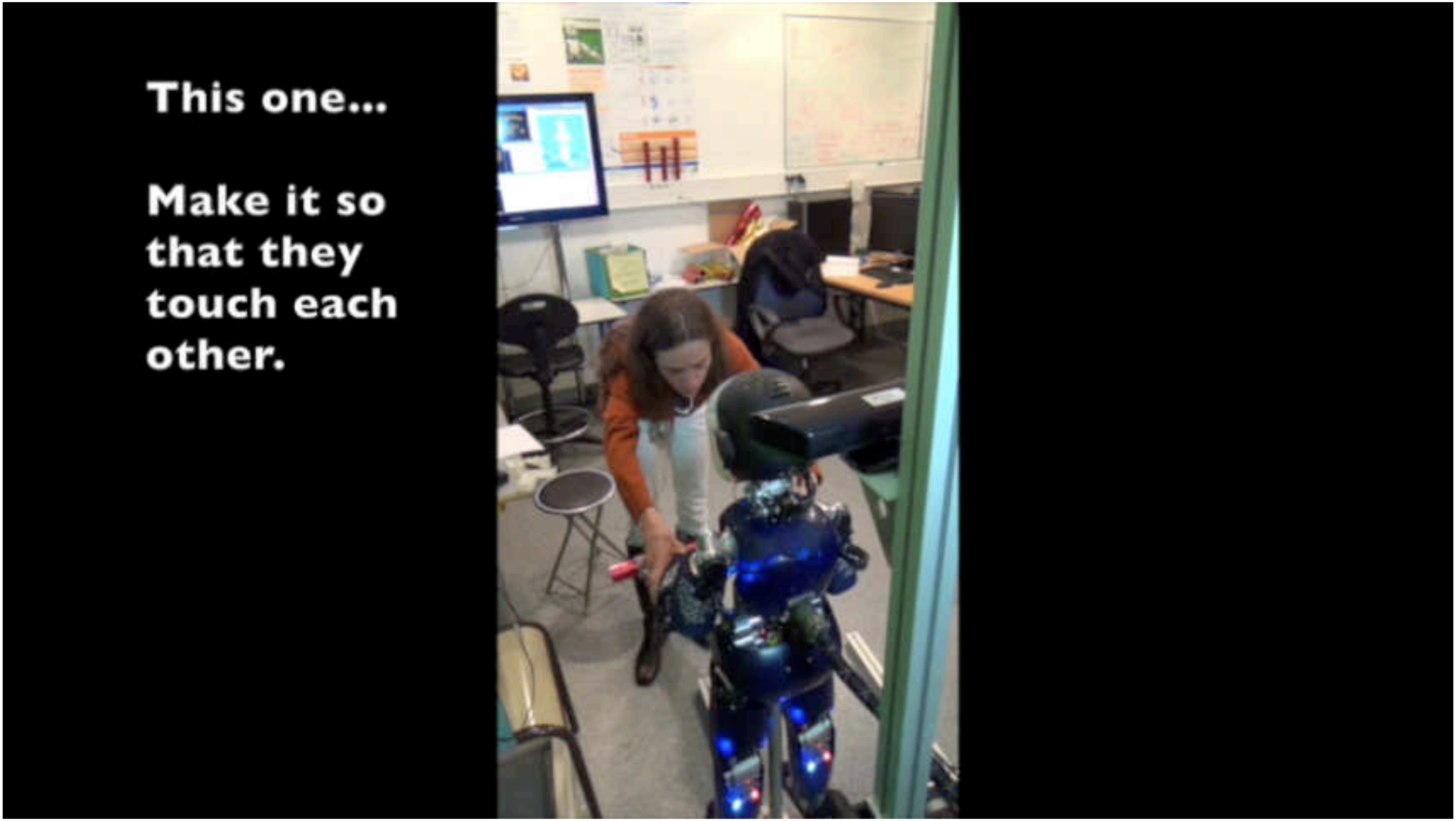
Negative attitude toward situations and interactions with robots

Negative attitude toward social influence of robots

Negative attitude toward emotions in interactions with robots

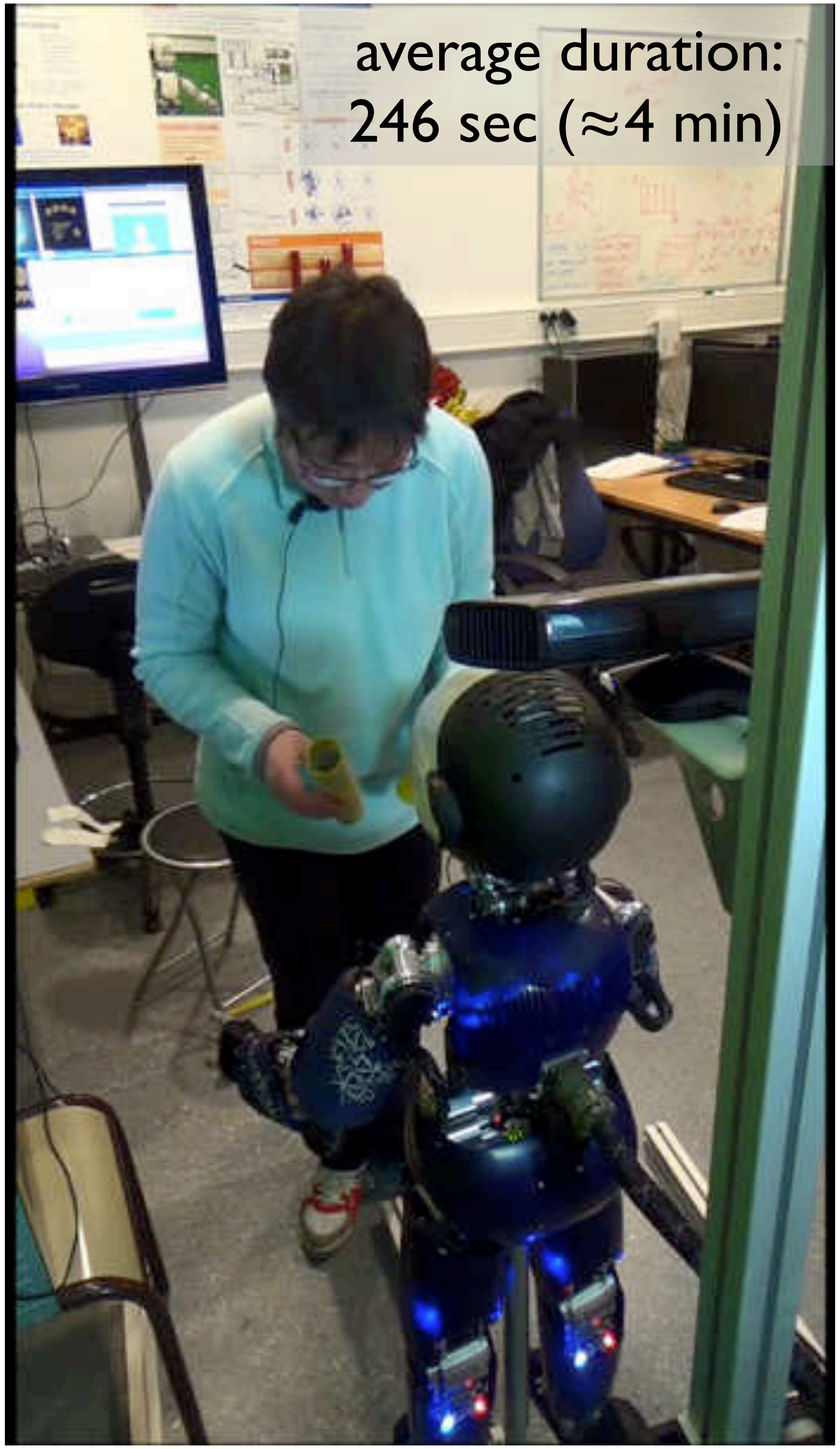
Nomura et al (2004)

Individual factors appear in the interaction



Ivaldi, S.; Lefort, S.; Peters, J.; Chetouani, M.; Provasi, J.; Zibetti, E. (2016) Towards engagement models that consider individual factors in HRI: on the relation of extroversion and negative attitude towards robots to gaze and speech during a human-robot assembly task. Int. Journal Social Robotics

Results and observations



Most relevant results:

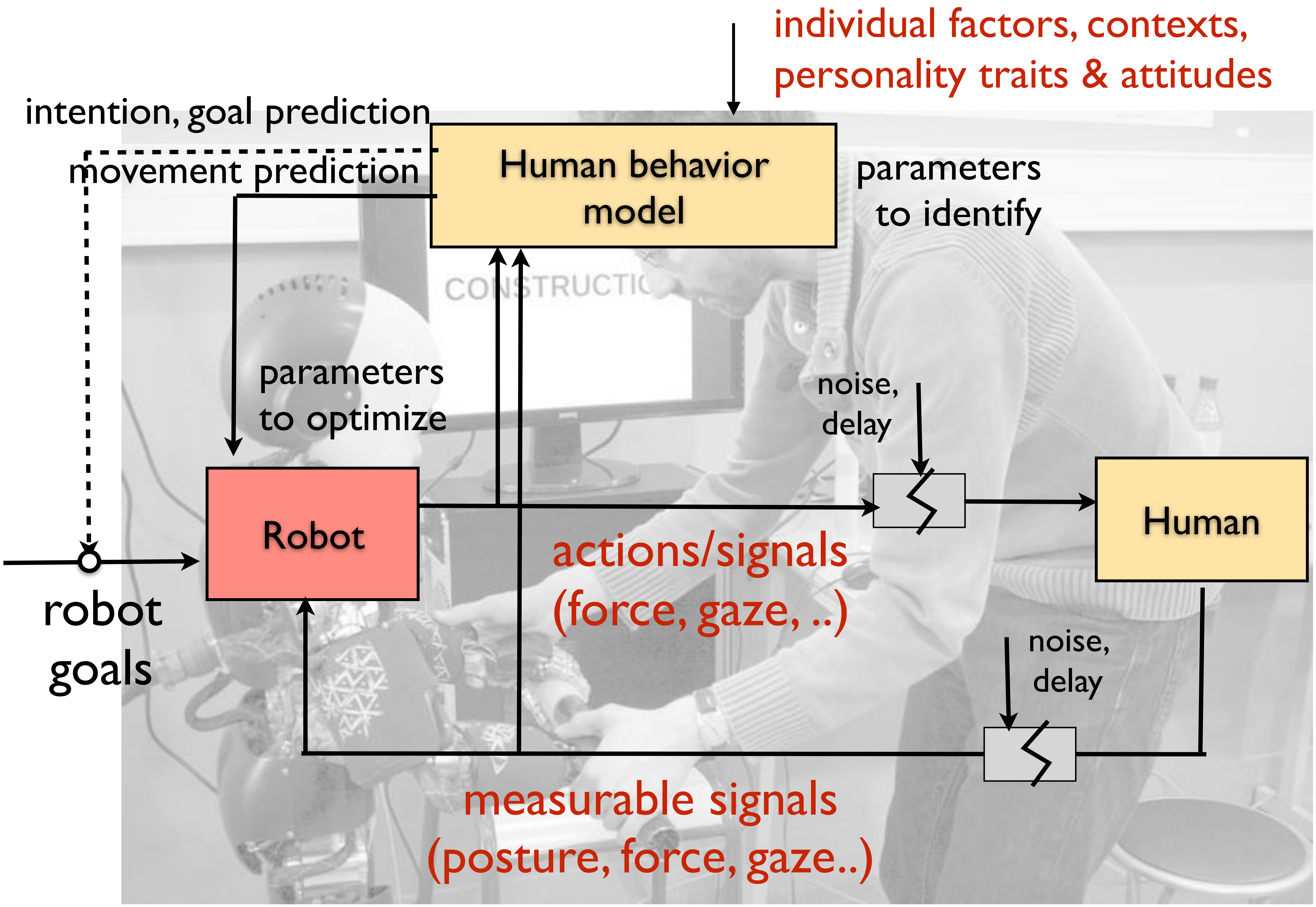
- Extroverts talk more
- **Negative attitude towards robots:**
  - avoid gazing at the robot's face
  - apply bigger forces
- Older people apply smaller forces
- **Learning effect in only 3 trials:**
  - smoothness
  - forces

Important observations:

- different strategies/behaviors
- **a lot of variability** in the recorded trajectories during haptic exchange

Ivaldi, S.; Lefort, S.; Peters, J.; Chetouani, M.; Provasi, J.; Zibetti, E. (2016) Int. Journal Social Robotics

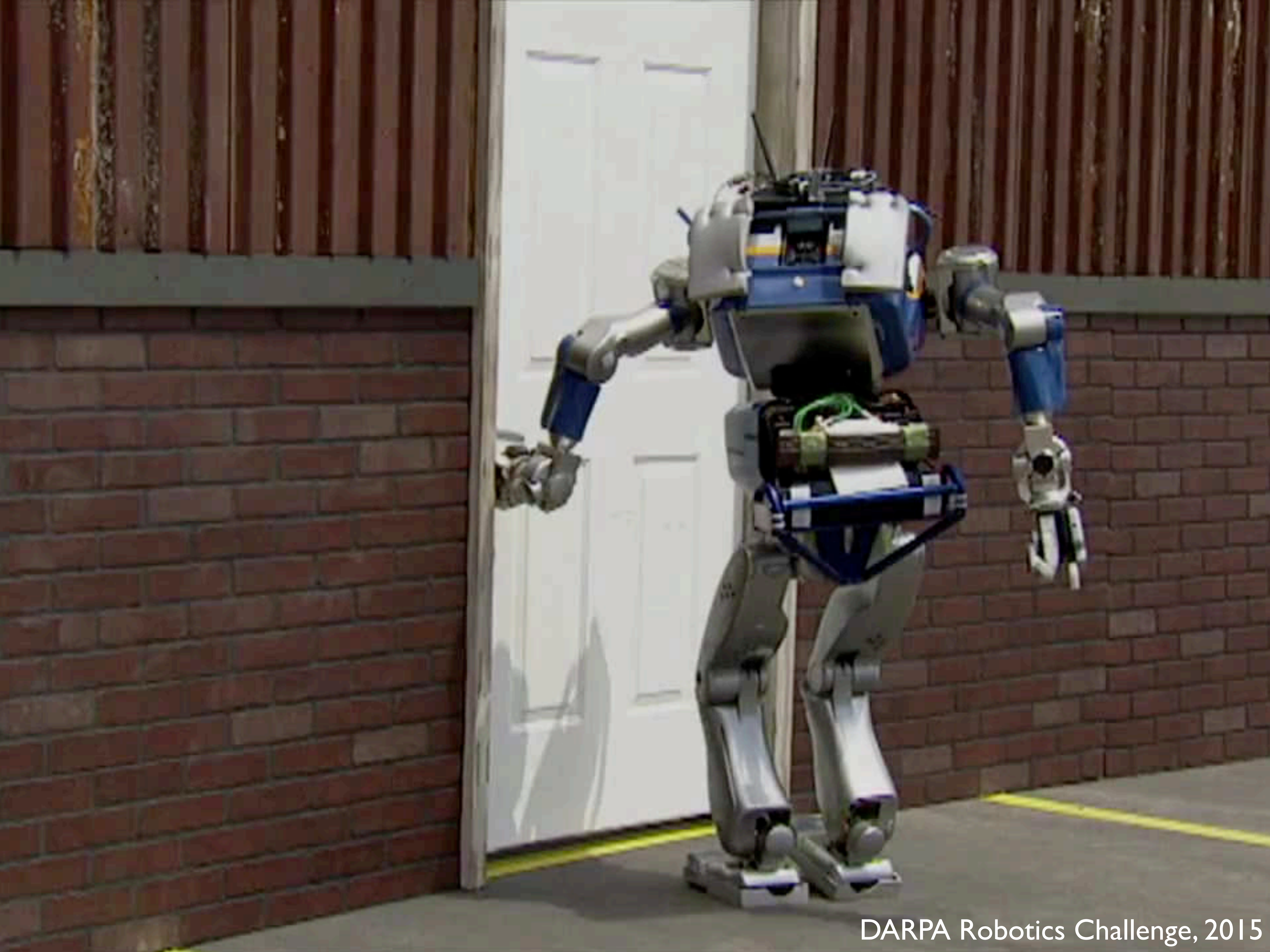
… the robot can adapt its policy to each human partner



Part 2:
Learning for damage recovery

Jean-Baptiste Mouret

jeanbaptiste.mouret@inria.fr

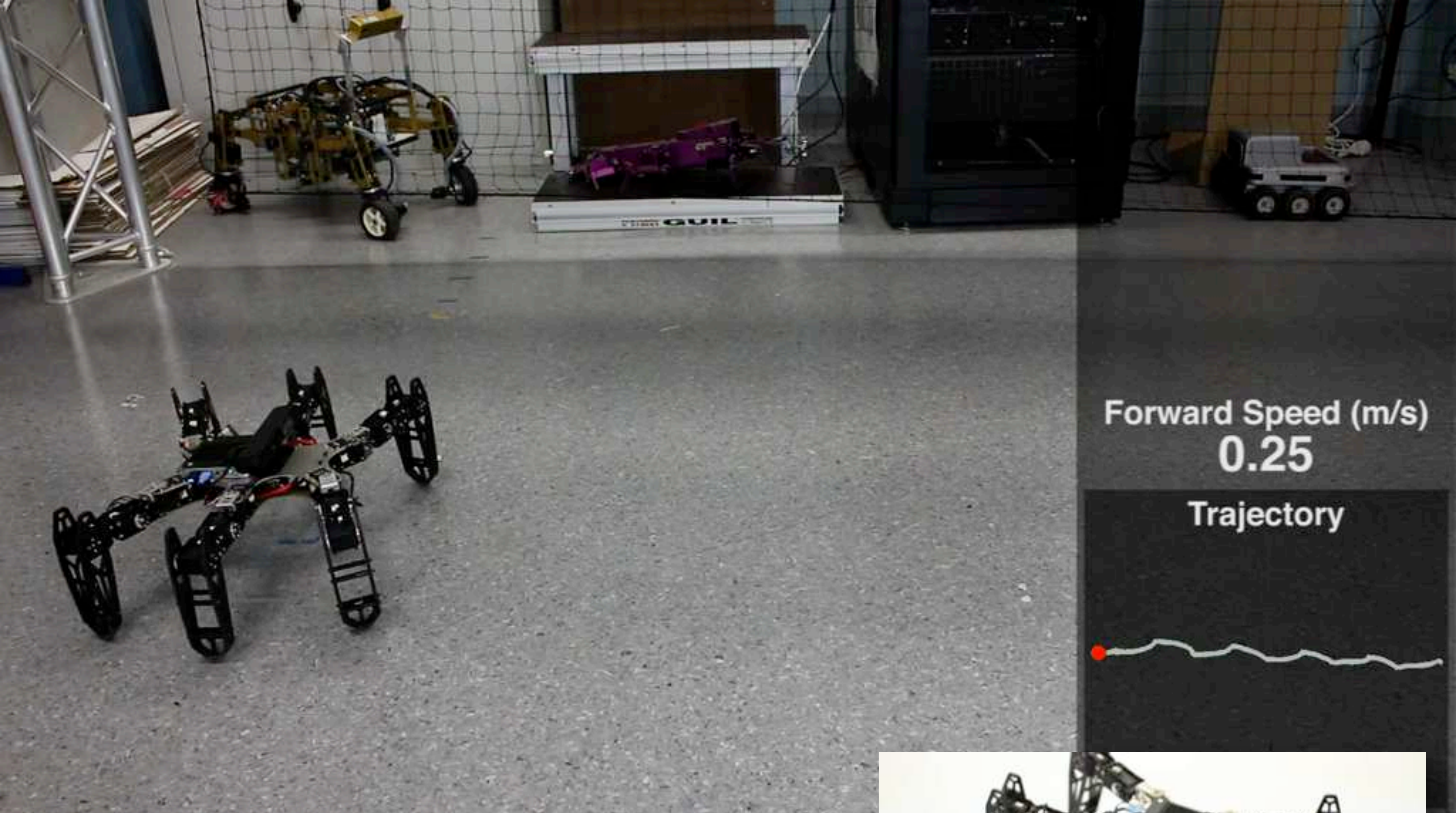


DARPA Robotics Challenge, 2015

The issue with robots is not that they fail & break…

… it is that they do not get back on their feet and try again

[If something unexpected happens, the mission is aborted!]



- Controller : periodical signals (36 parameters)
- Performance: covered distance in 5 seconds
- Performance evaluated onboard (RGB-D visual odometry)

What can we do?

- **The medical approach:**
  - diagnose the problem
  - try to fix it

- expensive (sensors)
- need to place the sensors "at the right place" = anticipate


Trial and error learning… in minutes!
(they do not « understand» the injury)

Micro-data learning

AlphaGo

30 million positions
+ self-play

« Big Data » **Amount of data** **« Micro data »**

38 days
of learning

1-20 trials

Deep learning ?

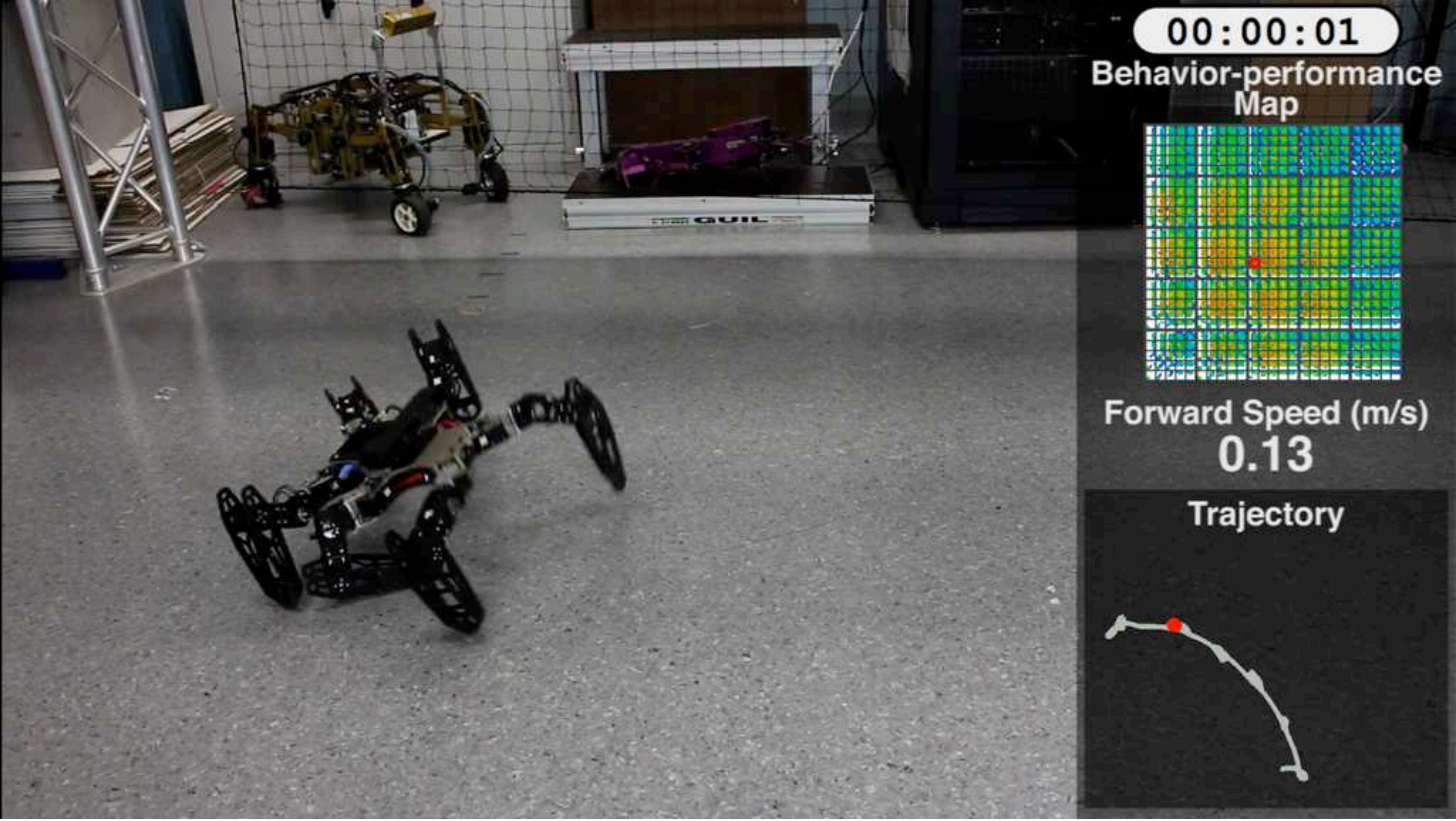
Learning with
robots

Five precepts for micro-data learning

1. Choose wisely what to test next (active learning)
   ➡ OK to trade data resources for computational resources
2. Know what you know
   ➡ Take the uncertainty into account when selecting what to test
3. Use prior knowledge
   i. use an easy search space (possibly, design it automatically)
   ii. make prior knowledge explicit
   iii. use everything we know (e.g. simulator of the intact robot)
4. Exploit every bit of information from each test
   ➡ e.g., use all the points of a trajectory
5. Only learn what is necessary
   ➡ e.g, do not reinvent control theory

All this precepts should be combined

JB Mouret. *Micro-data learning: the other end of the spectrum.* ERCIM News. 2016



Two main ideas:

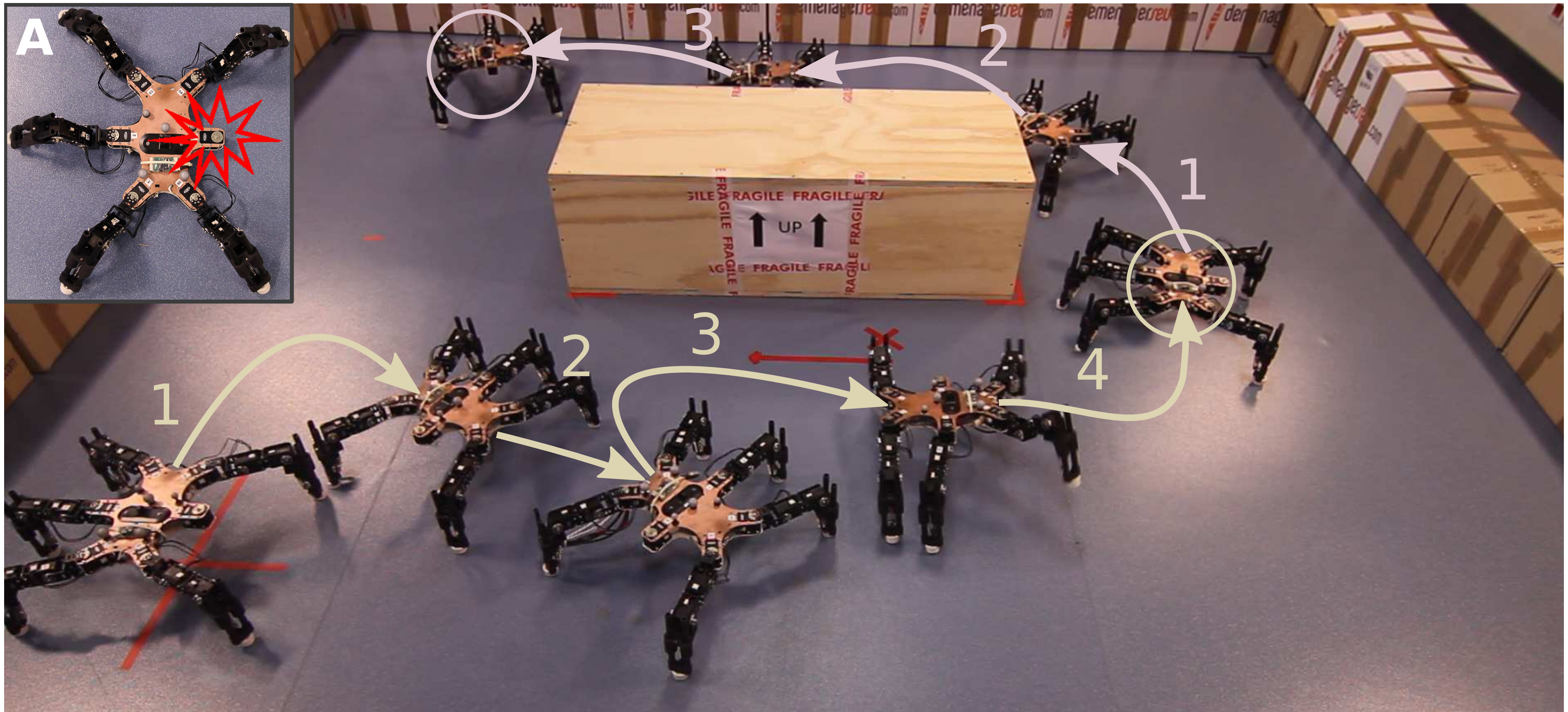
1. generate priors with a simulation of the intact robot
2. choose the next trial using Bayesian optimization (i.e. take uncertainty of predictions into account)

Cully, A., Clune, J. , Tarapore, D. and Mouret, J.-B.
Robots that can adapt like animals.
Nature. Vol 521 Pages 503-507. (2015).

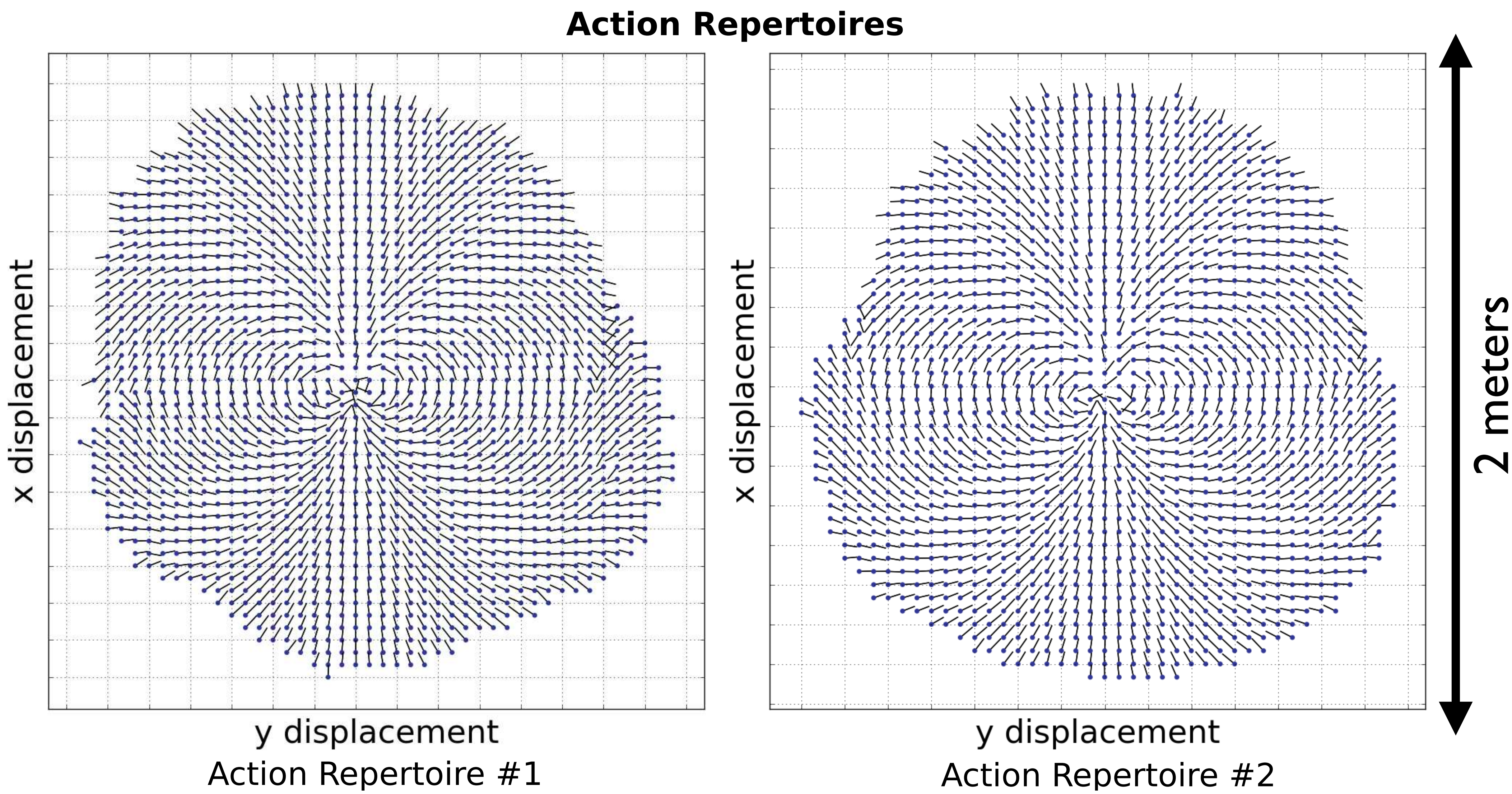
Trial & error damage recovery in ~10 trial but…

- This is episodic learning: the robot is reset after each trial
  - ➡ learn without reset
  - ➡ … while taking the environment into account (obstacles)
  - ➡ "learn while doing": trials useful for the task

- We know a dynamics simulator of the intact robot & the environment
- We don't know the damage (could be anything)



Results of MAP-Elites



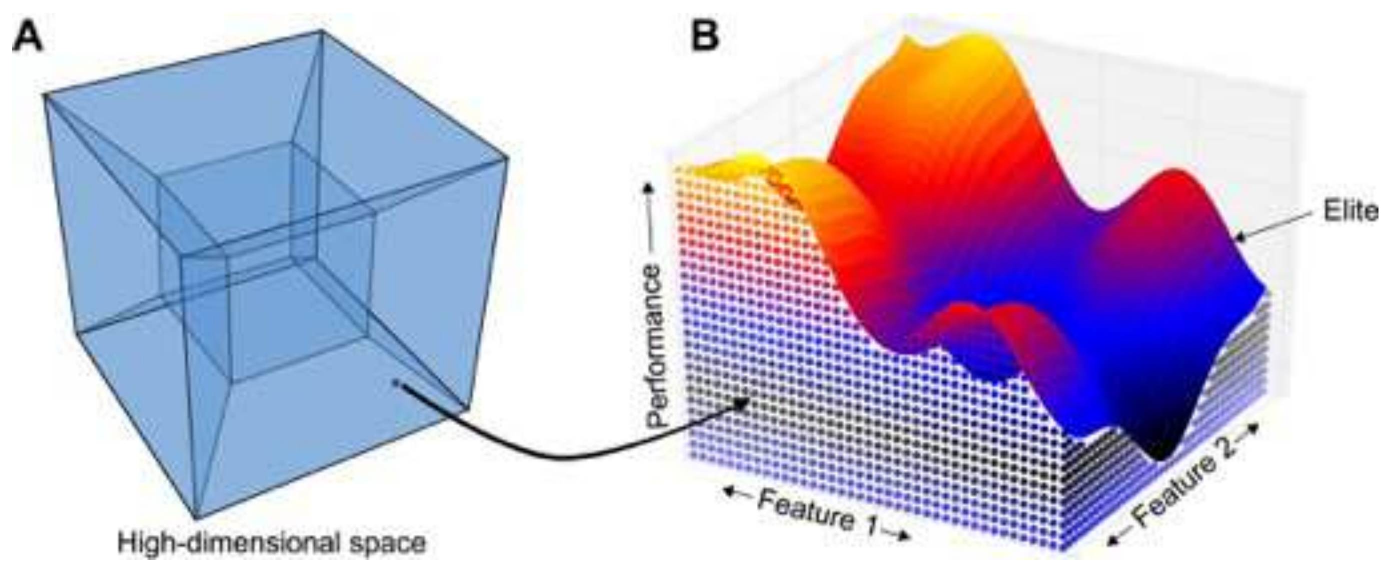
36 parameters ➟ 1500 good controllers in a 2D space

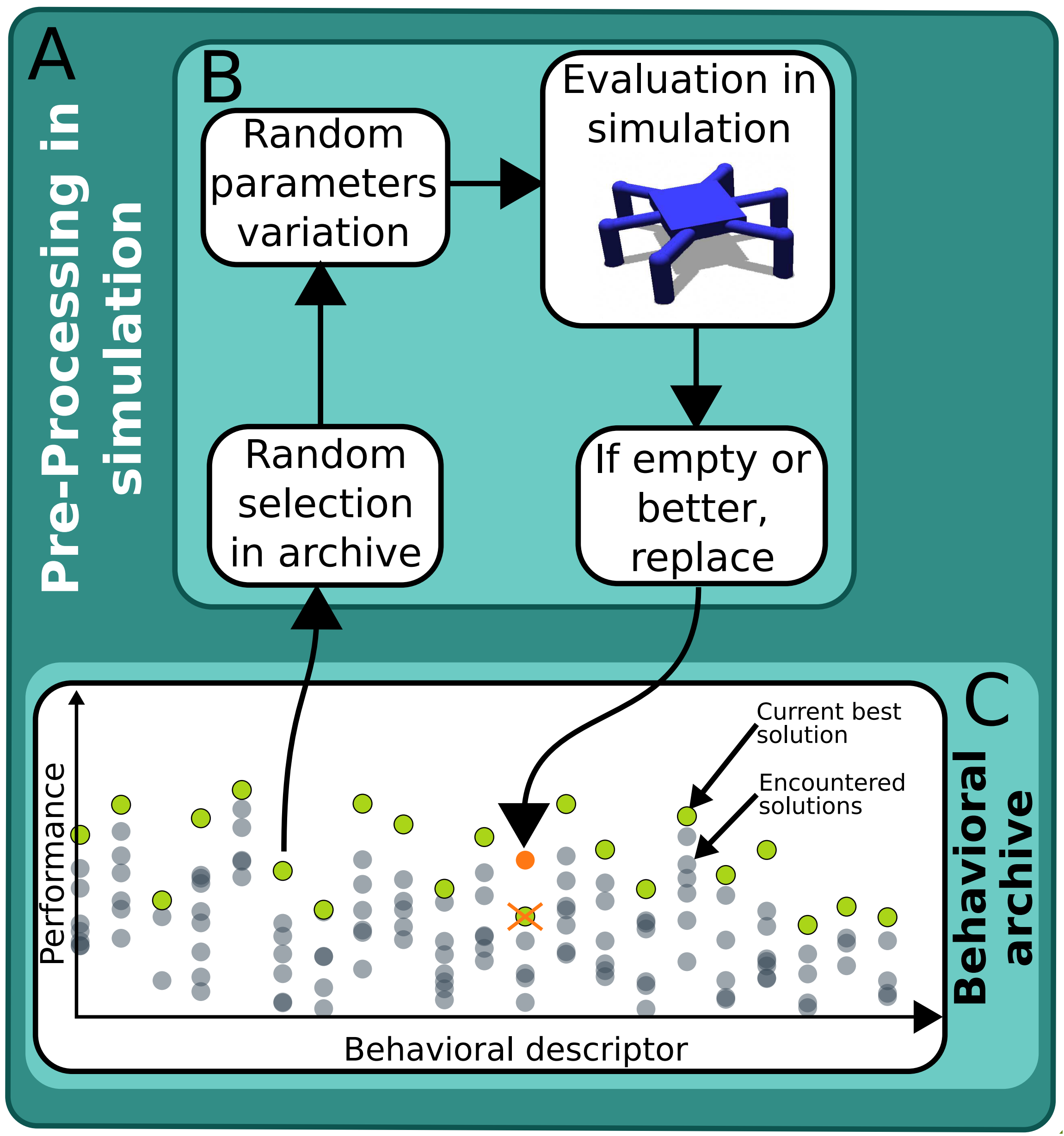
Performance: does the robot follow a circular trajectory?

Learn with a simulation of the intact robot

Breaking the complexity: pre-computing a repertoire

Multi-dimensional Archive of Phenotypic Elites





Goal: find many <u>good</u> alternatives
➠ The **elites** of the search space

Not random sampling at all: you do not find good walking controllers "by chance"

Mouret, J.-B., and J. Clune. "Illuminating search spaces by mapping elites." arXiv preprint arXiv:1504.04909 (2015).

… but the repertoire needs to be corrected

- Learn a modification of the repertoire with a Gaussian process (one GP for each dimension — x, y)

$$P(f(\mathbf{x})|\mathbf{P}_{1:t+1}),\mathbf{x}) = \mathcal{N}(\mu_{t+1}(\mathbf{x}), \sigma^2_{t+1}(\mathbf{x}))$$

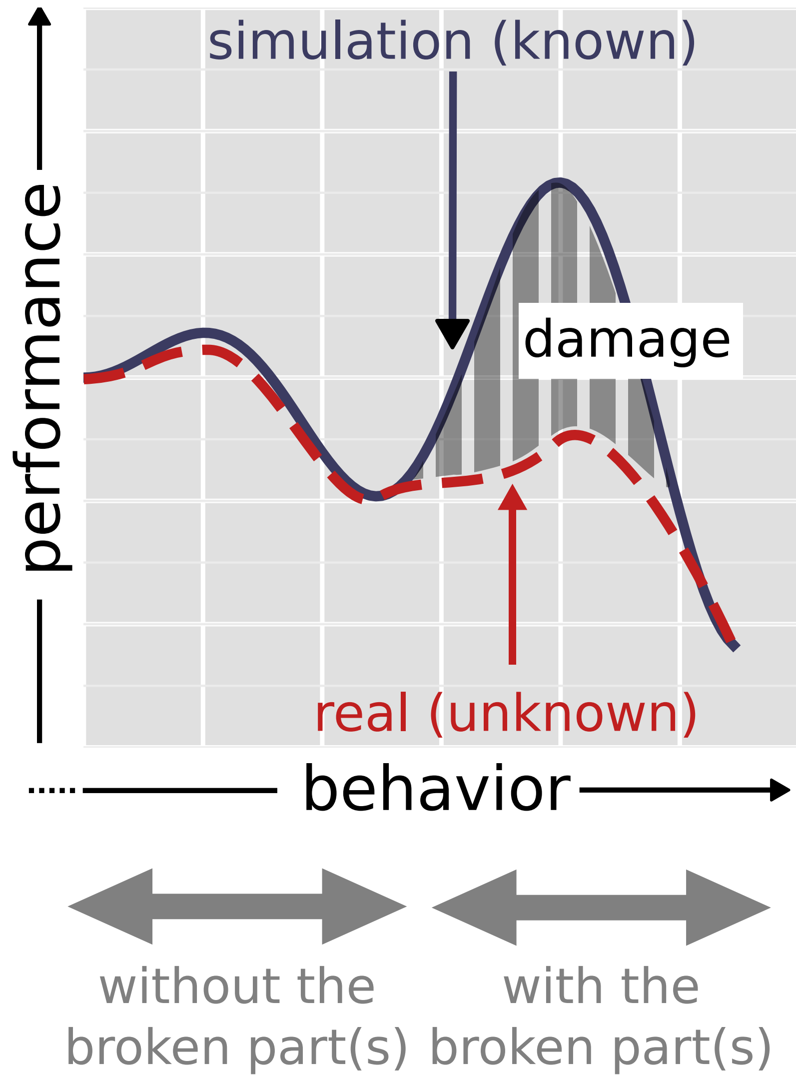
where

$$\mu_{t+1}(\mathbf{x}) = \mathcal{A}(\mathbf{x}) + \mathbf{k}^t\mathbf{K}^{-1}(\mathbf{P}_{1:t+1} - \mathcal{A}(\mathbf{y}_{1:t+1}))$$

$$\sigma^2_{t+1}(\mathbf{x}) = k(\mathbf{x},\mathbf{x}) - \mathbf{k}^t\mathbf{K}^{-1}\mathbf{k}$$

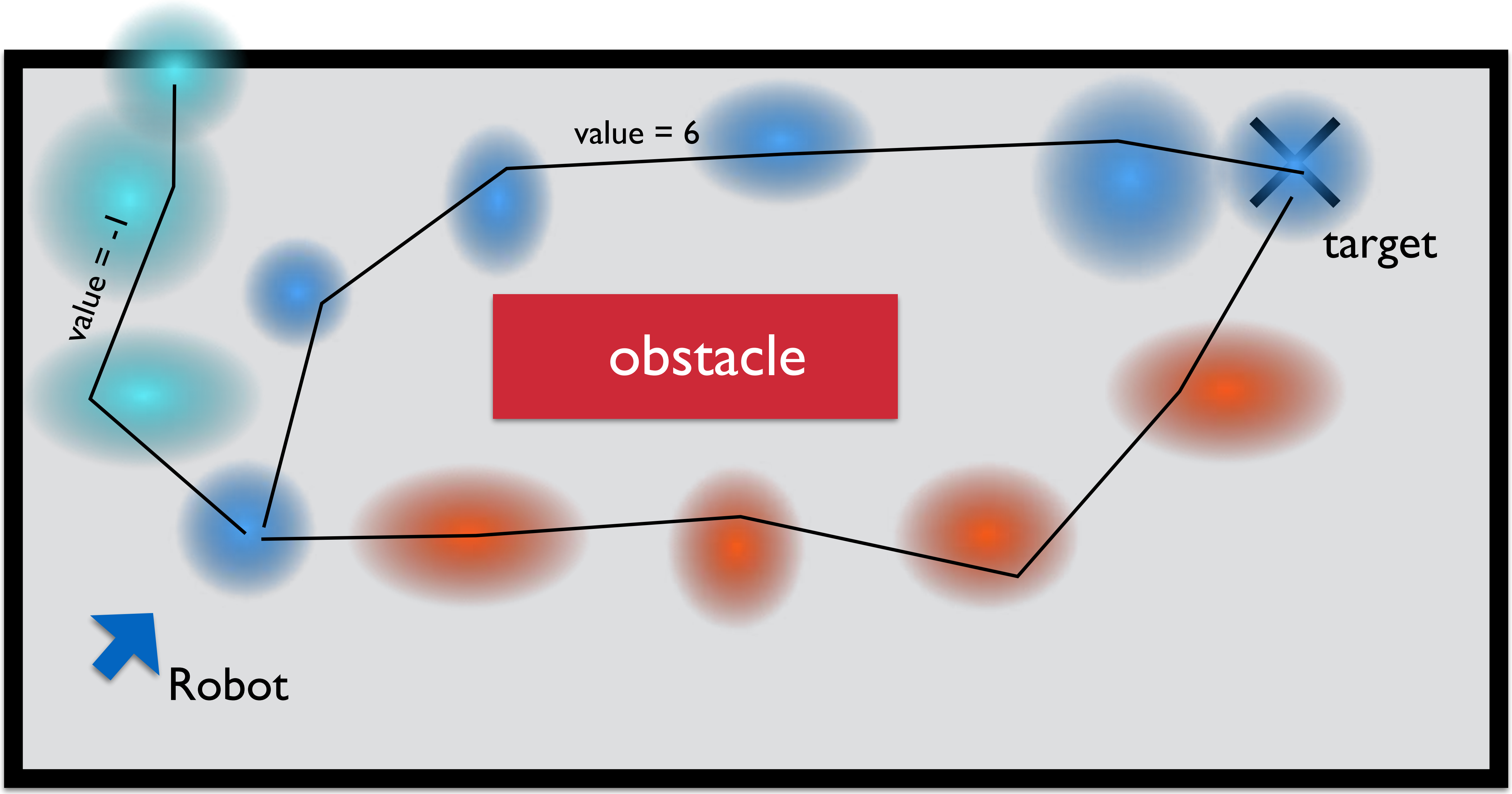
$$\mathbf{K} = \begin{bmatrix} k(\mathbf{y}_1,\mathbf{y}_1) + \sigma^2_{noise} & \cdots & k(\mathbf{y}_1,\mathbf{y}_t) \\ \vdots & \ddots & \vdots \\ k(\mathbf{y}_t,\mathbf{y}_1) & \cdots & k(\mathbf{y}_t,\mathbf{y}_t) + \sigma^2_{noise} \end{bmatrix}$$

$$\mathbf{k} = \begin{bmatrix} k(\mathbf{x},\mathbf{y}_1) & k(\mathbf{x},\mathbf{y}_2) & \cdots & k(\mathbf{x},\mathbf{y}_t) \end{bmatrix}$$



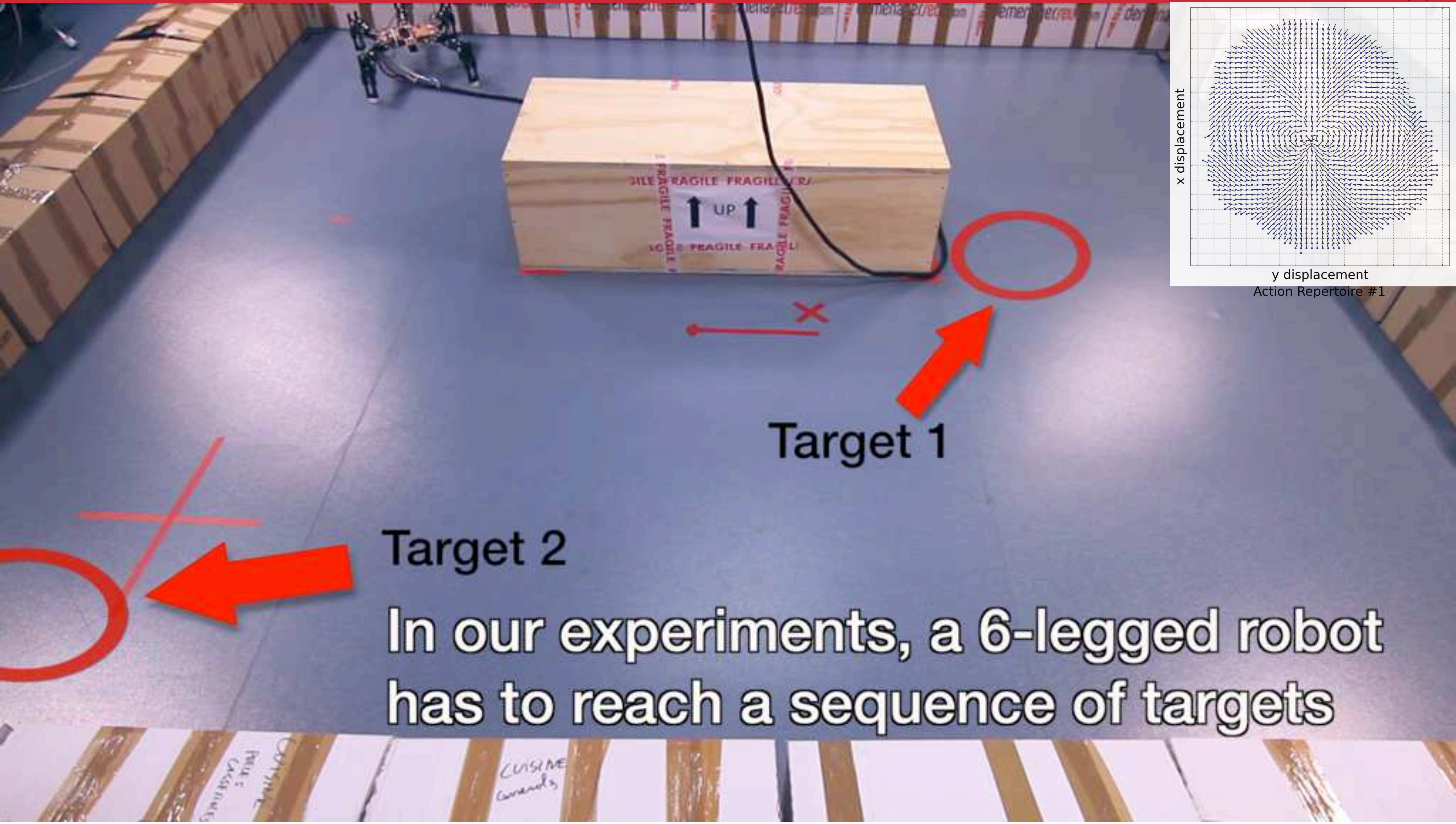
What to try next?

This is a planning problem (with uncertainty because of the GP)
➟ can be solved with Monte Carlo Tree Search (MCTS)



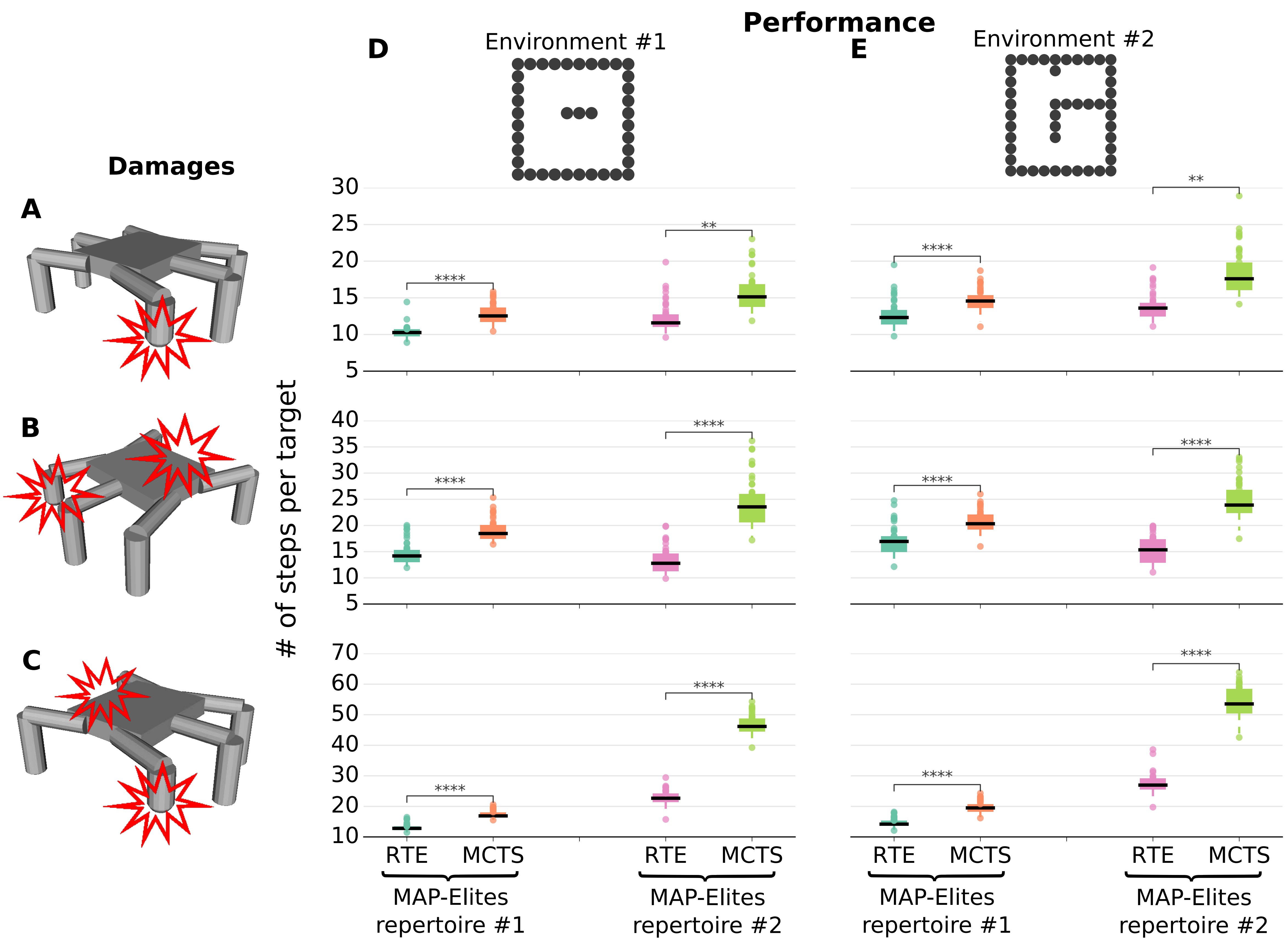
1. Use Monte-Carlo rollouts to evaluate the probability distribution of value of each behavior / policy
2. Choose the most interesting action / policy
3. Run it on the robot
4. Update the models (GP), which reduces the uncertainty & improve predictions
5. …

Reset-free Trial & Error (RTE)



K. Chatzilygeroudis, V. Vassiliades, and J.-B. Mouret (2016). *Reset-free Trial-and-Error Learning for Data-Efficient Robot Damage Recovery*. ArXiv.

Does learning help?



Conclusion

No need to model a damage to continue the mission!

Reset-free damage recovery

… on a "complex" robot / policy (36 parameters to learn)

… in a few minutes

…. with reasonable computation times (< 30 s)

➡ **a "realistic scenario" for damage recovery**

Future work

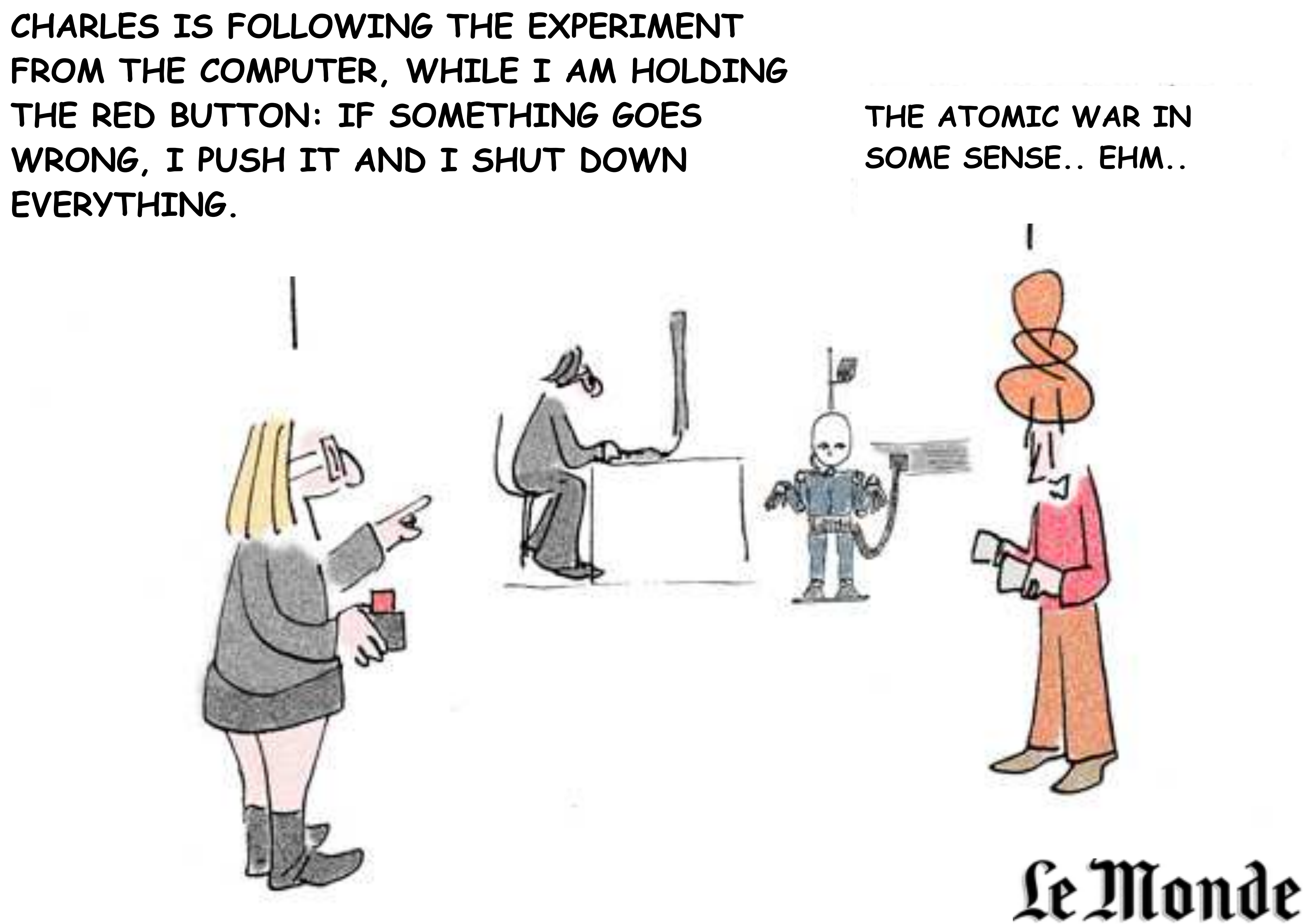
- ➡ humanoid robots (iCub)
- ➡ include safety constraints (cf Paspaspyros et al. NIPS Workshop 2016)
- ➡ use trajectories to improve predictions (use more from each trial)

Papaspyros V, Chatzilygeroudis K, Vassiliades V, Mouret JB. *Safety-Aware Robot Damage Recovery Using Constrained Bayesian Optimization and Simulated Priors*. Proc. Of the NIPS workshop on Bayesian Optimization. 2016

Acknowledgements

R. Calandra (TU Darmstadt), K. Chatzilygeroudis (Inria), J. Clune (U. Wyoming), A. Cully (Imp. College), M. P. Deisenroth (Imp. College), L. Natale (IIT), F. Nori (IIT), J. Peters (TU Darmstadt), D. Tarapore (U. Southampton), V. Vassiliades (Inria), E. Zibetti (Paris 8)

Thank you!
Questions ?



Comics by Fiamma Luzzati - Le Monde - April 2014