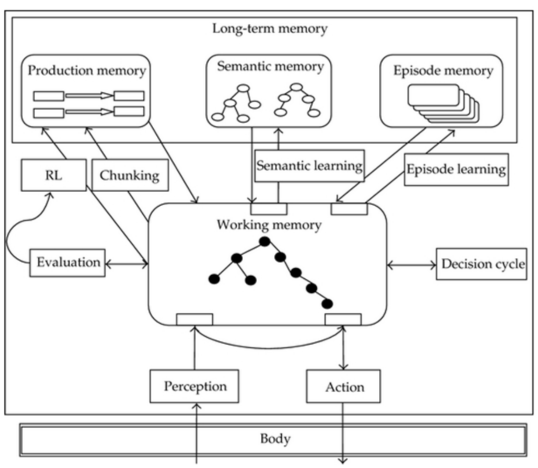
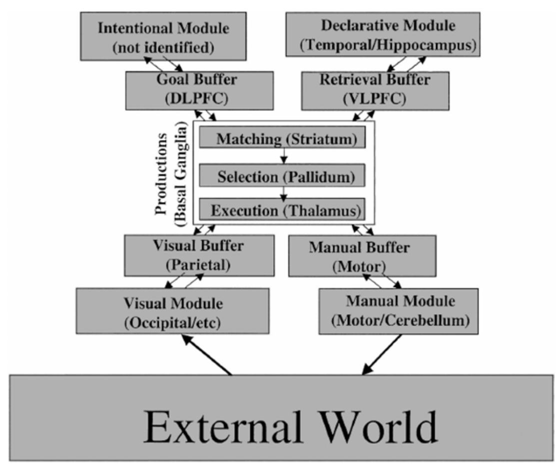
Cognitive Robotics

- Cognitive Robotics
  - Embodied AI/Embodied CS
    - Robot capable of perception, reasoning, learning, deliberation, planning, acting, interacting, etc.
  - Cognitive Architecture
    - Unified Theory of Cognition
    - Cognitive Models





Cognitive Systems and Robotics

- Since 2001: Cognitive Systems intensely funded by the EU "Robots need to be more robust, context-aware and easy-to-use. Endowing them with advanced learning, cognitive and reasoning capabilities will help them adapt to changing situations, and to carry out tasks intelligently with people"

Research Areas

- Biorobotics
  - Bio-inspirata, biomimetica, etc.

- Enactive Robotics
  - Dynamic interaction with the environment

icub

- Developmental Robotics (Epigenetic)
  - Robot learns as a baby
  - Incremental sensorimotor and cognitive ability [Piaget]

Cog

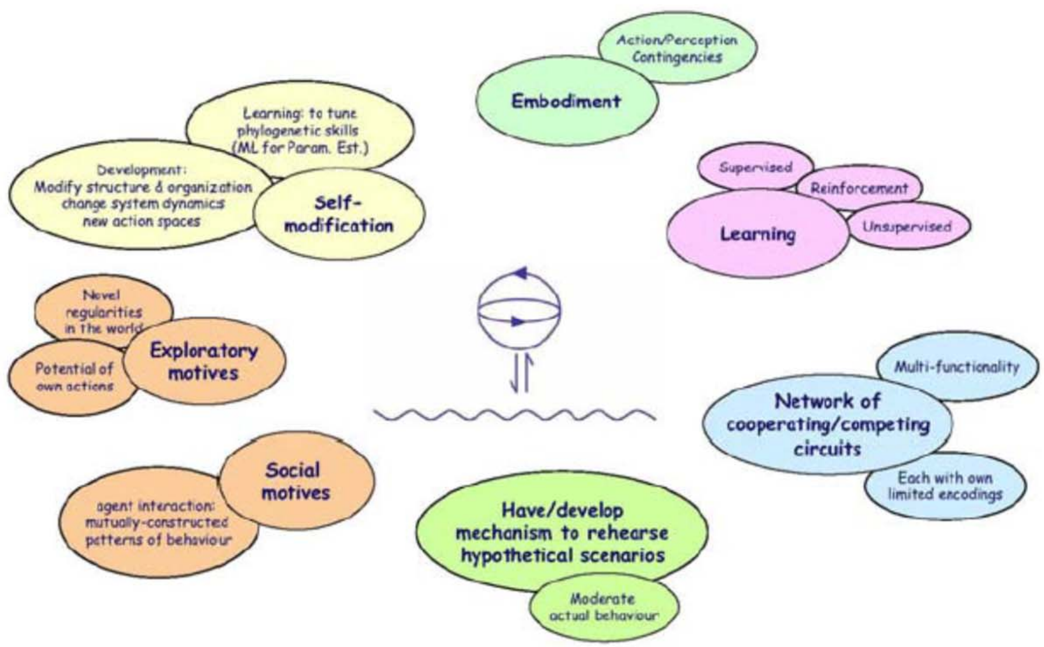
- Neuro-Robotics
  - Models from cognitive neuroscience
  - Prosthesis, wearable systems, BCI, etc.

Developmental Robotics

- iCub Platform
  - Italian Institute of Technology
  - EU project RobotCub: *open source cognitive humanoid platf.*

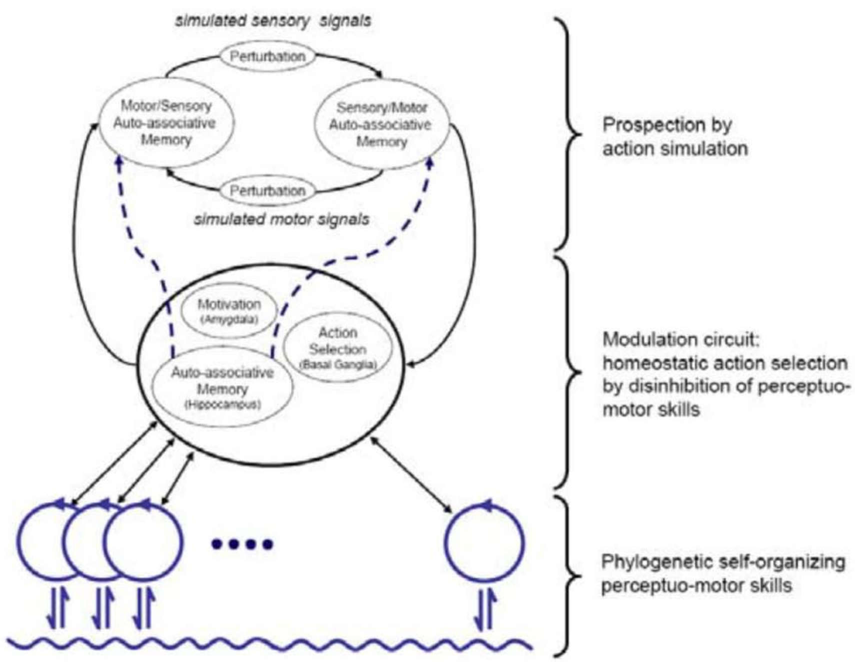
Developmental Robotics

- iCub Platform
  - Italian Institute of Technology
  - EU project RobotCub: *open source cognitive humanoid platf.*



Developmental Robotics

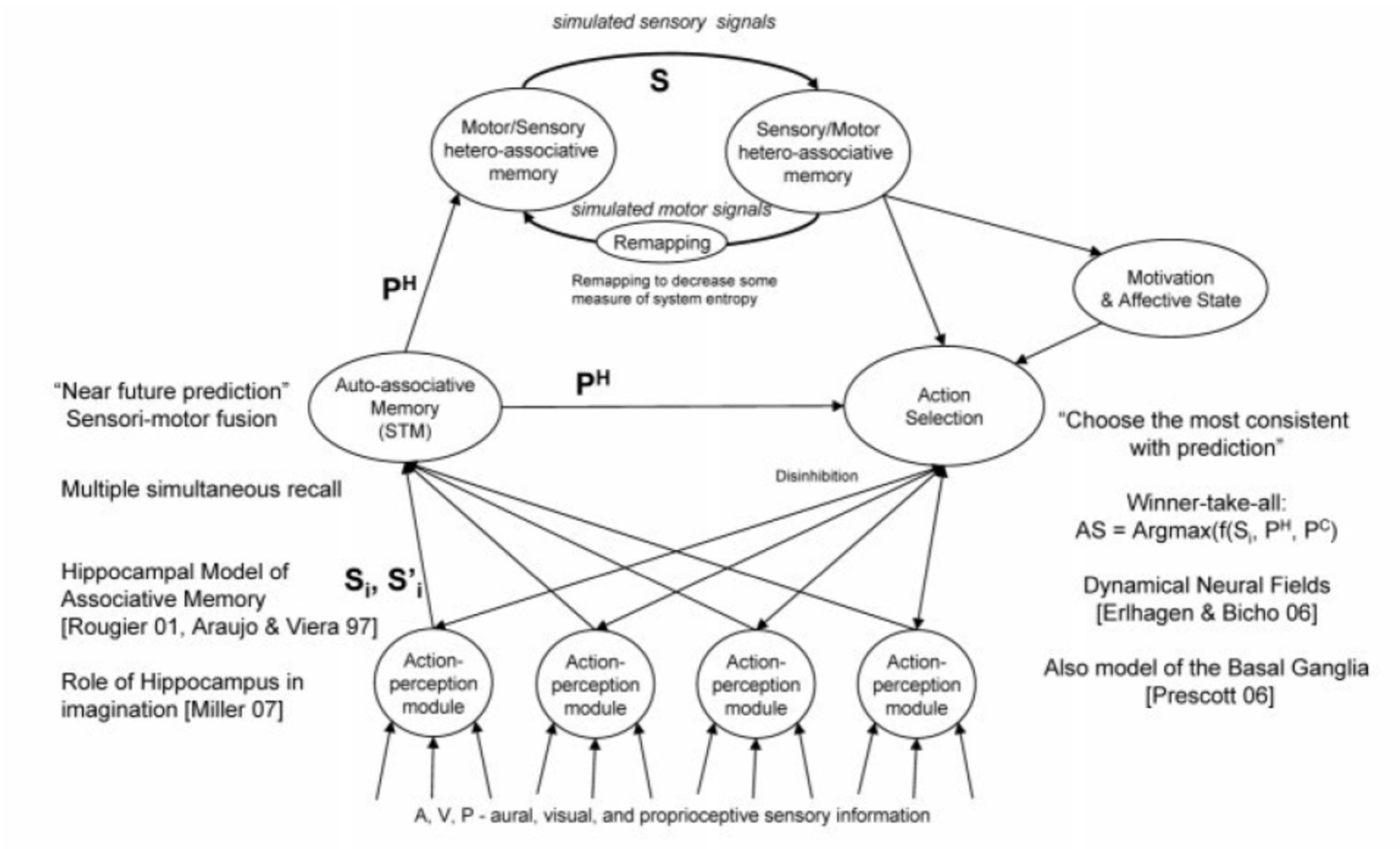
- iCub Platform



Scenario Capabilities: cognitive perception/action behaviours
Object tracking through occlusion (smooth pursuit & saccades)
Learn to coordinate VOR & tracking
Learn to reach towards a fixation point
Attention and action selection by modulation of capabilities
Conditional modulation based on anticipation
Construct sensorimotor maps & cross-modal maps
Learn by demonstration (crawling & constrained reaching)
Exploratory, curiousity-driven, action
Experience-based action selection based on interaction histories
Navigate based on local landmarks and ego-centric representations
Quasi-independent Phylogenetic Capabilities
Saccadic re-direction of gaze towards salient multi-modal events
Focus attention and direct gaze on human faces
Ocular modulation of head pose to centre eye gaze
Move the hand(s) towards the centre of the visual field
Stabilize & integrate of saccadic percepts
Stabilize gaze with respect to self-motion (VOR)
Create attention-grabbing stimuli
Gait control
Component Capabilities
Compute optical flow
Compute visual motion with ego-motion compensation
Segmentation of the flow-field based on similarity of flow parameters
Segmentation based on the presence of a temporally-persistent boundary
Fixation and vergence
Gaze control: smooth pursuit with prediction; possibly tuned by learning
Classification of groups of entities based on low numbers
Classification of groups of entities based on gross quantity
Detection of mutual gaze
Detection of biological motion

Developmental Robotics

- iCub Platform

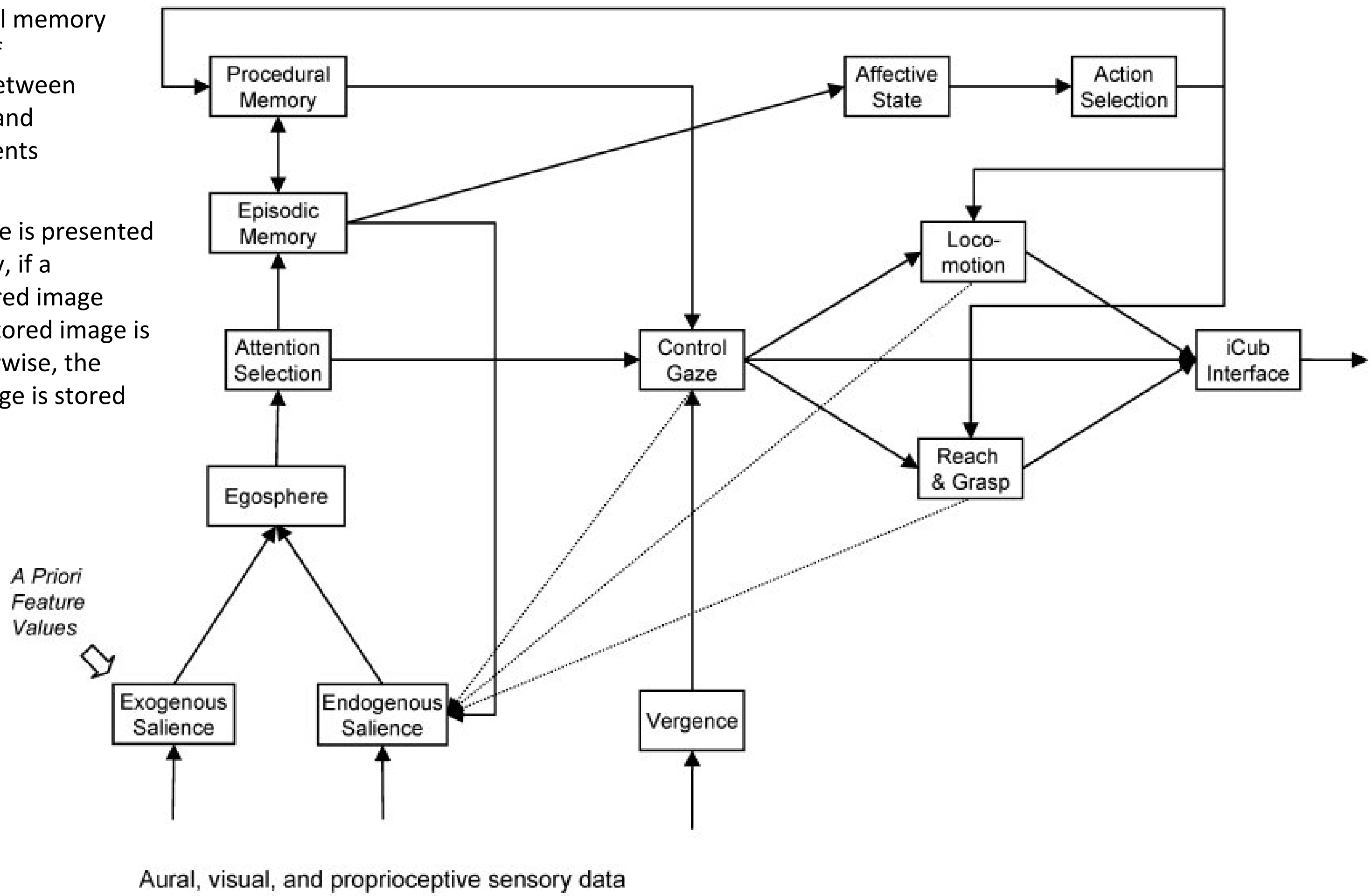


Developmental Robotics

- Cognitive Architecture:

The procedural memory is a network of associations between action events and perception events

When an image is presented to the memory, if a previously stored image matches the stored image is recalled; otherwise, the presented image is stored



Affordance

- **Affordance:** property of an object or a feature of the immediate environment, that indicates how to interface with that object or feature ("action possibilities" latent in the environment [Gibson 1966])
  - "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, but the noun affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment."

  - Interaction design (*The Design of Everyday Things* [Norman 1988]): concept of affordance subjective rather than objective
  - Developmental psychology:
    - Means-end behaviors, predicting effects and imitation
  - Robotics: exploration, learning, grasping [Jamone et al. 2016]
    - Learning affordance, perception of affordance

Affordance

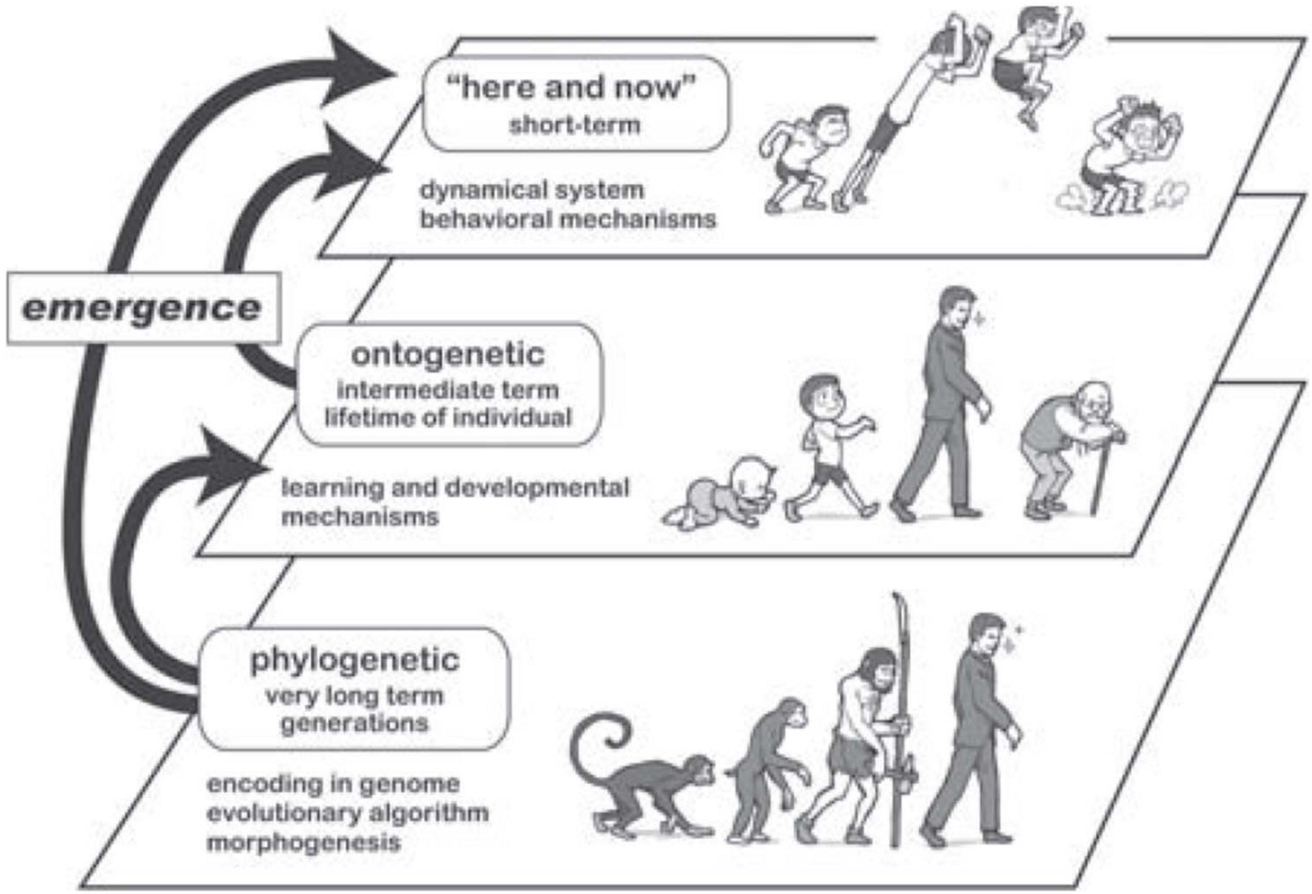
- **Affordance:** property of an object or a feature of the immediate environment, that indicates how to interface with that object or feature ("action possibilities" latent in the environment [Gibson 1966])
  - Robotics: exploration, learning, grasping [Jamone et al. 2016]
    - Learning affordance, perception of affordance

Developmental Robotics

- JST ERATO Asada Synergistic Intelligence Project:
  - Child-robot with Biomimetic Body [Asada et al. 2006]

Evolution, Development, Learning

- Phylogeny Ontogenesis



Rolf Pfeifer and Josh Bongard

Research Areas

- Biorobotics
  - Bio-inspirata, biomimetica, etc.
- Enactive Robotics
  - Dynamic interaction with the environment

icub

- Developmental Robotics (Epigenetic Robotics)
  - Robot learns as a baby
  - Incremental sensorimotor and cognitive ability [Piaget]
- Neuro-Robotics
  - Models from cognitive neuroscience
  - Prosthesis, wearable systems, BCI, etc.

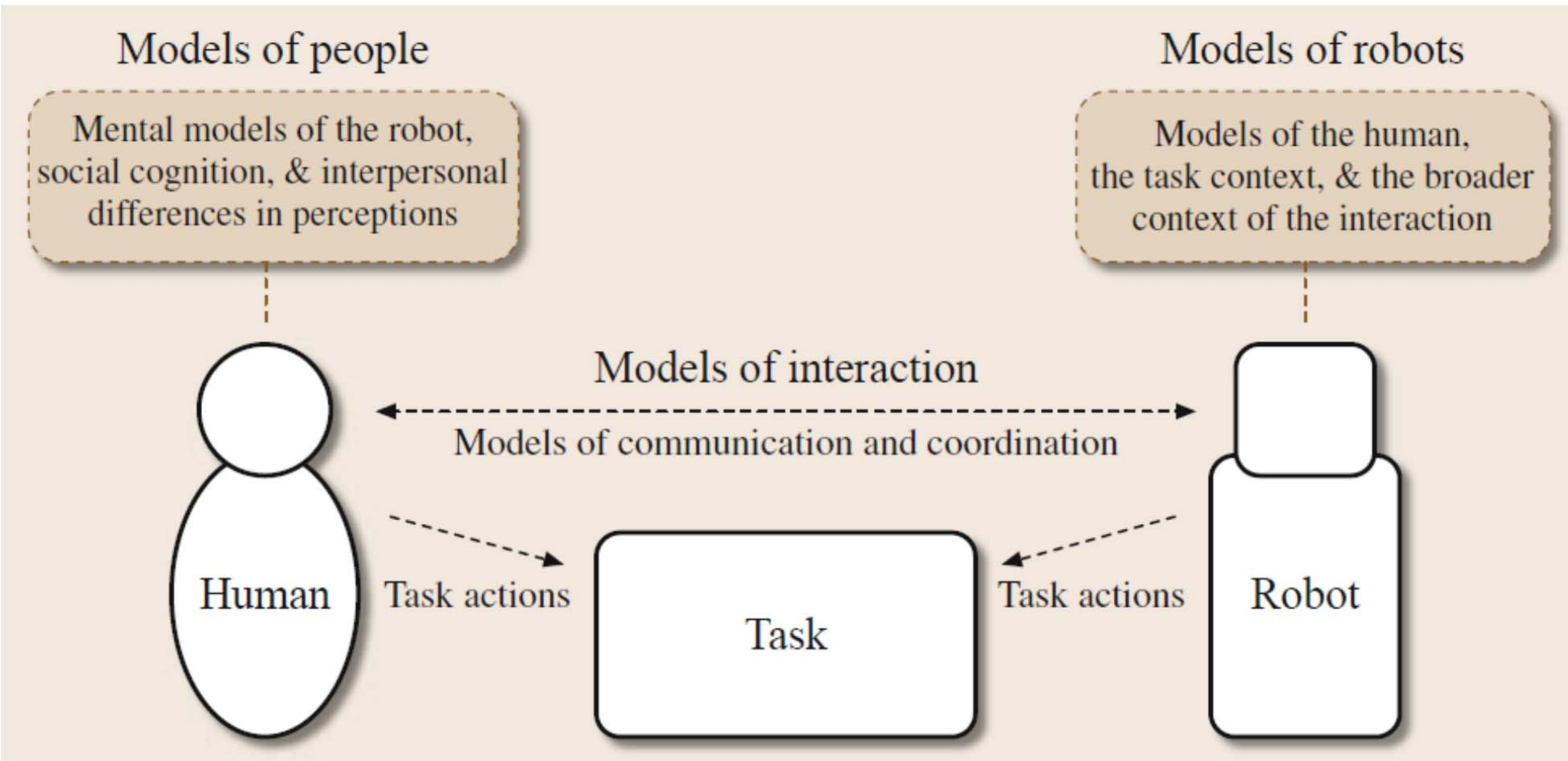
Cog

Cognitive Science as Enabler Cognitive Robotics

- Hypothesis:
  - A system using human-like representations and processes will enable better collaboration with people than a computational system that does not
    - Similar representations and reasoning mechanisms make it easier for humans to work with the system; more compatible
  - For close collaboration, systems should act “naturally”
    - i.e. not do something or say something in a way that detracts from the interaction/collaboration with the human
    - Robot should accommodate humans; not other way around
  - Solving tasks from “first principles”
    - Humans are good at solving some tasks; let’s leverage human’s ability

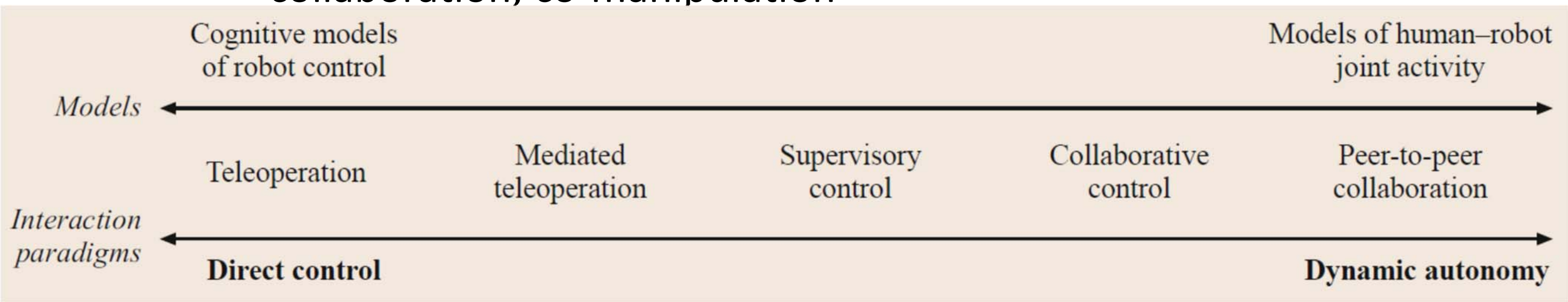
Cognitive Robotics and HRI

- Human-Robot Interaction as a uniform Cognitive System
  - Human models, Robot models, interaction models



Cognitive Robotics and HRI

- Interaction type:
  - Physical, Cognitive, Social
  - Proximity:
    - Same place, different place
  - Critical:
    - Search and Recue, Service, Entertainment, etc.
  - Autonomy:
    - Teleoperation, supervision, interaction, coordination, collaboration, co-manipulation



Human-Robot Interaction

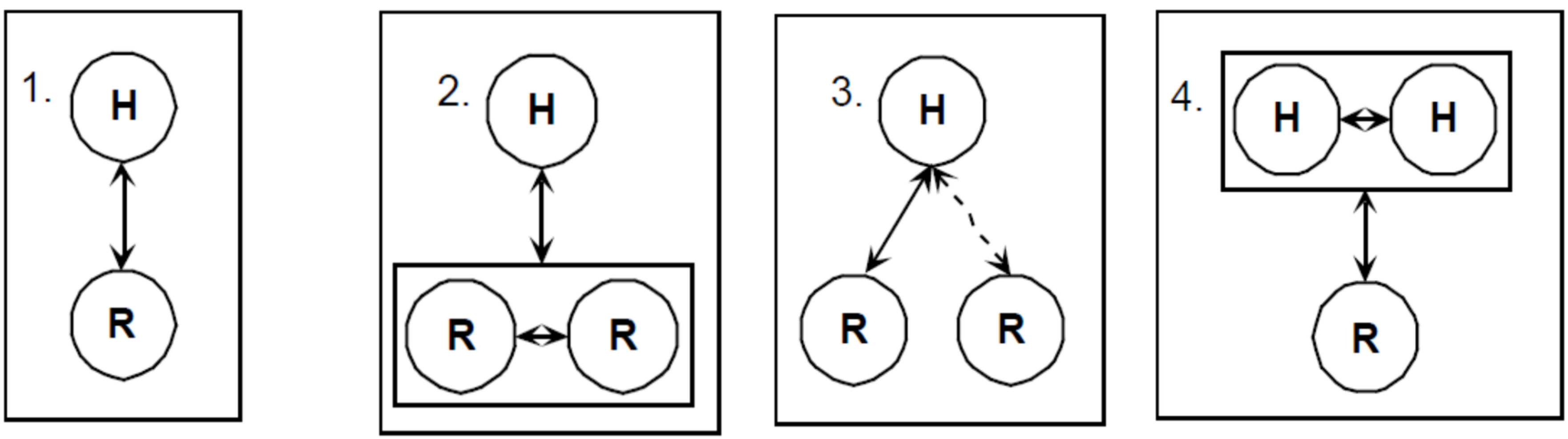
- Specific features:
  - Physical/cognitive interaction with embodied intelligence
  - Social relation between humans and robots
  - Complex, dynamic, unpredictable environment/human

Taxonomies in HRI

- Physical vs. Cognitive
- Co-located vs. Remote
- Team configurations [Yanco2002]

Taxonomies in HRI

- Team configurations [Yanco2002]

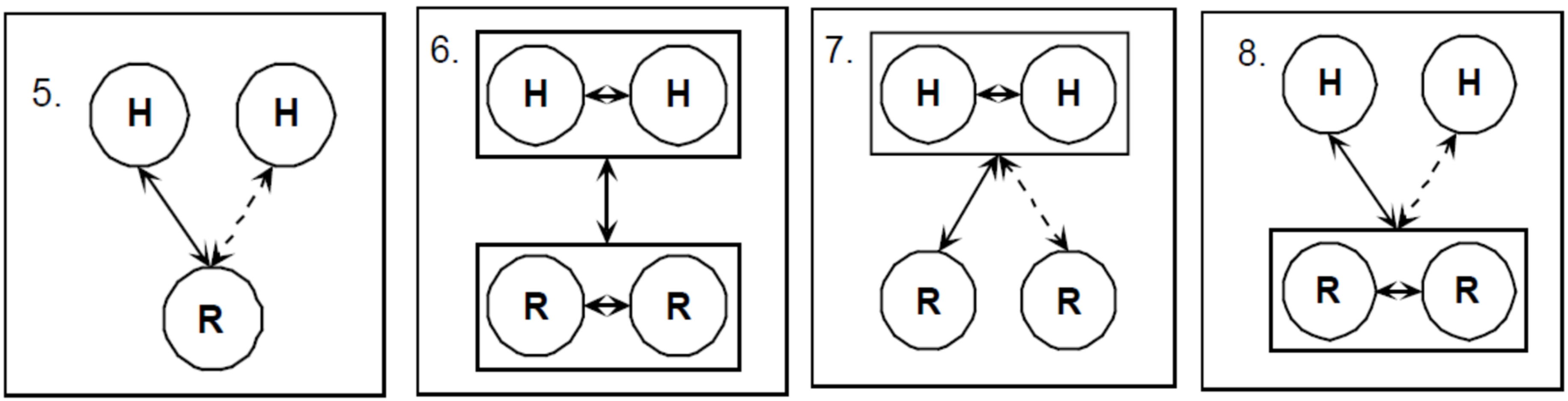


Human-Robot   Human-Team   Human-Robots   Robot-Team

HUMAN-ROBOT-RATIO
ROBOT-TEAM-COMPOSITION
(homogeneous/heterogeneous)
ROBOT-MORPHOLOGY
(anthropomorphic,zoomorphic,functional)

Taxonomies in HRI

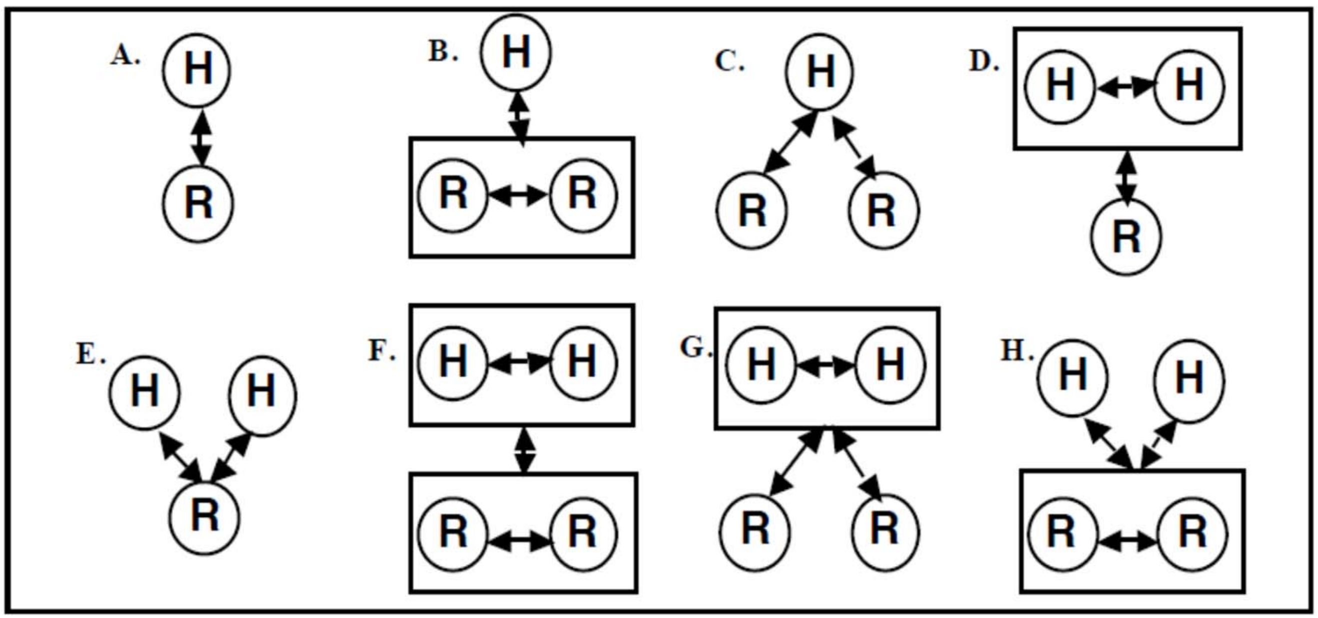
- Team configurations [Yanco2002]



mHuman-1Robot TeamR-TeamH TeamH-mRobot mHuman-TeamR

Taxonomies in HRI

- Team configurations [Yanco2002]



one human, one robot;
one human, robot team; *one human, multiple robots*;
human team, one robot; *multiple humans, one robot*;
human team, robot team; *human team, multiple robots*;
and *multiple humans, robot team*.

Taxonomies in HRI

- Space and Time:

		Time	
		Same	**Different**
Space	**Same**	Robot Wheelchair	Manufacturing robots
	Different	Urban Search and Rescue	Mars Rover

Table 1: Time-space taxonomy category, with examples.

- Criticality:
  - High: Search and Rescue
  - Medium: Service
  - Low: Game/Social

Taxonomies in HRI

- Space and Time:

		Time	
		Same	**Different**
Space	**Same**	Robot Wheelchair	Manufacturing robots
	Different	Urban Search and Rescue	Mars Rover

Table 1: Time-space taxonomy category, with examples.

- Criticality:
  - High: Search and Rescue
  - Medium: Service
  - Low: Game/Social

Taxonomies in HRI

- PHYSICAL_PROXIMITY:
  - *avoiding*, *passing*, *following*, *approaching*, and *touching*, none
- Decision Support:
  - AVAILABLE-SENSORS, PROVIDED-SENSORS,
  - SENSOR-FUSION, PRE-PROCESSING;

Taxonomies in HRI

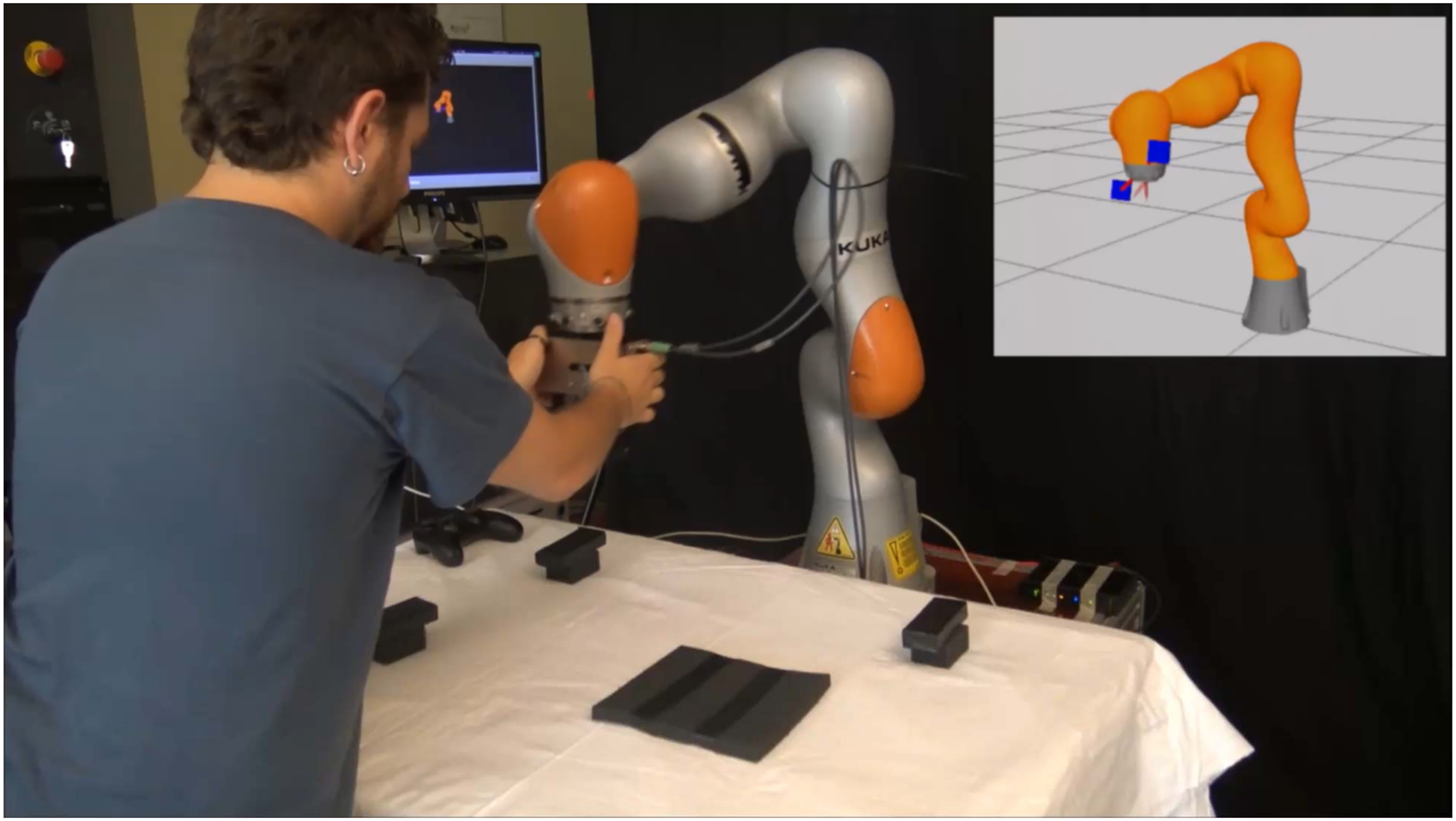
- Team hierarchy
  - Conflict resolution
  - Especially for peer-based relationships (co-X)
  - Active roles
  - Supervisor
  - Operator
  - Mechanic / Assistant
  - Peer
  - Slave
  - Passive roles
  - Patients
  - Visitors
  - Bystanders

Service Robotics

- Robot co-worker (CoBot):
  - Human monitoring
  - Intention/activity recognition
  - Physical and cognitive interaction
  - Communication and dialogue
  - Collaboration
  - Teaching

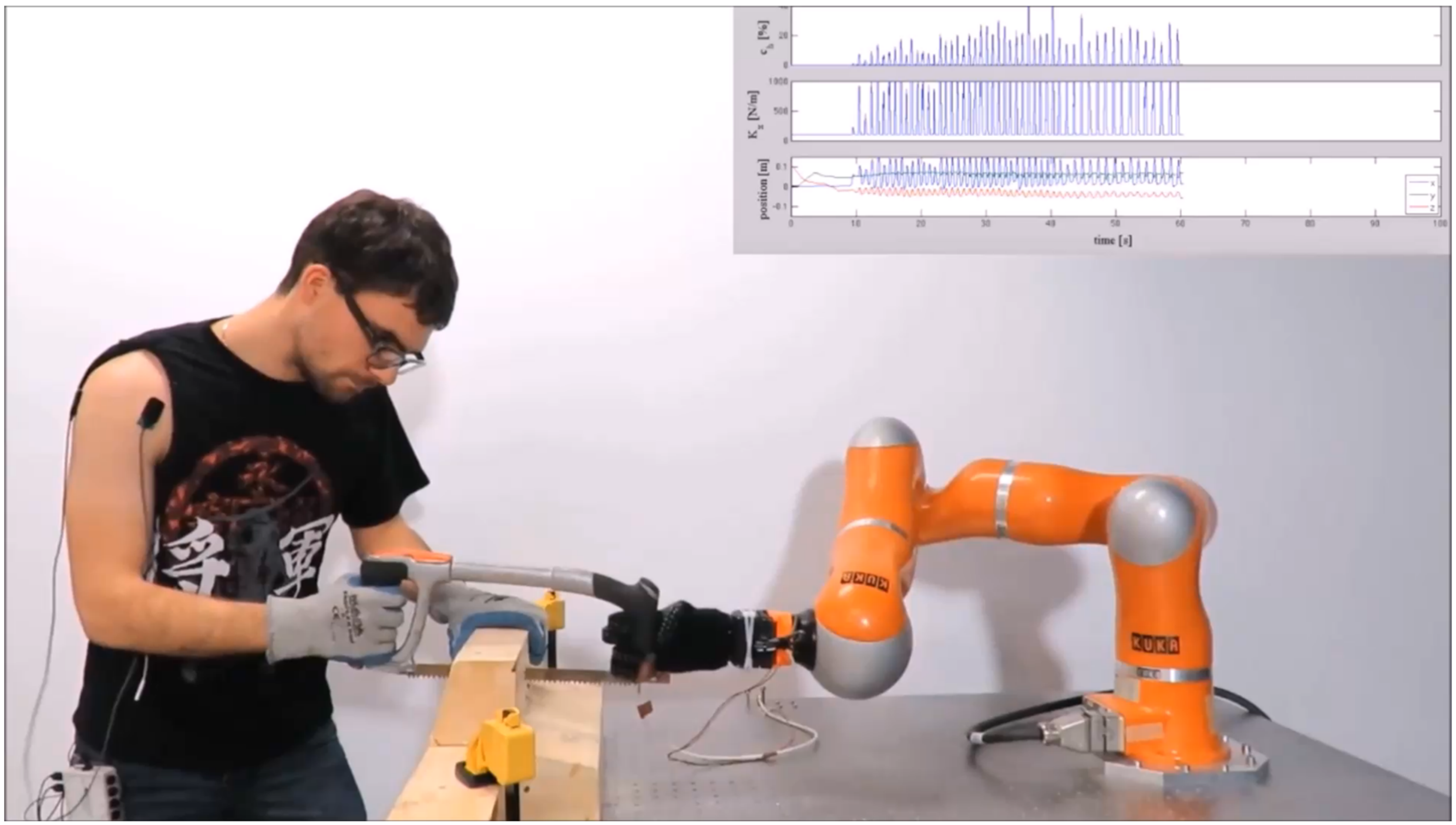
Service Robotics

- Hand guidance, co-manipulation:



Service Robotics

- Co-manipulation:
  - Physical interaction, safety, learning, adaptation



IIT [Peternel et al. 2016]

Collaborative Robotics

- **Four main categories:**
  - **Safety-rated monitored stop**
    - Robot stops when operator enters the collaborative workspace and continues when the operator has left the collaborative workspace. No physical separation is required (fences).
  - **Hand guiding**
    - Robot movements are directly guided and controlled by the operator.
  - **Speed and separation monitoring**
    - The contact between operator and moving robot is prevented by the robot by controlling the speed as a function of the separation distance.
  - **Power and force limiting**
    - Contact forces between operator and robot are technically limited to a safe level.

These basic principles of protection in HRC are in the standard EN ISO 10218 “Robots and robotic devices - Safety requirements for industrial robots” Parts 1 and 2. ISO/TS 15066 provides details

Collaborative Robotics

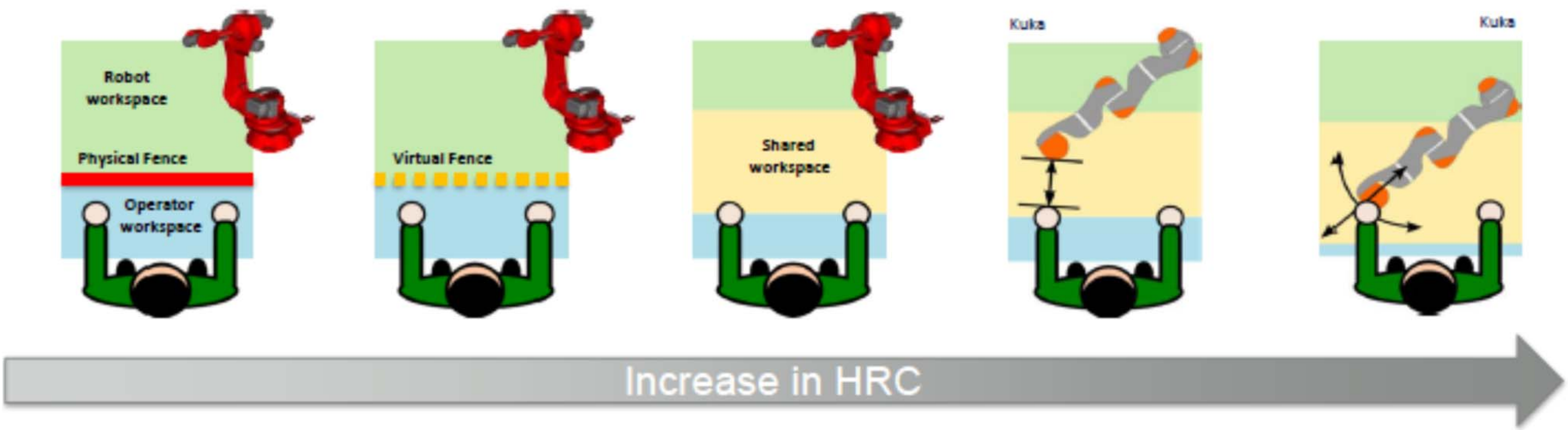
Contact not possible
→ Fixed safety fence

SMS (Safety Monitored Stop)
Contact not desired
No need to have a physical protection
→No fixed guard, **virtual** safety fence

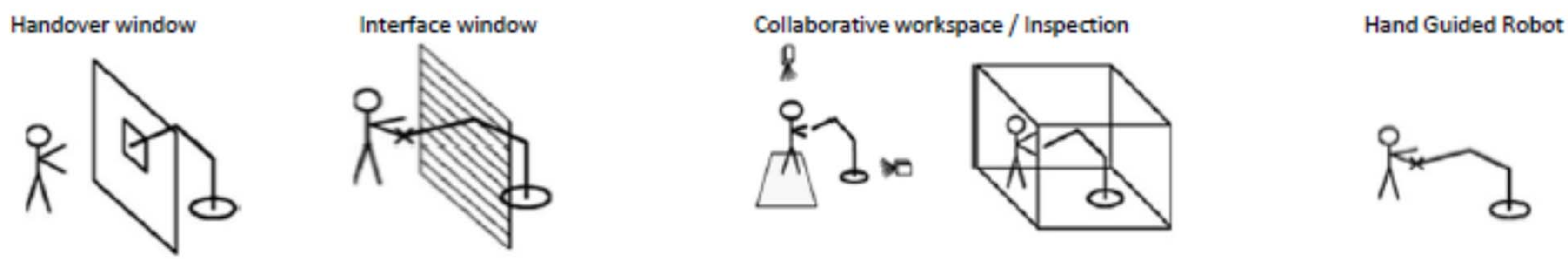
SMS Need of shared space.
Contact accepted only with stationary robot
→Shared workspace, but exclusive motion

SSM (Speed Separation Monitoring) or SMS
Need of shared space.
Contact not desired, but possible
→ Shared workspace

PFL (Power Force Limiting) or HG (Hand Guiding)
Contact desired
→ Shared workspace
, simultaneous motion
e.g. manual guidance



Collaboration examples according to EN ISO 10218-2, ANNEX E



Collaborative Robotics

Origin

moving elements

Potential consequences

- crushing
- impact
- shearing

Free collision

- Dynamic due to high velocities
- Dependent on velocity, mass and shape (of robot, gripper and manipulated object)
- Risks from the end-effector and the transported part
- Collision detection must function reliably and in time (before consequential damage)

Crushing

- No high velocities, virtually static
- Max. force must be parameterizable and not exceeded

Self activation

- External approval for activation required
- Light advices prior to motion

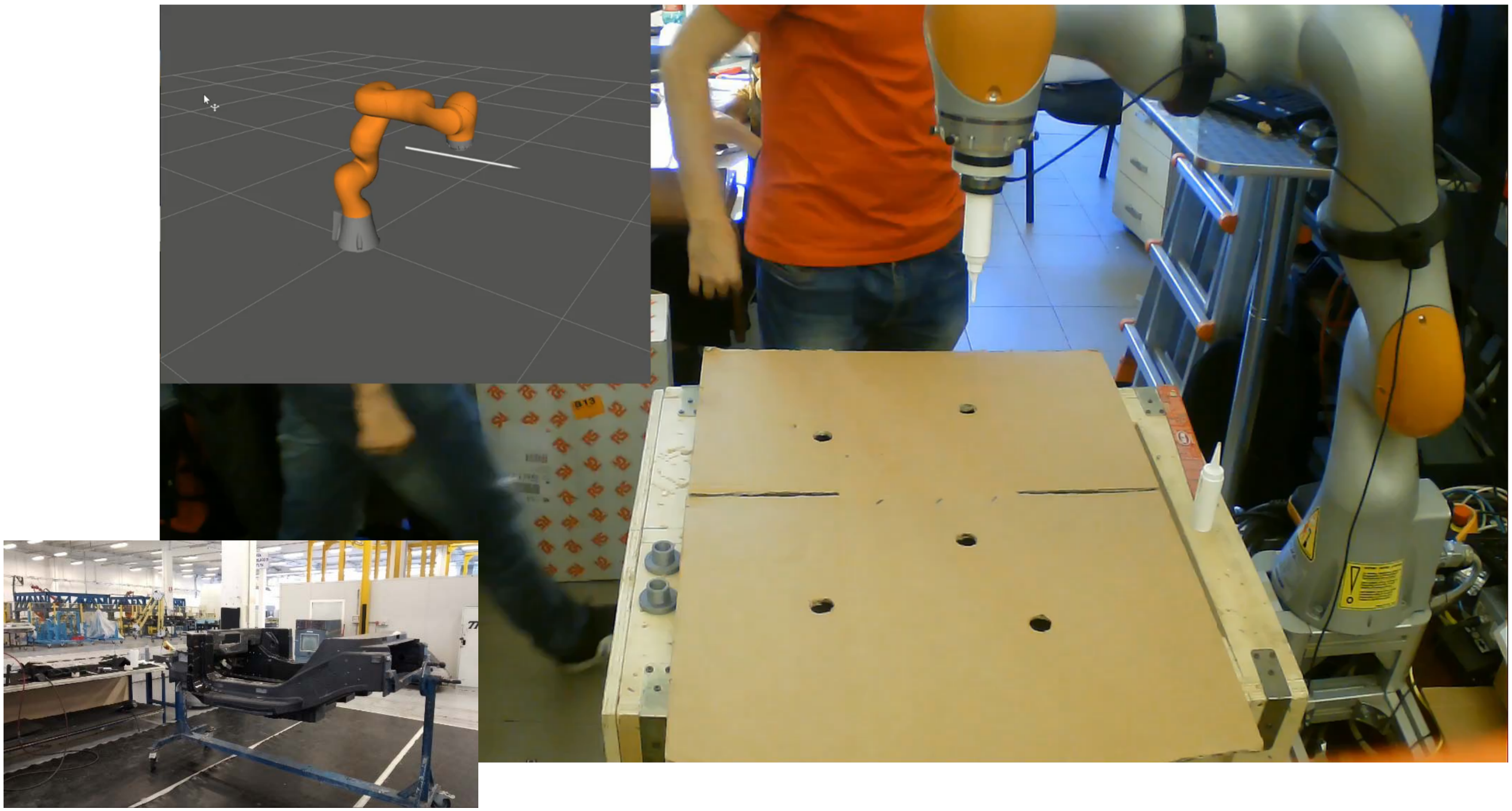
CASE STUDY: ICOSAF project

Integrated and collaborative systems for the intelligent factory
PON – Fondi Strutturali EU (2018-2021)

- Hand-guidance and co-manipulation:

ICOSAF

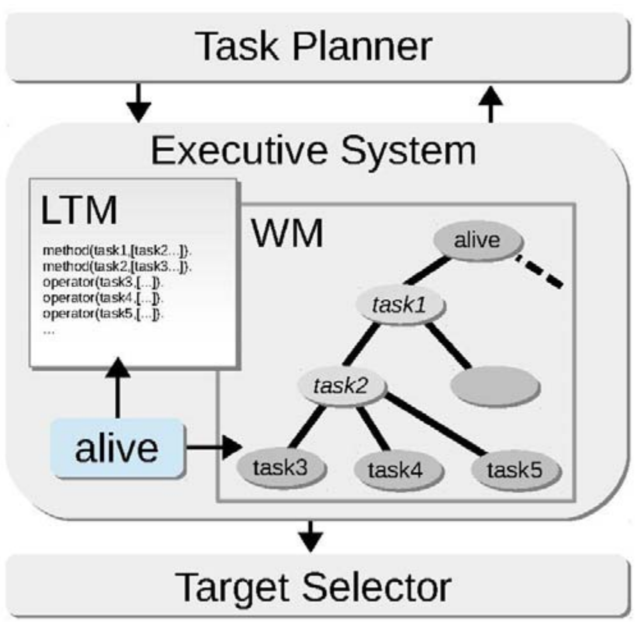
pon 2014-2020 FONDI STRUTTURALI EUROPEI

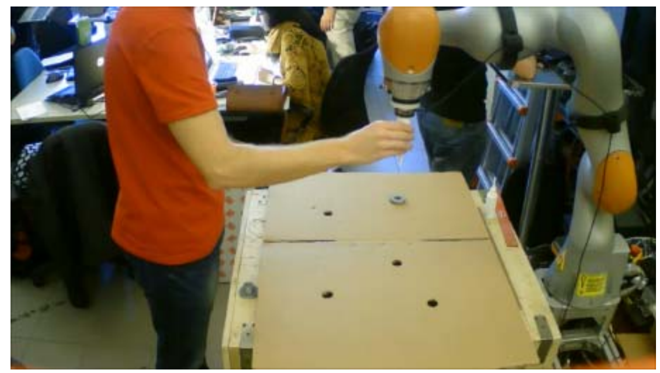


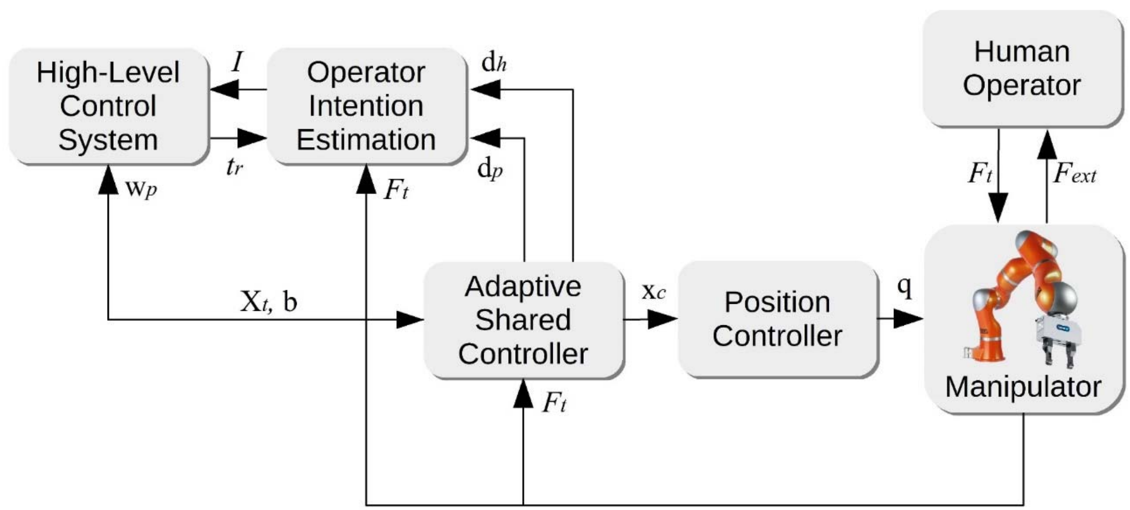
Co-manipulation and Physical Guidance

J. Cacace, R. Caccavale, A. Finzi, V. Lippiello. Interactive plan execution during human-robot cooperative manipulation. IFAC-PapersOnLine 51 (22), 500-505, 2018
J Cacace, A Finzi, V Lippiello . Enhancing Shared Control via Contact Force Classification in Human-Robot Cooperative Task Execution. Human Friendly Robotics, 167-179, 2018

- Integrated Human-Robot (Physical) Interaction Framework:
  - High Level Control, Intention Estimation, Adaptive Shared Controller

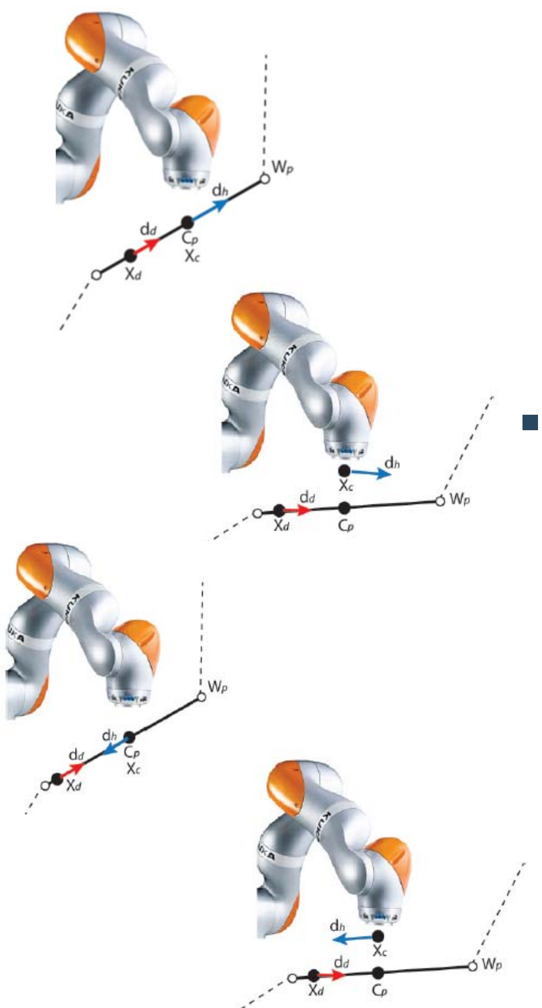




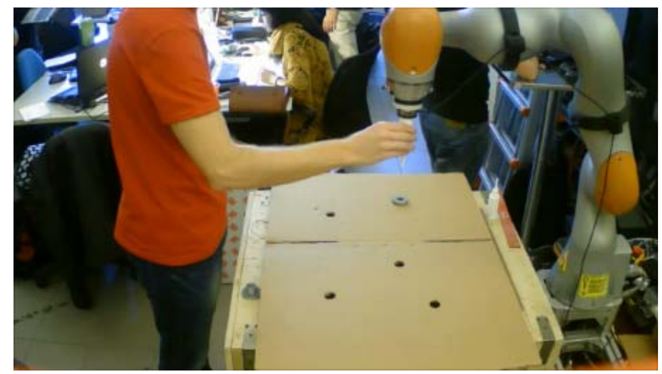


Contact Force Classification

J. Cacace, R. Caccavale, A. Finzi, V. Lippiello. Interactive plan execution during human-robot cooperative manipulation. IFAC-PapersOnLine 51 (22), 500-505, 2018
J Cacace, A Finzi, V Lippiello . Enhancing Shared Control via Contact Force Classification in Human-Robot Cooperative Task Execution. Human Friendly Robotics, 167-179, 2018



- Assess the intention of the operator physical guidance with respect to the robotic target and trajectory:
  - Coinciding (C).
  - Deviating (D).
  - Opposite (O).
  - Opposite Deviation (OD).



- Artificial Neural Network:
  - 3-layer NN
    - 3 input layers
    - 25 hidden nodes
    - 4 output layers
    - ~80%
  - LSTM
    - 3 input nodes
    - 16 hidden nodes
    - 4 outputs
    - ~86%

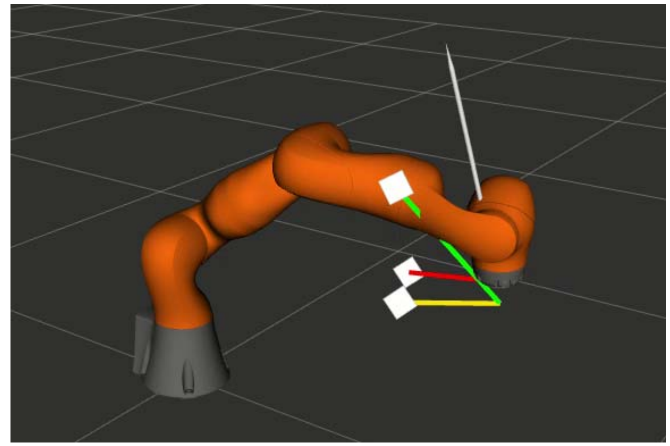
Network inputs:

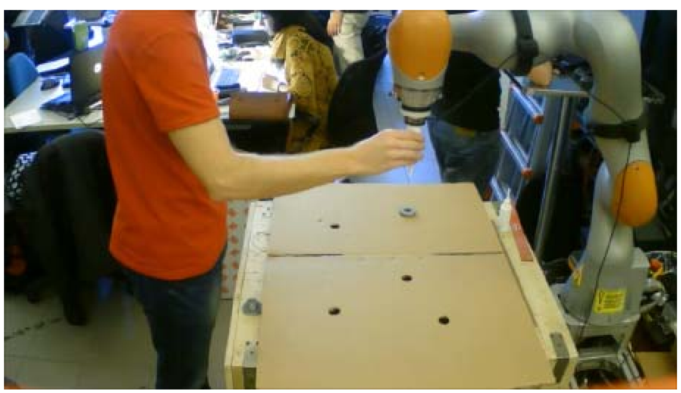
- $||F_t||$: magnitude of the force
- $|X_c - C_p|$: distance between current pos. of end eff. and closest point of the trajectory
- $\overrightarrow{d_c}$^$\overrightarrow{d_d}$: deviation from the planned path (angle between the two movements)

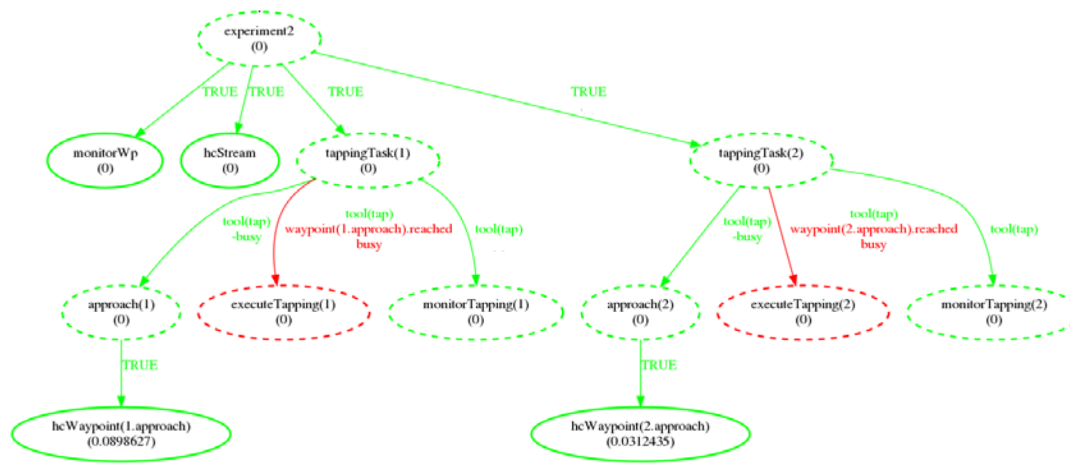
Target Selection

J. Cacace, R. Caccavale, A. Finzi, V. Lippiello. Interactive plan execution during human-robot cooperative manipulation. IFAC-PapersOnLine 51 (22), 500-505, 2018

- Target Selection:
  - Selected Target:
    - Only one target enabled with <u>Coinciding</u> or <u>Deviating</u>
    - No selection otherwise
  - Execution Mode:
    - Estimated intention if target selected (Coinciding or Deviating)
    - Passive otherwise

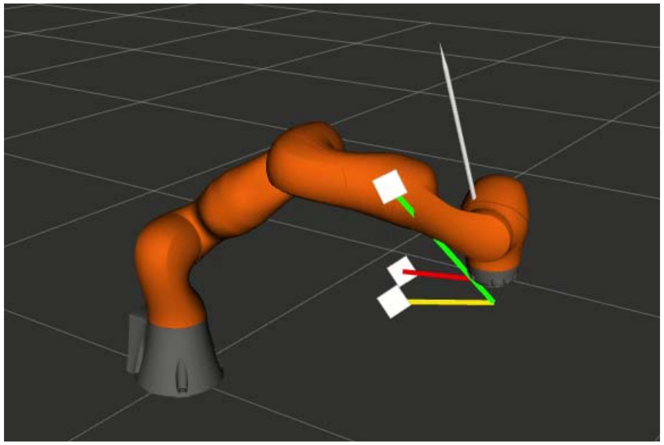




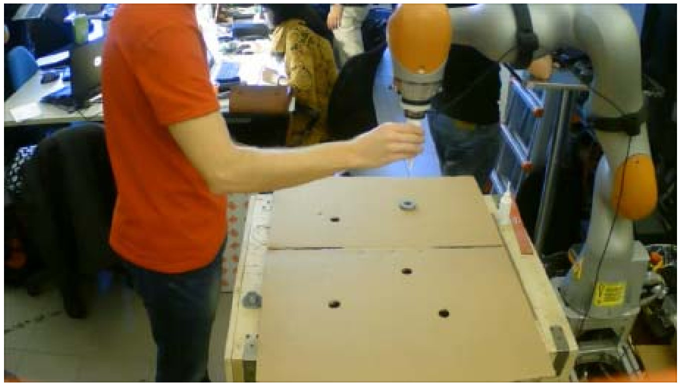


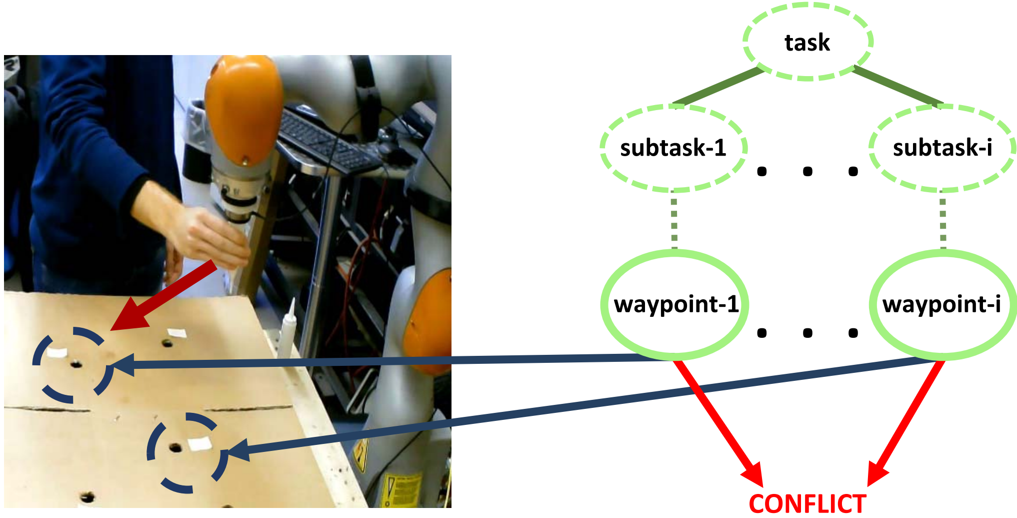
Target Selection

J. Cacace, R. Caccavale, A. Finzi, V. Lippiello. Interactive plan execution during human-robot cooperative manipulation. IFAC-PapersOnLine 51 (22), 500-505, 2018



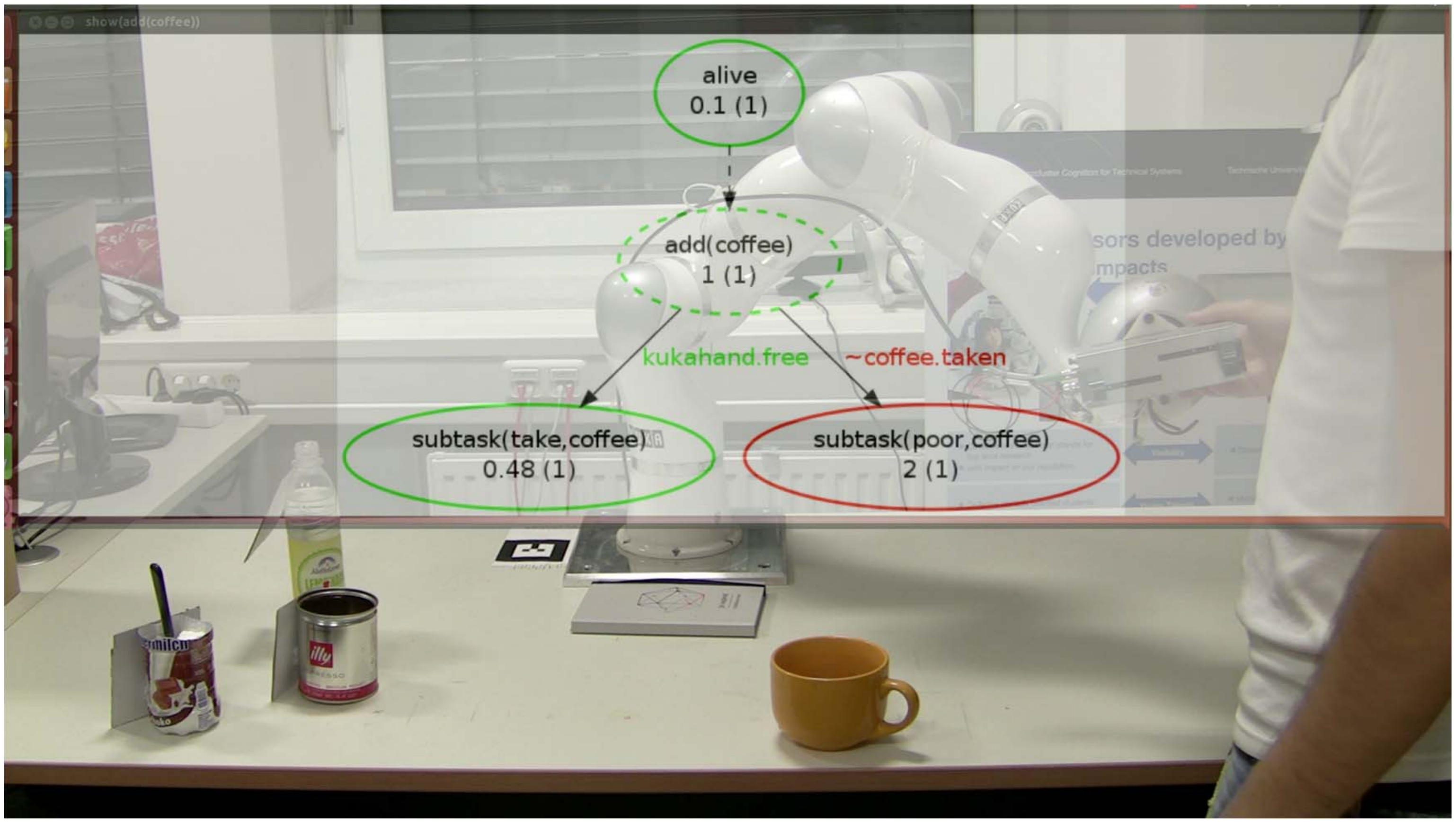
- Target Selection:
  - Selected Target:
    - Only one target enabled with Coinciding or Deviating
    - No selection otherwise
  - Execution Mode:
    - Estimated intention if target selected (Coinciding or Deviating)
    - Passive otherwise





Service Robotics

- Kinesthetic Teaching
  - Physical and cognitive interaction, attention, multimodal



Unia-TUM [Caccavale et al. 2018]

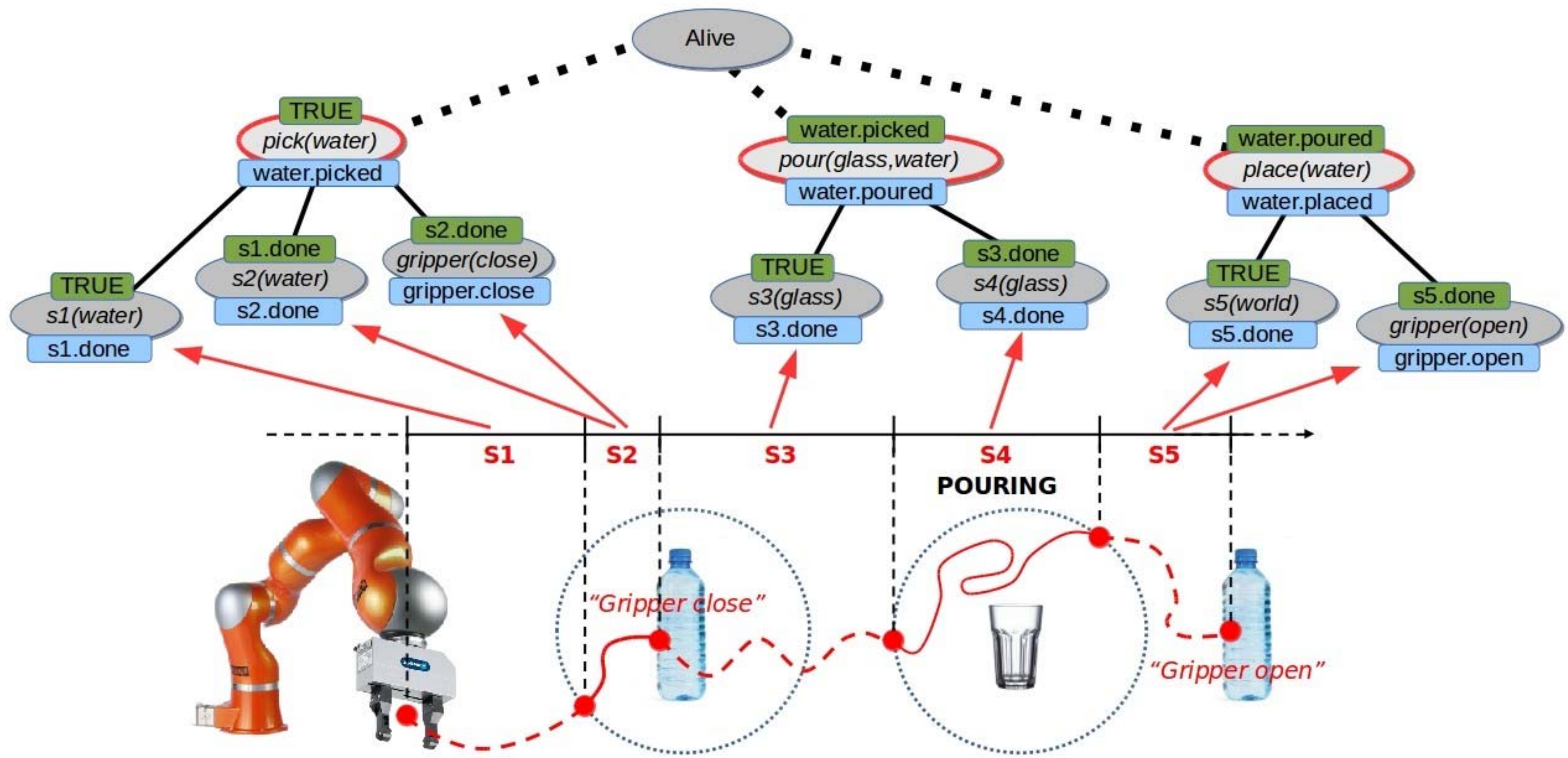
Kinesthetic Teaching

R. Caccavale, M. Saveriano, A. Finzi, D. Lee. Kinesthetic teaching and attentional supervision of structured tasks in human-robot interaction. Auton. Robots 43(6): 1291-1307 (2018)

Segments linked to the active and most emphasized subtasks

- **Far-Object-Action (FOA):** executed with a point-to-point motion (less accurate).
- **Near-Object-Action (NOA):** executed with Dynamic Movement Primitives (more accurate) [Ijspeert et al. 2013].

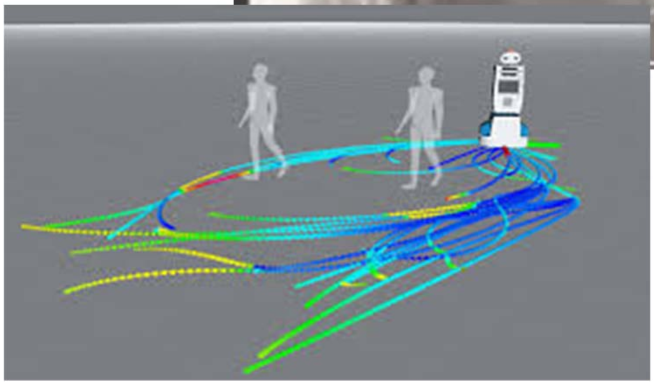
The **attentional system** monitors both the human and the robot motions, associating new segments to the *open subtasks* in WM



Service/Social Robotics

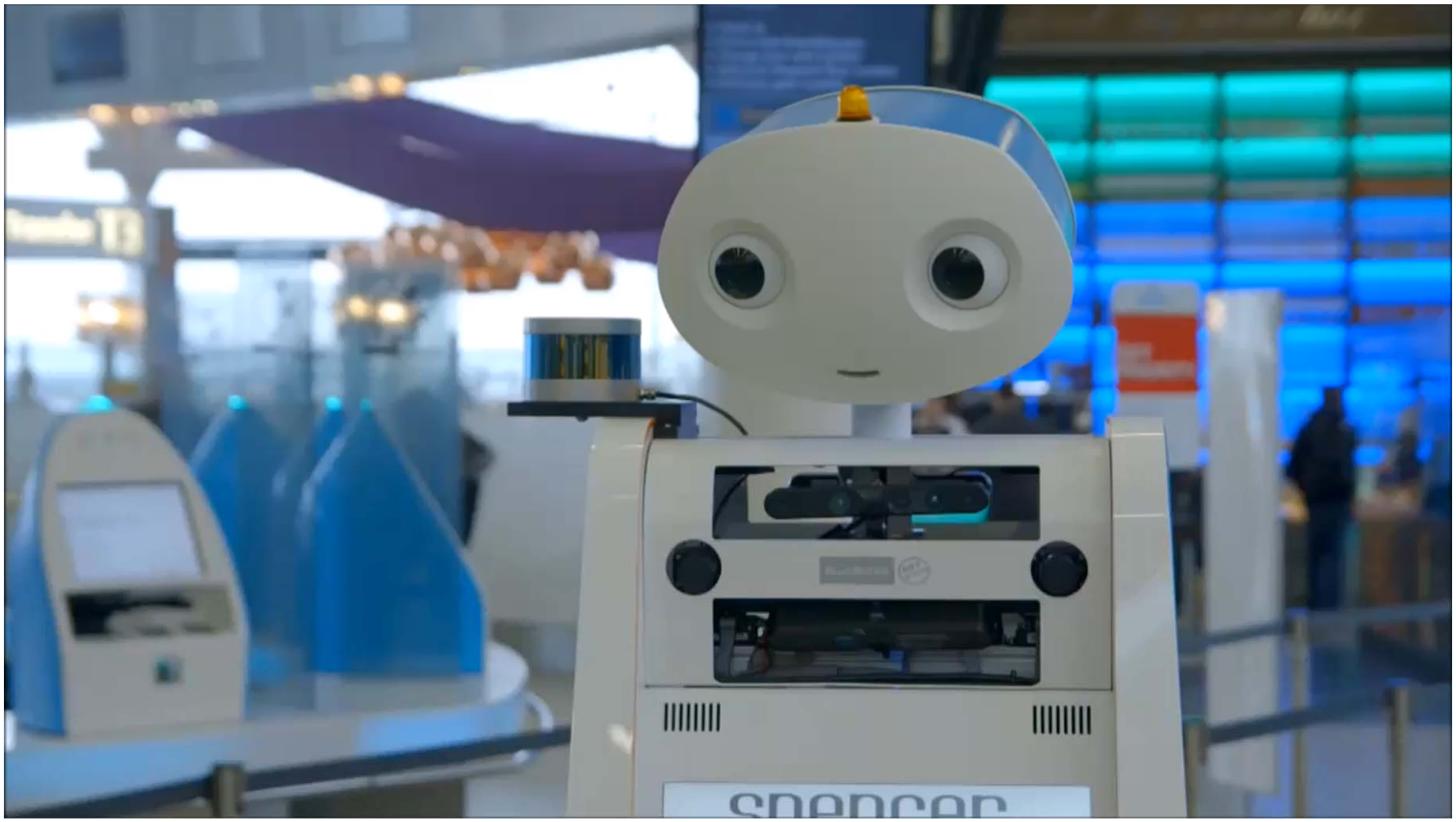
- Navigation in populated environments
  - Safety
  - Human anticipation
  - Robot action readability
  - Proxemics
  - Social constraints

Minerva



Service/Social Robotics

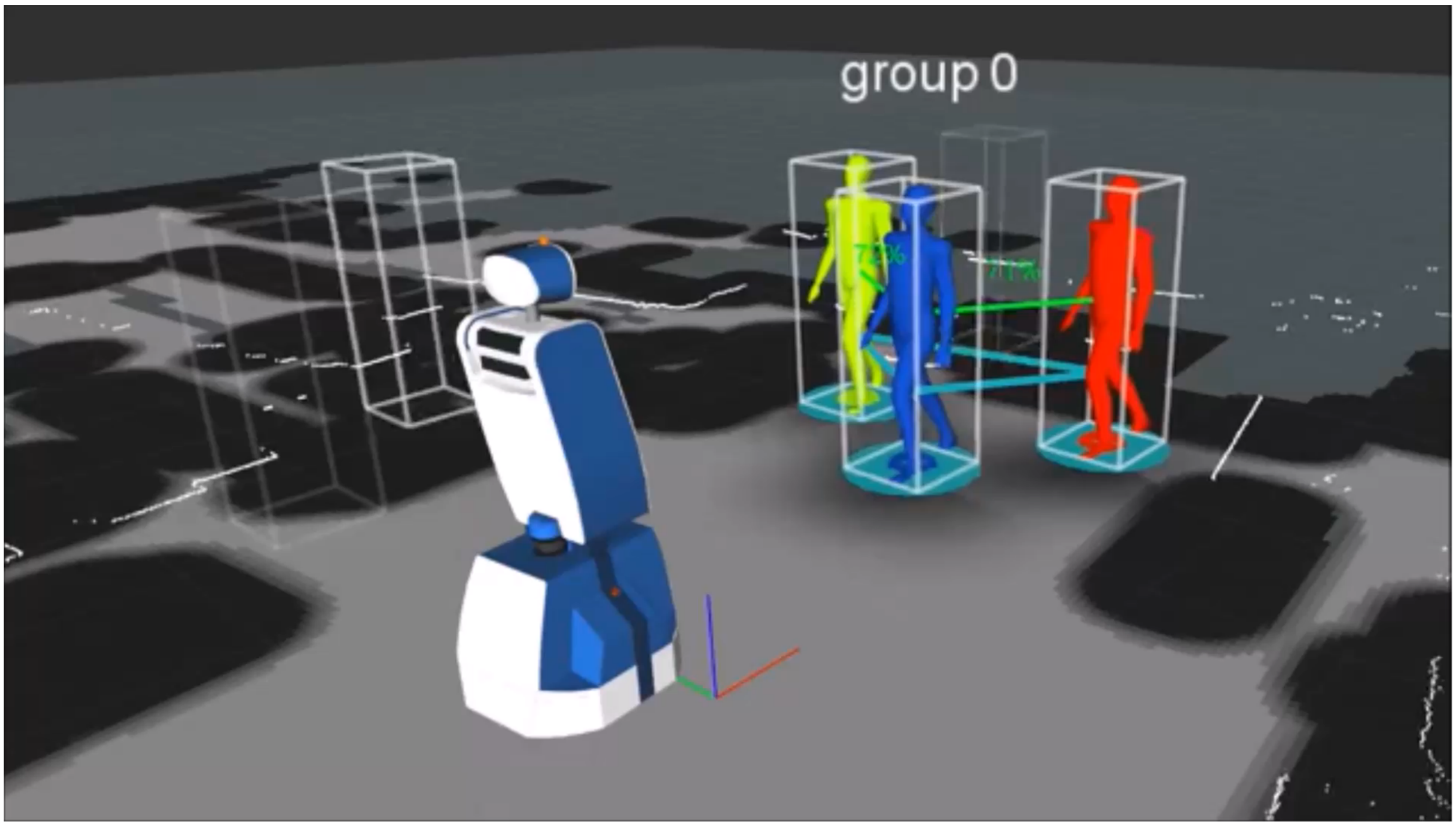
- Navigation in populated environments



Social situation-aware perception and action for cognitive robots (SPENCER) FP7

Service/Social Robotics

- Navigation in populated environments



Service/Social Robotics

- Navigation in populated environments
- Challenges [Kruse et al. 2013]
  - Comfort: absence of annoyance and stress for humans interaction with robots
  - Naturalness: similarity between robots and humans in low-level behavior patterns
  - Sociability: adherence to explicit cultural conventions
  - Interaction:
    - Follow a person, guide, navigate in dense crowd, move in formation, etc.

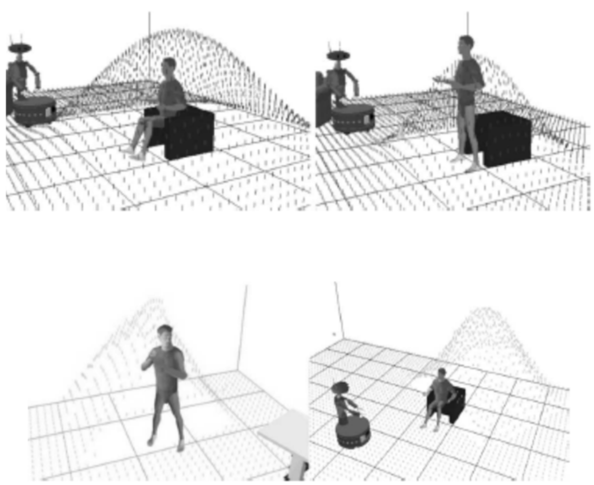
Service/Social Robotics

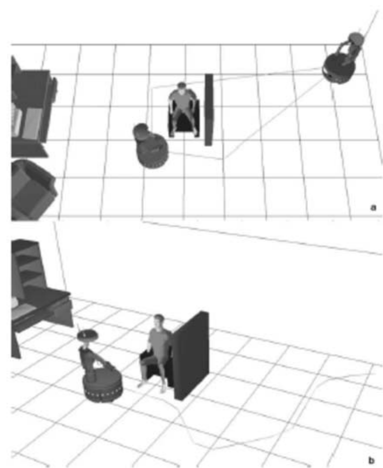
- Navigation in populated environments
- Methods:
  - Human monitoring and predictions (motion models)
  - Social activity assessment (groups)
  - Cost functions (cost of disturbance for humans)
  - Path planning (social constraints, proxemics, cost functions)
  - Decision Theory Planning
  - Strategic Reasoning

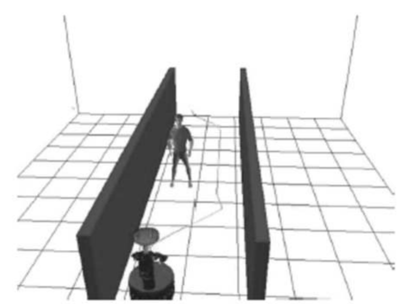
Service/Social Robotics

- Navigation in populated environments
- Methods:
  - Cost functions (cost of disturbance for humans)
  - Path planning (social constraints, proxemics, cost functions)

$$
\begin{aligned}
Cost_{\text{merged}}(x, y) &= w_1\, Cost_{\text{safety}}(x, y) \\
&+ w_2\, Cost_{\text{visibility}}(x, y)
\end{aligned}
$$







A Human Aware Mobile Robot Motion Planner, Sisbot et al. 2007

Service/Social Robotics

- Navigation in populated environments
- Methods:
  - Human monitoring and predictions (motion models)
  - Path planning (social constraints, proxemics, cost functions)
  - Decision Theory Planning
    - autonomous vehicles must estimate unknown pedestrian intentions and hedge against the uncertainty in intention estimates in order to choose actions that are effective and robust
    - POMDP planning in near real time to control the speed of the vehicle along the planned path
    - A* to plan the path
    - Pedestrian Modeling $(x_i, y_i), g_i, v_i$ probabilistic dynamic process $p\left(x'_i, y'_i \middle| x_i, y_i, g_i, v_i\right)$
    - Reward model

Intention-Aware Online POMDP Planning for Autonomous Driving in a Crowd, Bai et al. 2015

Autonomy

- Automotive:
  - Driverless car
    - Mapping, localization, vision/LiDAR-based detection, avoidance, path planning, navigation, decision making, etc.

  - Assisted Driving
    - Autopilot, cruise control, vision-based detection, avoidance, alerting, HMI, etc.

Autonomy

- Autonomy level
  - The amount of interventions required for controlling a robot
  - the percentage of time that a human operator must be controlling the robot (e.g. AUTONOMY=0% and INTERVENTION=100%, AUTONOMY=75% and INTERVENTION=25%, etc.)
  - adjustable autonomy, sliding scale autonomy and mixed initiative
  - Human operators may wish to override the robot’s decisions, or the robot may need to take over additional control during a loss of communications

Autonomy

- Automotive:
  - **Level 0:** The driver controls it all: steering, brakes, throttle, power.
  - **Level 1: (**”hands on”**)** Most of the functions are still controlled by the driver, but a specific function (like steering or accelerating) can be done automatically by the car
  - **Level 2:** (”hands off”) full control of the vehicle (accelerating, braking, and steering). The driver must monitor the driving and be prepared to immediately intervene at any time if the automated.
  - **Level 3:** (”eyes off”): The driver can safely turn their attention away from the driving tasks, e.g. the driver can text or watch a movie.
  - **Level 4:** (”mind off”): No driver attention is ever required for safety, i.e. the driver may safely go to sleep or leave the driver's seat. The system is "designed to perform all safety-critical driving functions and monitor roadway conditions for an entire trip." Limited to the "operational design domain (ODD)".
  - **Level 5: (**”wheel optional”): No human intervention is required (robotic taxi).

Autonomy

- Planning for Autonomous Cars that Leverage Effects on Human Actions [Sadigh et. al 2016]
  - robot's actions have immediate consequences on the state of the car, but also on human actions
  - human as an optimal planner, with a reward function acquired through Inverse Reinforcement Learning.

$$x' = f_{\mathcal{R}}(x, u_{\mathcal{R}})$$

$$x'' = f_{\mathcal{H}}(x', u_{\mathcal{H}})$$

$$x^{t+1} = f_{\mathcal{H}}(f_{\mathcal{R}}(x^t, u^t_{\mathcal{R}}), u^t_{\mathcal{H}}) \qquad r_{\mathcal{R}}(x^t, u^t_{\mathcal{R}}, u^t_{\mathcal{H}})$$

The robot will use Model Predictive Control (MPC) at every iteration, it will compute a finite horizon sequence of actions to maximize its reward. It will then execute the first one, and replan

$$\mathbf{x} = (x^1, \ldots, x^N)^\top \qquad \mathbf{u}_{\mathcal{H}} = (u^1_{\mathcal{H}}, \ldots, u^N_{\mathcal{H}})^\top \qquad R_{\mathcal{R}}(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}) = \sum_{t=1}^{N} r_{\mathcal{R}}(x^t, u^t_{\mathcal{R}}, u^t_{\mathcal{H}}) \qquad \mathbf{u}^*_{\mathcal{R}} = \arg\max_{\mathbf{u}_{\mathcal{R}}} R_{\mathcal{R}}(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}^*_{\mathcal{H}}(x^0, \mathbf{u}_{\mathcal{R}}))$$

$$\begin{aligned} \mathbf{u}^{*t}_{\mathcal{H}}(x^0, \mathbf{u}_{\mathcal{R}}) &= \mathbf{u}^{*t}_{\mathcal{H}}(\mathbf{x}^0, u^{0:t}_{\mathcal{R}}, u^{*0:t-1}_{\mathcal{H}}) \\ &= \arg \max_{\mathbf{u}^{t:t+N-1}_{\mathcal{H}}} r_{\mathcal{H}}(x^t, u^t_{\mathcal{R}}, u^t_{\mathcal{H}}) + \\ &\quad \sum_{i=t+1:t+N-1} r_{\mathcal{H}}(x^i, \tilde{u}^i_{\mathcal{R}}, u^i_{\mathcal{H}}) \end{aligned}$$

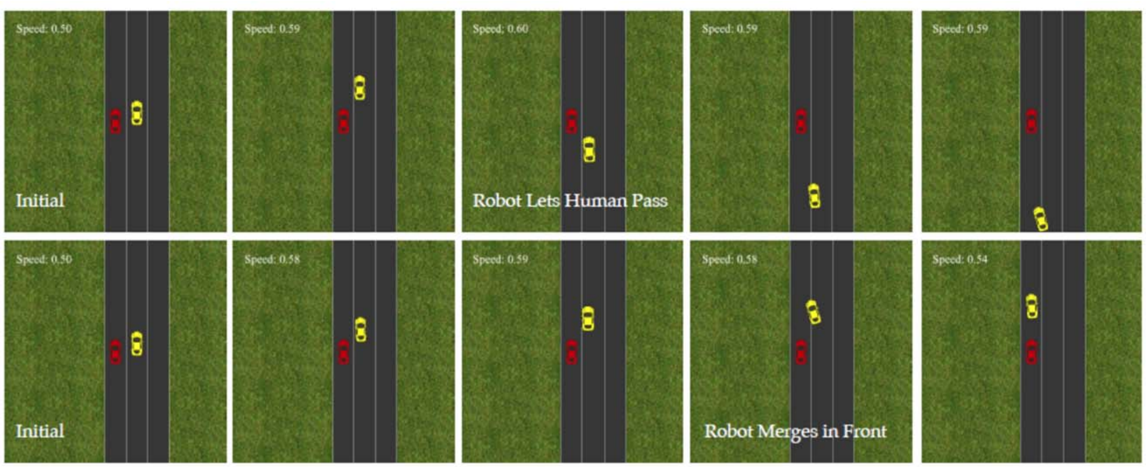
Approx.

$$\mathbf{u}^*_{\mathcal{H}}(x^0, \mathbf{u}_{\mathcal{R}}) = \arg\max_{\mathbf{u}_{\mathcal{H}}} R_{\mathcal{H}}(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}})$$

Inverse RL to learn R_H

Autonomy

- Planning for Autonomous Cars that Leverage Effects on Human Actions [Sadigh et. al 2016]
  - robot's actions have immediate consequences on the state of the car, but also on human actions
  - human as an optimal planner, with a reward function acquired through Inverse Reinforcement Learning.



The autonomous vehicle's goal is to reach a final point in the left lane. In the top scenario, the autonomous vehicle has a simple model of the human driver that does not account for the influence of its actions on the human actions, so it acts more defensively, waiting for the human to pass first. In the bottom, the autonomous vehicle uses the learned model of the human driver, so it acts more aggressively and reaches its goal faster.

Metrics for HRI

- *Fan out:* [Goodrich 2003]
  - how many robots can be controlled by a human
- *Intervention response time:*
  - (1) time to deliver the request from the robot,
  - (2) time for the operator to notice the request,
  - (3) situation awareness and planning time,
  - (4) execution time.
- *Level of autonomy discrepancies:*
  - multiple levels of control and autonom

Metrics for HRI

- Quantitative evaluation:
  - Effectiveness: the percentage of the mission that was accomplished with the designed autonomy
  - Efficiency: the time required to complete a task.
- Subjective evaluation:
  - Quality of the effort
- Appropriate utilization of mixed-initiative:
  - Percentage of requests for assistance made by robot
  - Percentage of requests for assistance made by operator
  - Number of interruptions of operator rated as non-critical

Metrics for HRI

- Operator performance:
  - *Situation awareness*
  - *Workload*
  - *Accuracy of mental models of device operation*
- Robot performance:
  - *Self-awareness*
  - *Human-awareness*
  - *Autonomy*

Information Exchange

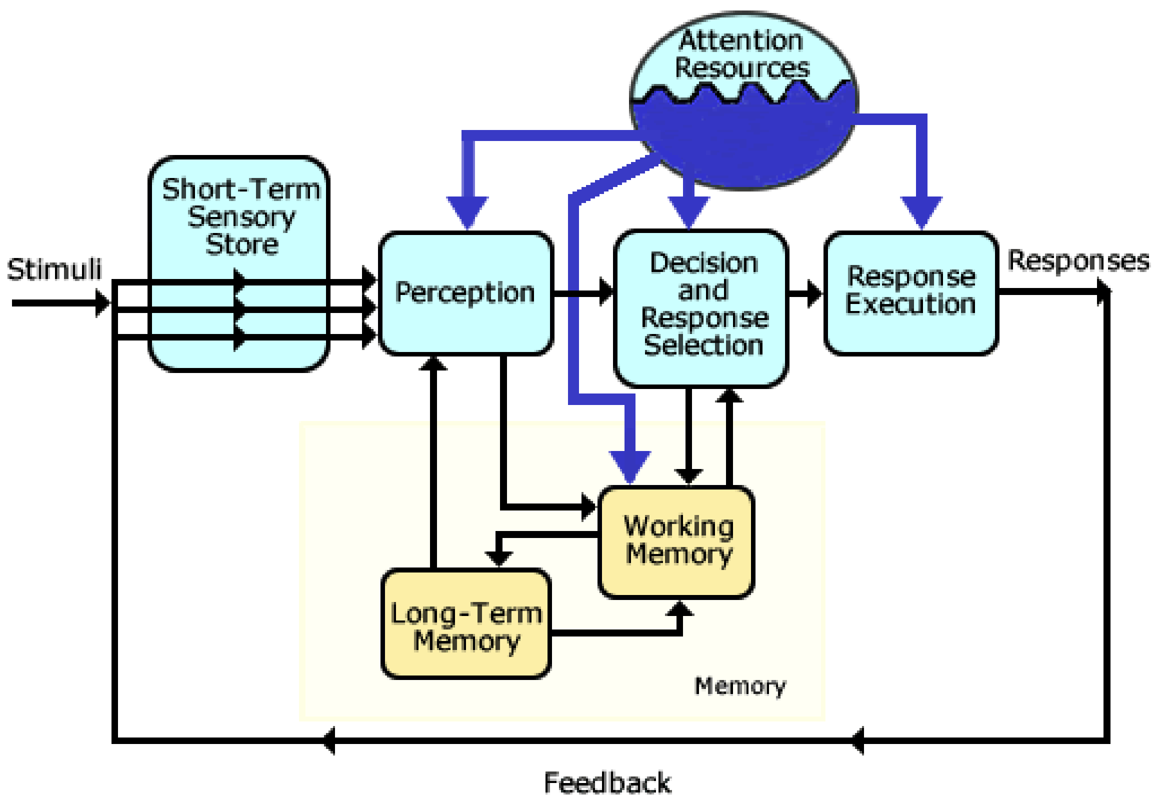
Intelligent interaction requires deliberate communication

- Interaction Time [Goodrich 2003]
- Switch attention to current task
- Establish context
- Plan actions
- Communicate plan to robot
- Workload
- Situational Awareness
- Shared Mental Model

Cognitive Workload

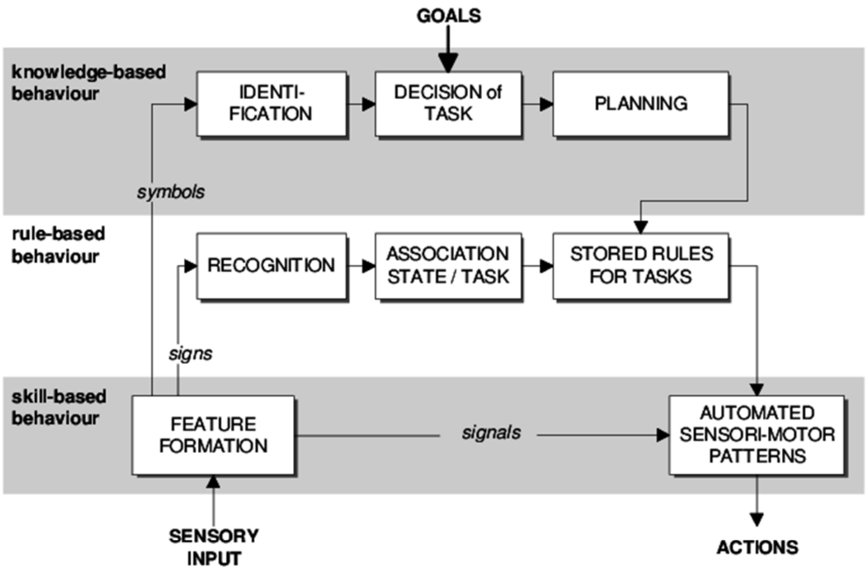
Model of human information processing [Wickens, 1992]:

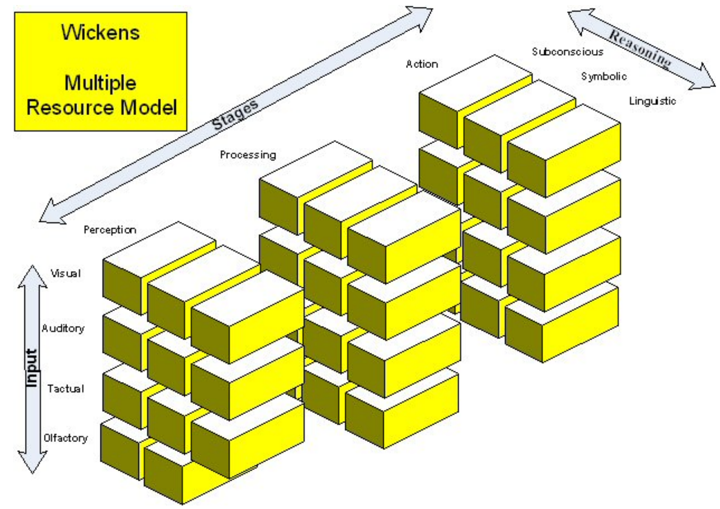
- Attention resources are limited.
- A resevoir with a finite amount of attentional resources.
- Attention determines what information is transmitted to working memory:
  - Expectancy and Saliency

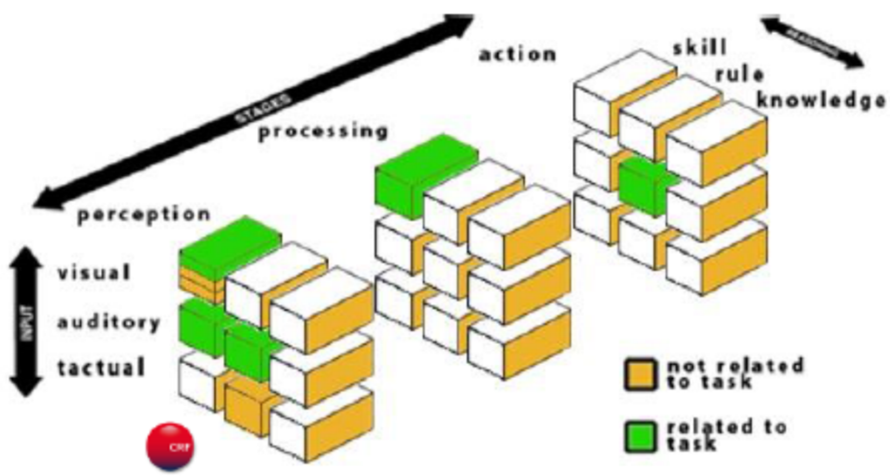


Cognitive Workload

Rasmussen + Wickens:







CASE STUDY: SHERPA Project

- Multimodal Human-Robot Interaction for Multi-UAVs Control in SHERPA





Coordinator: Prof. Marconi, UNIBO

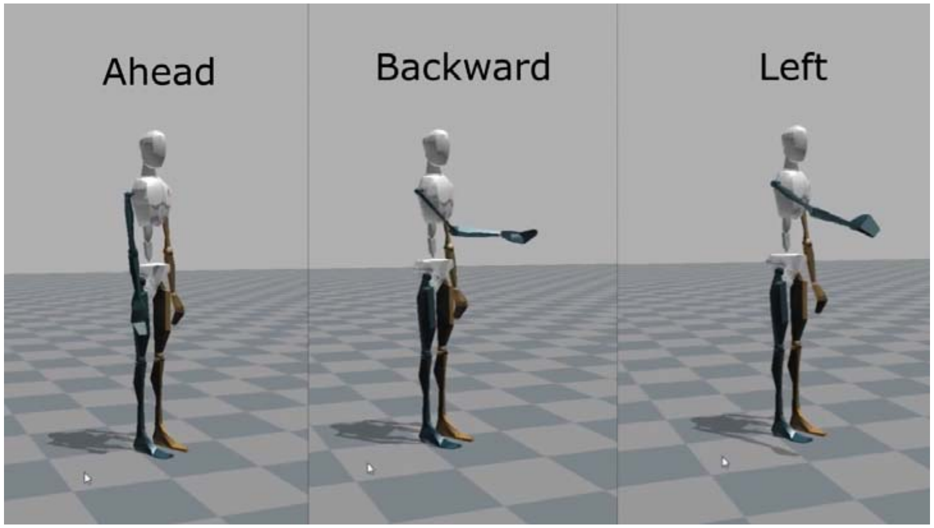
Mixed **ground** and **aerial** robotic platform to support **search and rescue** activities in the alpine scenario

The human rescuer is the "busy genius" working in team with the ground vehicle and with the aerial platforms"

"able to provide sketchy, though high-value, inputs towards the achievement of the team goal."

Multimodal Human Robot Interaction

- Speech recognition:
  - Bag of word model implemented via Pocketsphinx.
- 2D Gesture recognition:
  - Gesture detected on capacitive touch screen.
- 3D Gesture recognition:
  - Gesture detected considering the movement on the operator’s arm.
- Hand pose recognition:
  - Detection of the pose of the operator’s hand using EMG classification.



Human-Robot Interaction in the Alps

- SHERPA Davos demonstration



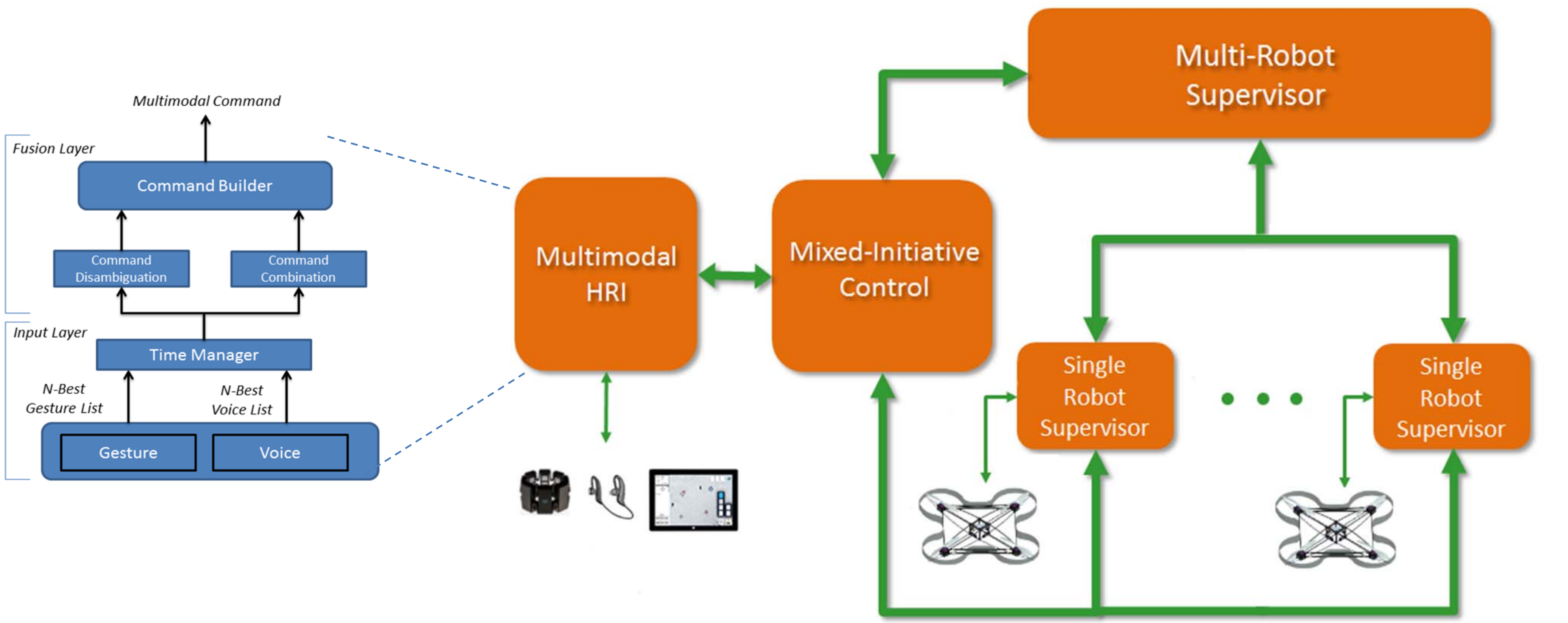

14 missions, ~15 min. for mission

Data	Value
Commands (#)	107
Generated Samples (#)	708
Success Rate (%)	96.8%
Rejected Samples (%)	66.9%
Correctly Rejected Samples (%)	74.3%



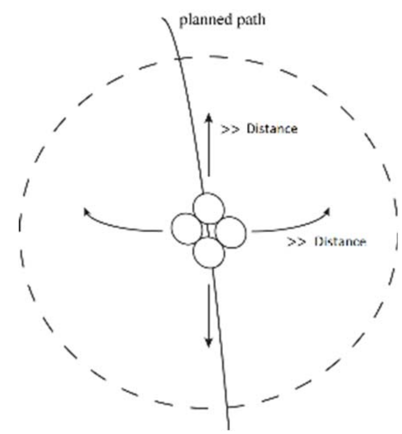
Multimodal Mixed-Initiative Interaction

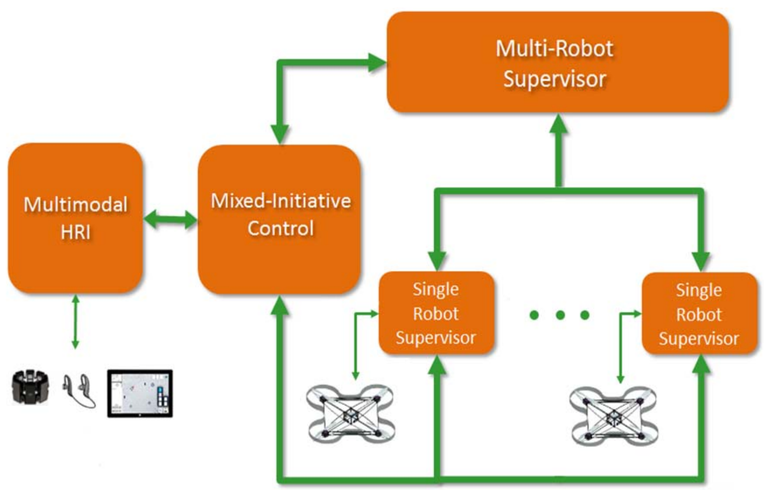
- Mixed Initiative Interaction:
  - Different modalities at different levels of astraction
    - Task, path, trajectory replanning

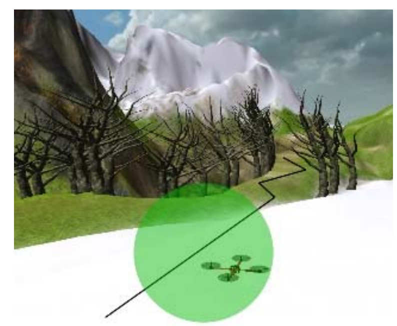


Multimodal Mixed-Initiative Interaction

- Mixed-initiative interaction with multiple drones:
  - Task level interaction
    - Mission-level task switching
  - Path and trajectory level interaction
    - Path/trajectory adjustament
    - Path/trajectory replanning

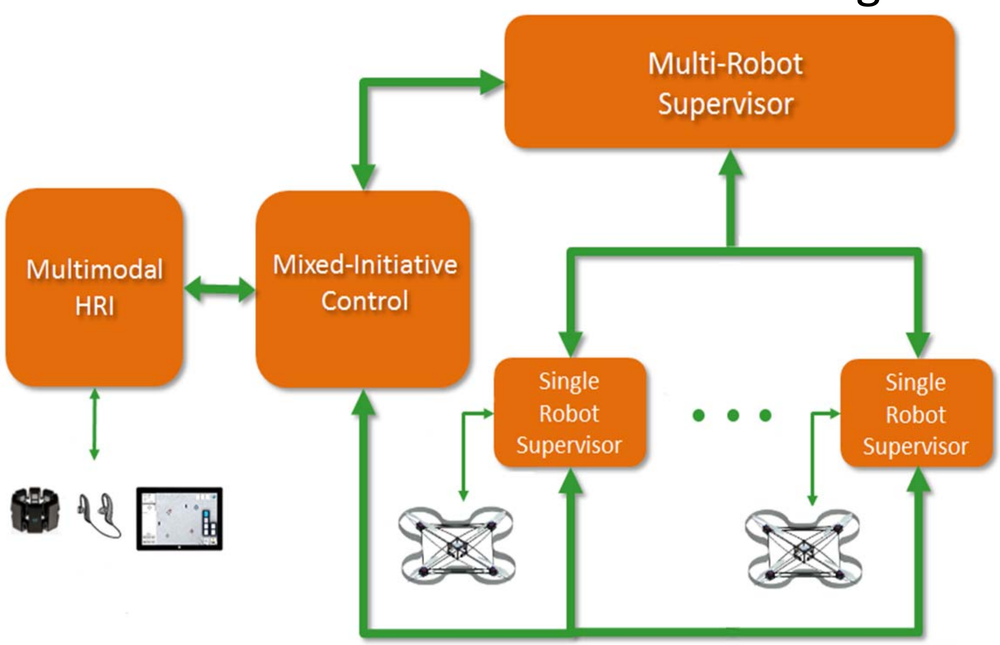




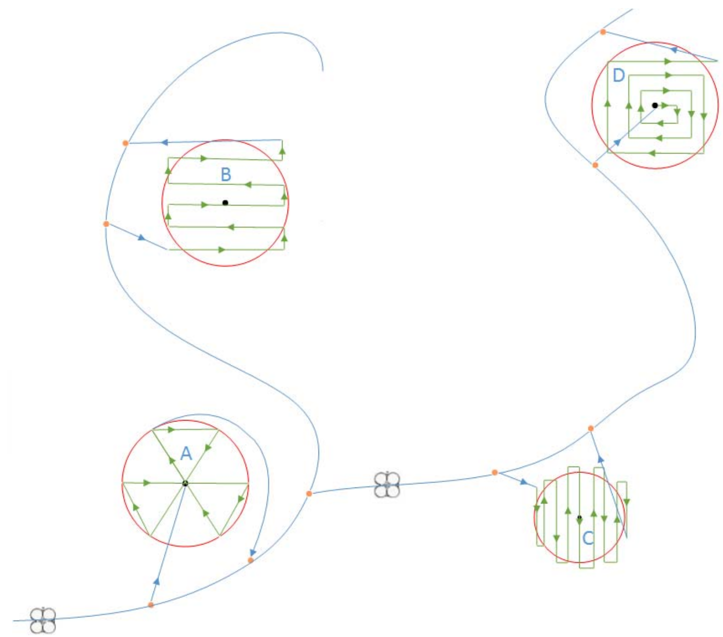


Multimodal Mixed-Initiative Interaction

- Mixed-initiative interaction with multiple remote drones in a simulated environment:
  - Task level interaction
    - Mission plan (e.g. explore a set of areas)
    - Mission-level task switching



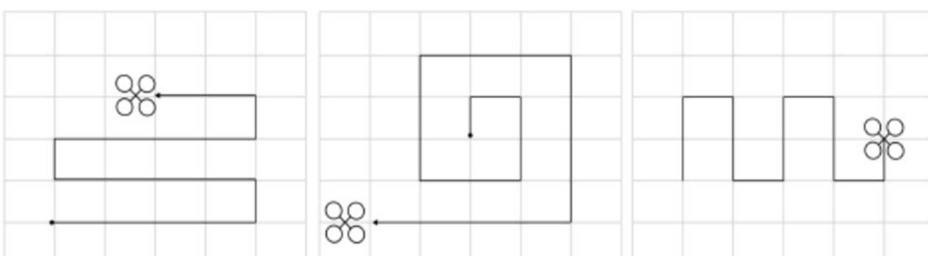
Architecture

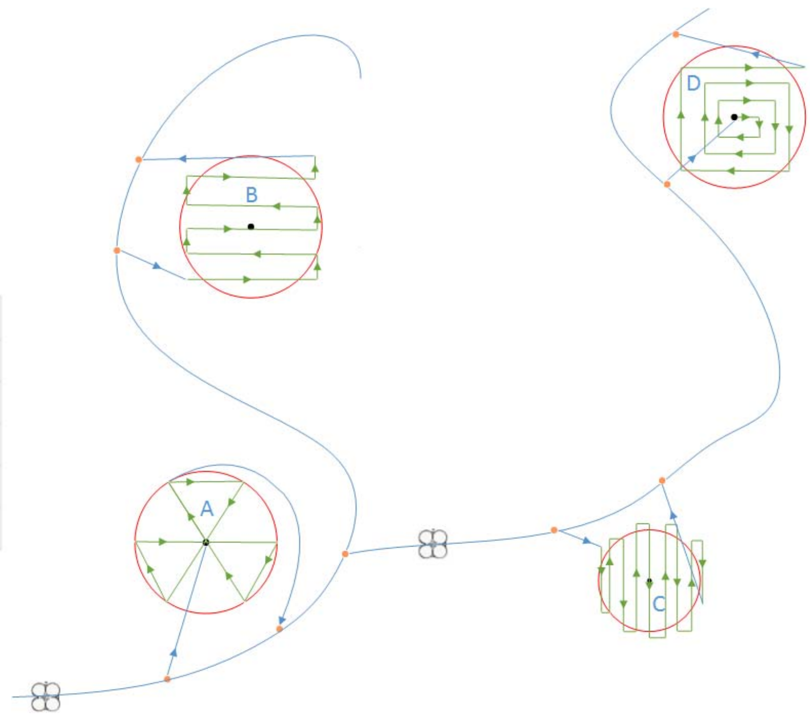


Search Strategy

Multimodal Mixed-Initiative Interaction

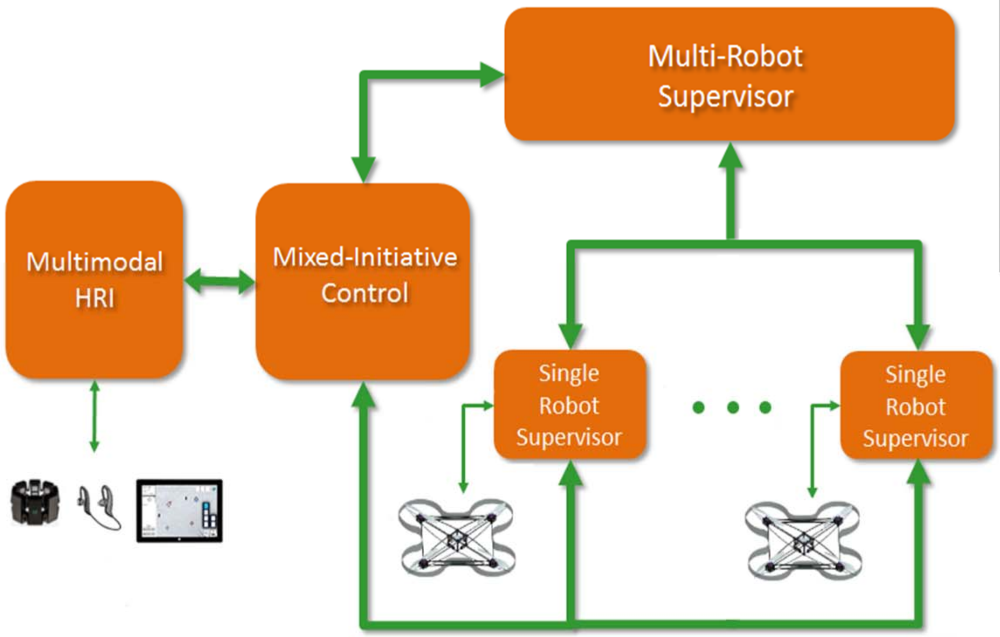
- Search and Rescue with UAVs
  - Guidelines for search and rescue [NATSAR 2011; CSAR-2000; NATO-1988]:
    - Search area for the targets
    - Sub-areas for assignment of search patterns
    - Assign specific search patterns to cover each sub-area
    - Sequence for the search patterns
    - Execute the patterns, marking the positions of the victims





Multimodal Mixed-Initiative Interaction

- Mixed-initiative interaction with multiple remote drones in a simulated environment:
  - Task level interaction
    - Mission-level tasks
    - Task Planning and Execution



Task	Description
SetSearchArea	Add/delete/modify a search area
SetSearchPath	Add/delete/modify a search path
SetSearchPattern	Change the exploration method
SearchPath	Search along the path
ScanArea	Scan the area with a pattern
ExploreMap	Explore the map
GoTo	Move towards a direction or an area
AbortMission	abort the overall mission

Mission Level Tasks

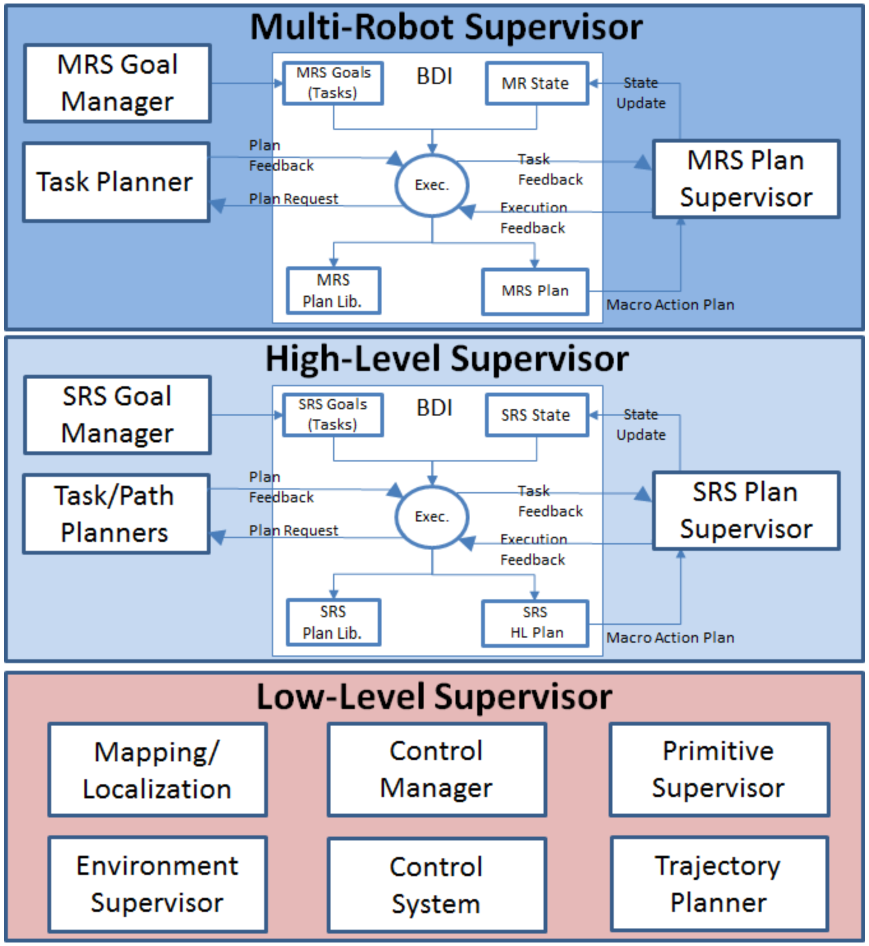
Incomplete and dynamic specification of an exploration task

Multimodal Mixed-Initiative Interaction

- Mixed-initiative interaction with multiple remote drones in a simulated environment:

Multi-Robot and Single-Robot Supervisors:
- BDI-based executive systems (PRS)
- Hierarchical task planner
- Mixed-initiative task planning
- Human interventions associated with replanning activities

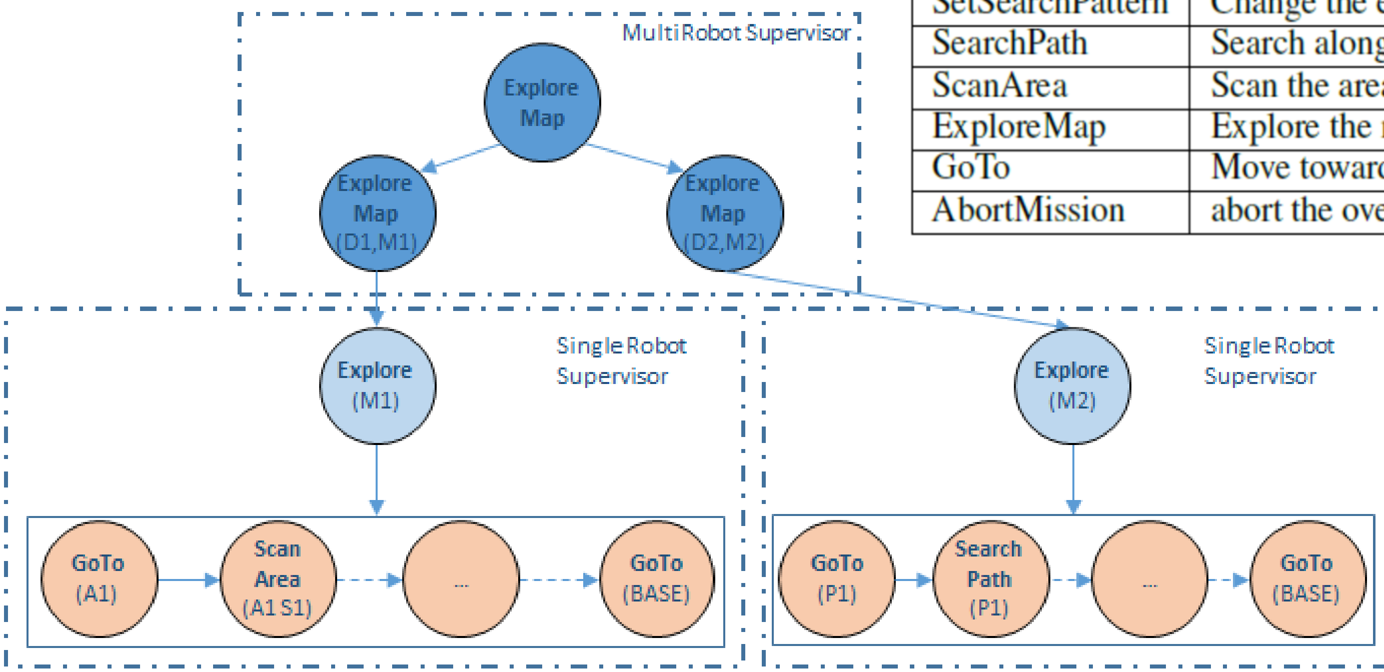
Internal system for testing



Multimodal Mixed-Initiative Interaction

- Mixed-initiative interaction with multiple remote drones in a simulated environment:
  - Hierarchical task planning
    - SHOP-like, Human-Aware (HATP)

Task	Description
SetSearchArea	Add/delete/modify a search area
SetSearchPath	Add/delete/modify a search path
SetSearchPattern	Change the exploration method
SearchPath	Search along the path
ScanArea	Scan the area with a pattern
ExploreMap	Explore the map
GoTo	Move towards a direction or an area
AbortMission	abort the overall mission



MRS task decomposition refined by the SRSs

Social Robotics

- Robotica per l’interazione sociale e la comunicazione:
  - Compagnia, intrattenimento, gioco, assistenza, monitoraggio
  - Interazione cognitiva e/o fisica
  - Comportamento sociale e proattivo
  - Cognizione ed emozione

Sota, Pepper and Jibo

Social Interaction

- Humans treat robots like people:
  - T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3-4):143–166, 2003
  - C. Nass and Y. Moon. Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1):81–103, 2000
- The more a robot interacts with people, the more life-like and intelligent it is perceived and the more excited users are
  - G. Schillaci, S. Bodiroza, and V. V. Hafner. Evaluating the effect of saliency detection and attention manipulation in human-robot interaction. *International Journal of Social Robotics*, 2012

Social Robotics

- Studio della relazione umana con i robot
- Tendenza a trattare i Robot come persone o animali
  - Forma fisica
  - Comportamento naturale e sociale
  - Relazione affettiva

Sota, Pepper and Jibo

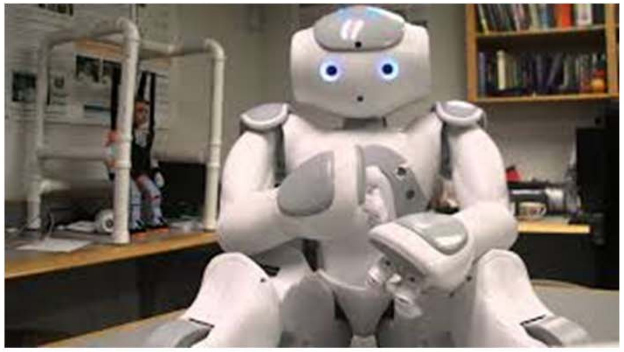
Social Robotics

- Antopomorfo
  - Per interagire con gli umani deve sembrare umani
- Caricaturale
  - Il realismo non è necessario, anzi, disturba
- Funzionale
  - Il corpo deve riflettere la funzione
- Zoomorfo
  - Robot come animale domestico
  - Più facile relazionarsi a robot se visti come animali
  - Si evita la "uncanny valley“

Social Robotics

- Robot Caricaturali

Minime caratteristiche e segnali per stimolare il comportamento sociale



Social Robotics

- Geminoid [Ishiguro]:
  - Tele-operated Android of an Existent Person

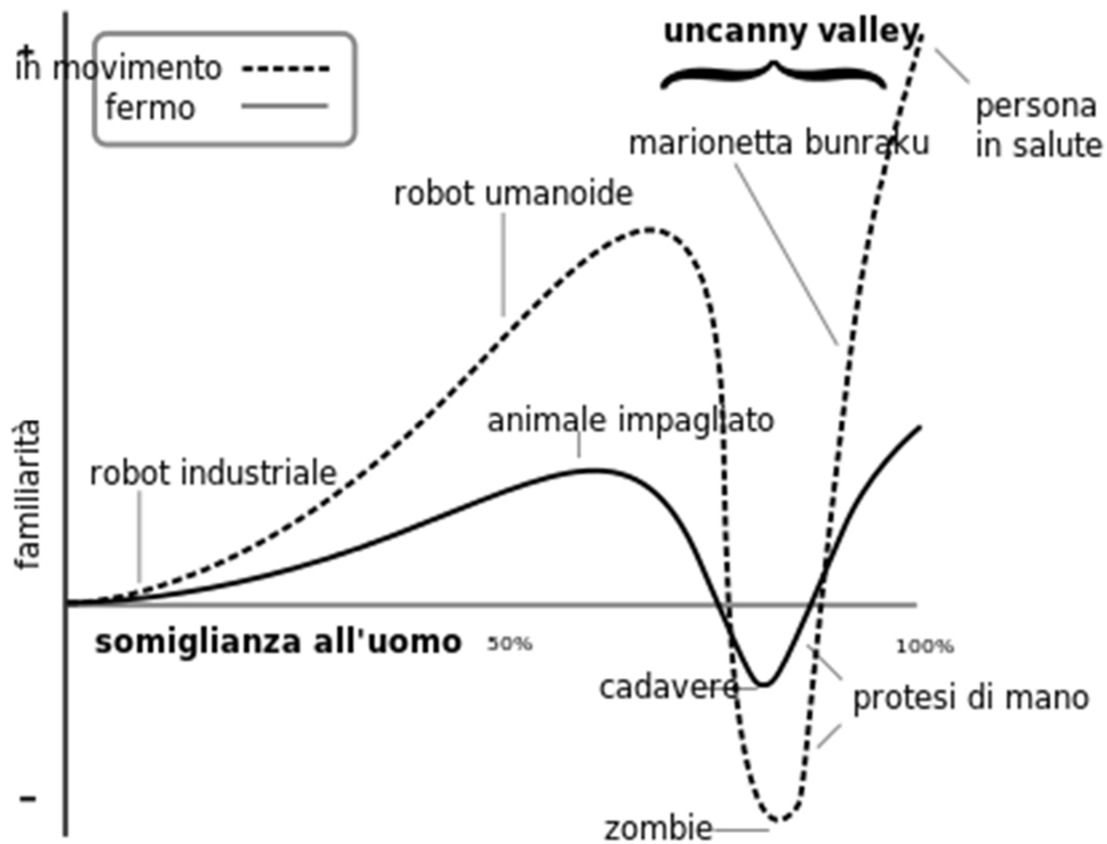
Effetti Telepresenza:
- Robot come medium

Identificazione:
- Rubber hand illusion

Human likeness:
- Espressioni facciali
- Postura
- Micro movimenti
- Chameleon effect
- Facial Feedback Hyp.

Social Robotics

- Uncanny valley



Social Robotics

- Comportamento simile a quello umano
  - Percezione ed azione simili a quelle umane
  - Sensori ed attuatori analoghi:
    - Testa, occhi, orecchie, braccia, gambe

- Comunicazione implicita ed esplicita
  - Verbale, gestuale, ma anche …
  - ... direzione dello sguardo, posizione del corpo, movimenti etc. (tutto comunica)
  - … ritmo dell’interazione, sincronizzazione, turnazione, etc.

Human-Oriented Perception

- Human-oriented perception:
  - Human Monitoring
  - Motion Capture
  - People Tracking
  - Facial Perception
  - Gaze tracking
  - Speech Recognition
  - Gesture Recognition
  - Intention Recognition
  - Plan Recognition

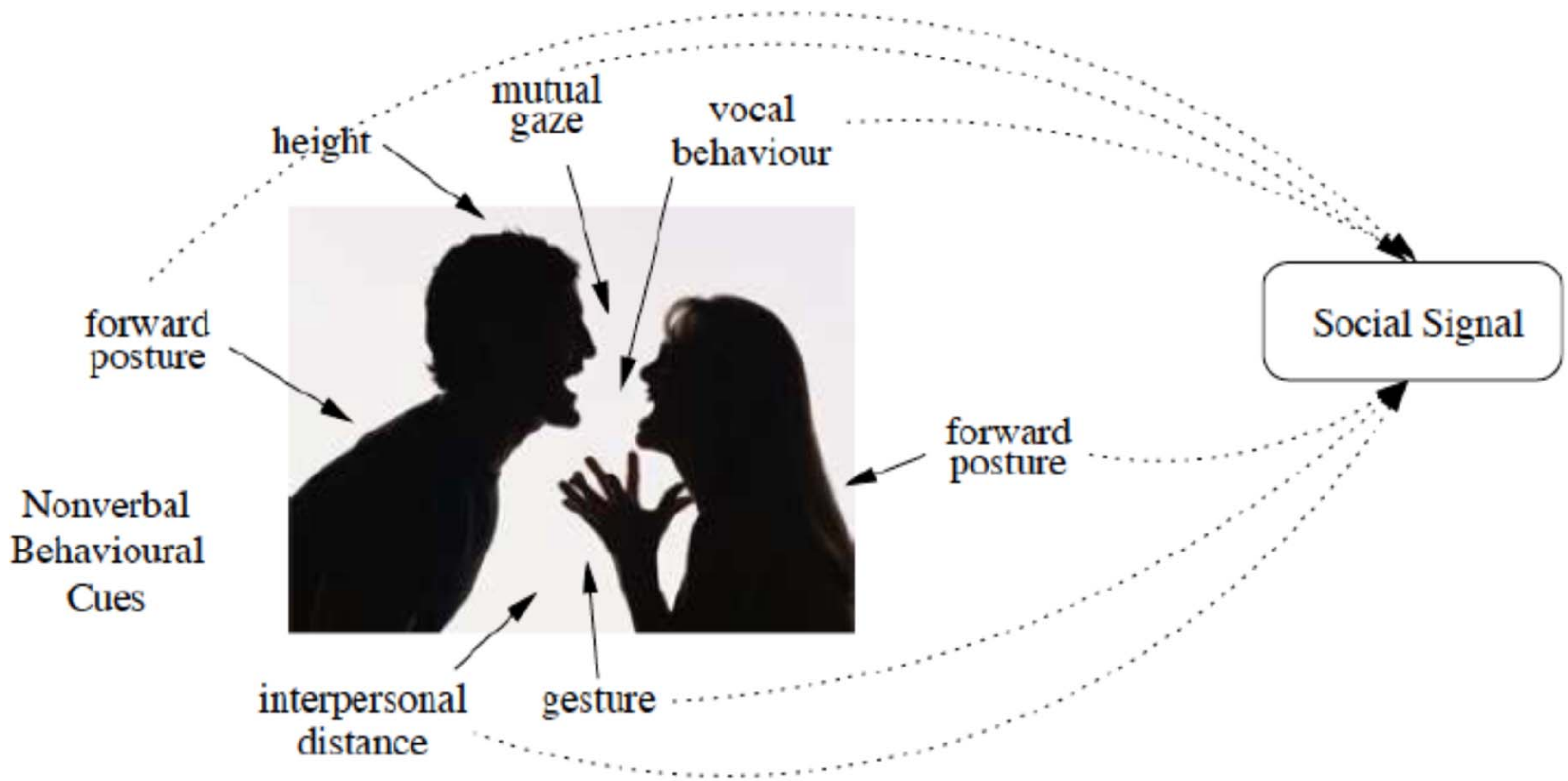
Social Interaction

- In order to interact socially with a human, a robot must convey intentionality, that is, the human must believe that the robot has **beliefs, desires, and intentions**.

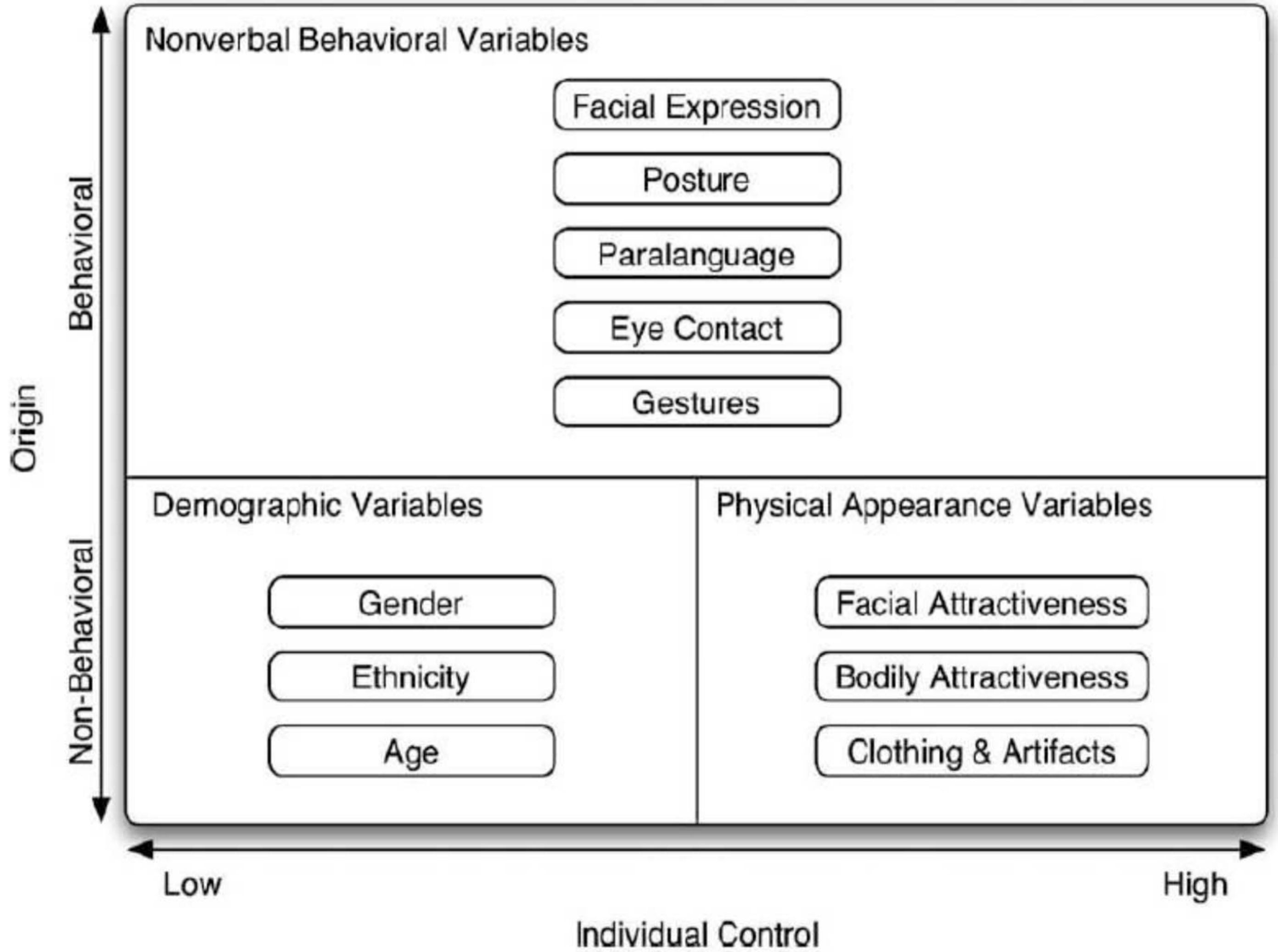
- A robot can exploits natural human social tendencies to convey intentionality through motor actions and facial expressions [Breazeal, Scassellati, 1999].

Social Signals

- Occorre integrare diversi segnali



Social Signals



DeMeuse (1987)

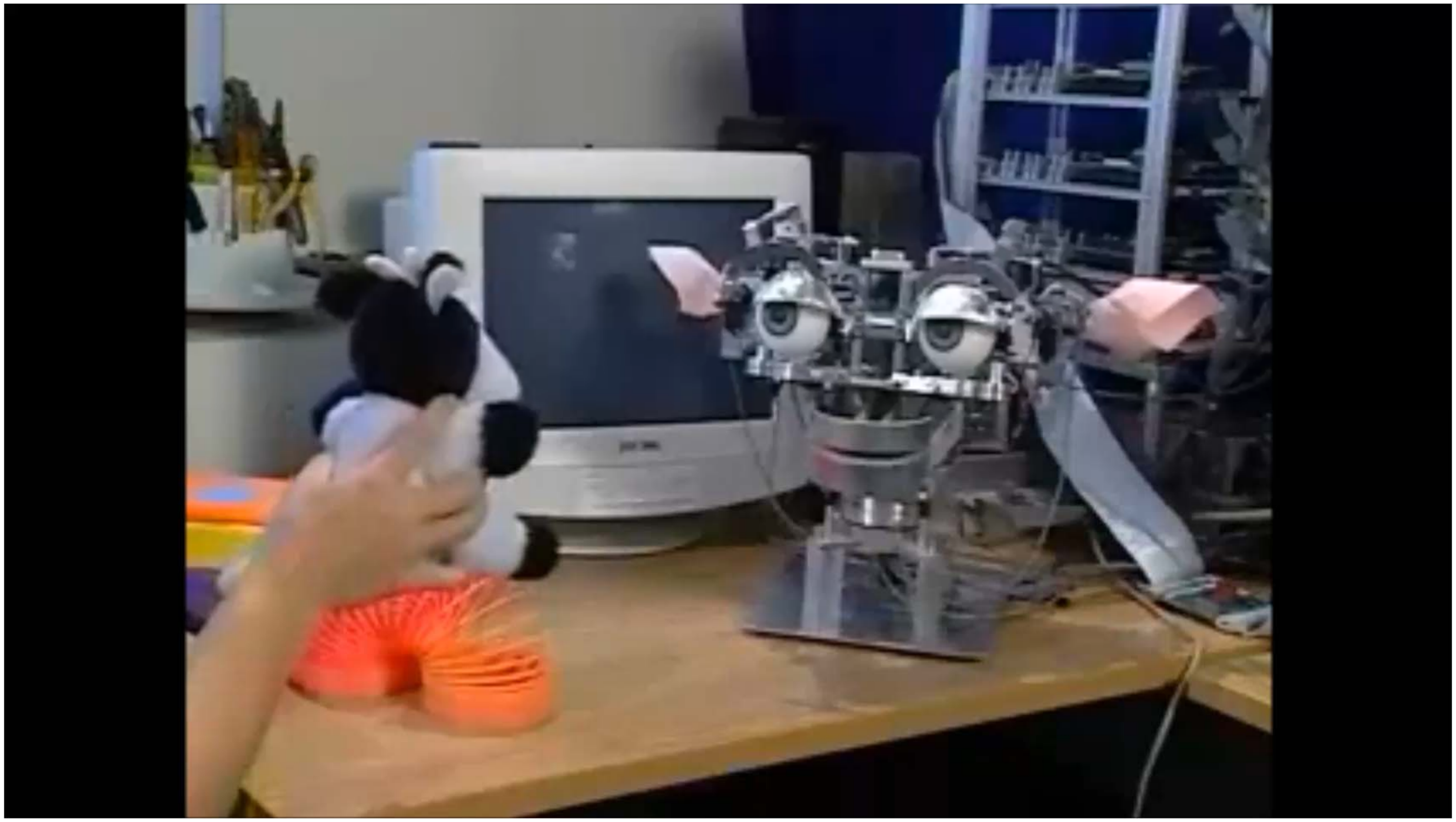
Social Interaction

To make robots acceptable to average users they should appear and behave “natural”

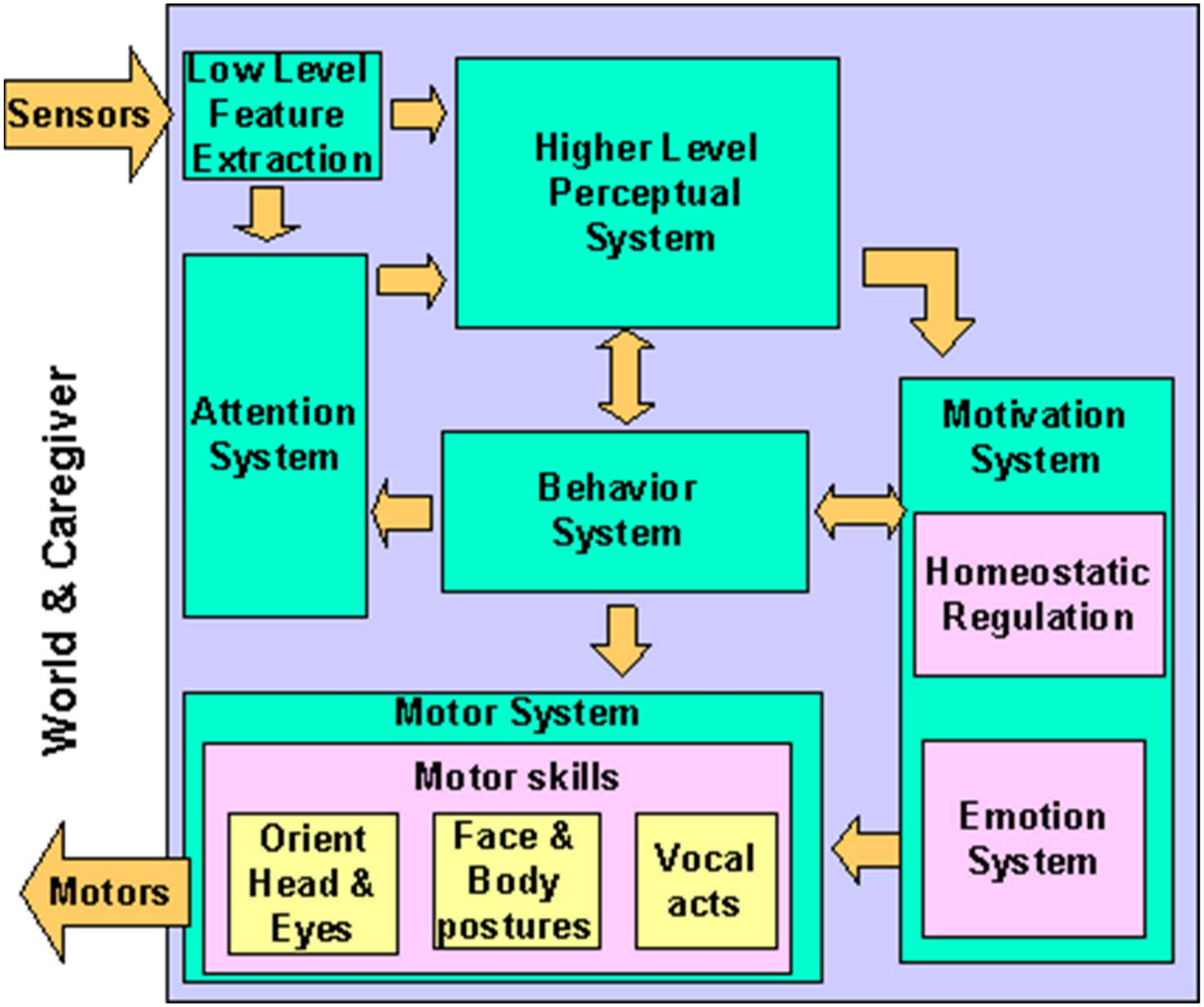
- "Attentional" Robots
  - Robot focuses on the user or the task
  - Attention forms the first step to imitation
- "Emotional" Robots
  - Robot exhibits “emotional” responses
  - Robot follows human social norms for behavior
    - Better acceptance by the user (users are more forgiving)
    - Human-machine interaction appears more “natural”
    - Robot can influence how the human reacts
    - Human can influence the robot behavior through emotional signaling

Social Interaction

- KISMET [Breazeal02]:
  - Interazione (comportamento infantile nell’apprendimento)
  - Attenzione visiva, motivazioni, emozioni, comportamento

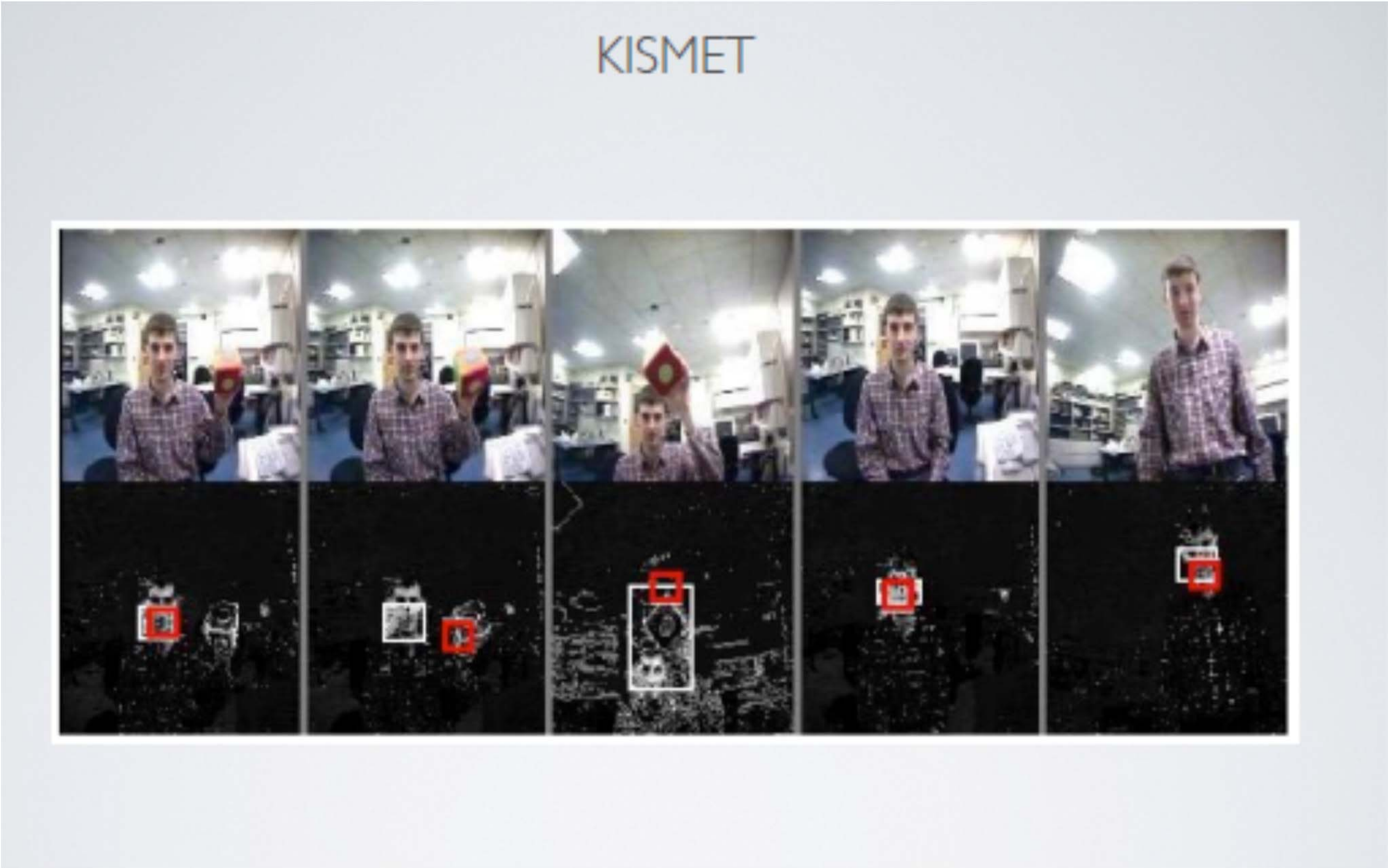


Social Interaction



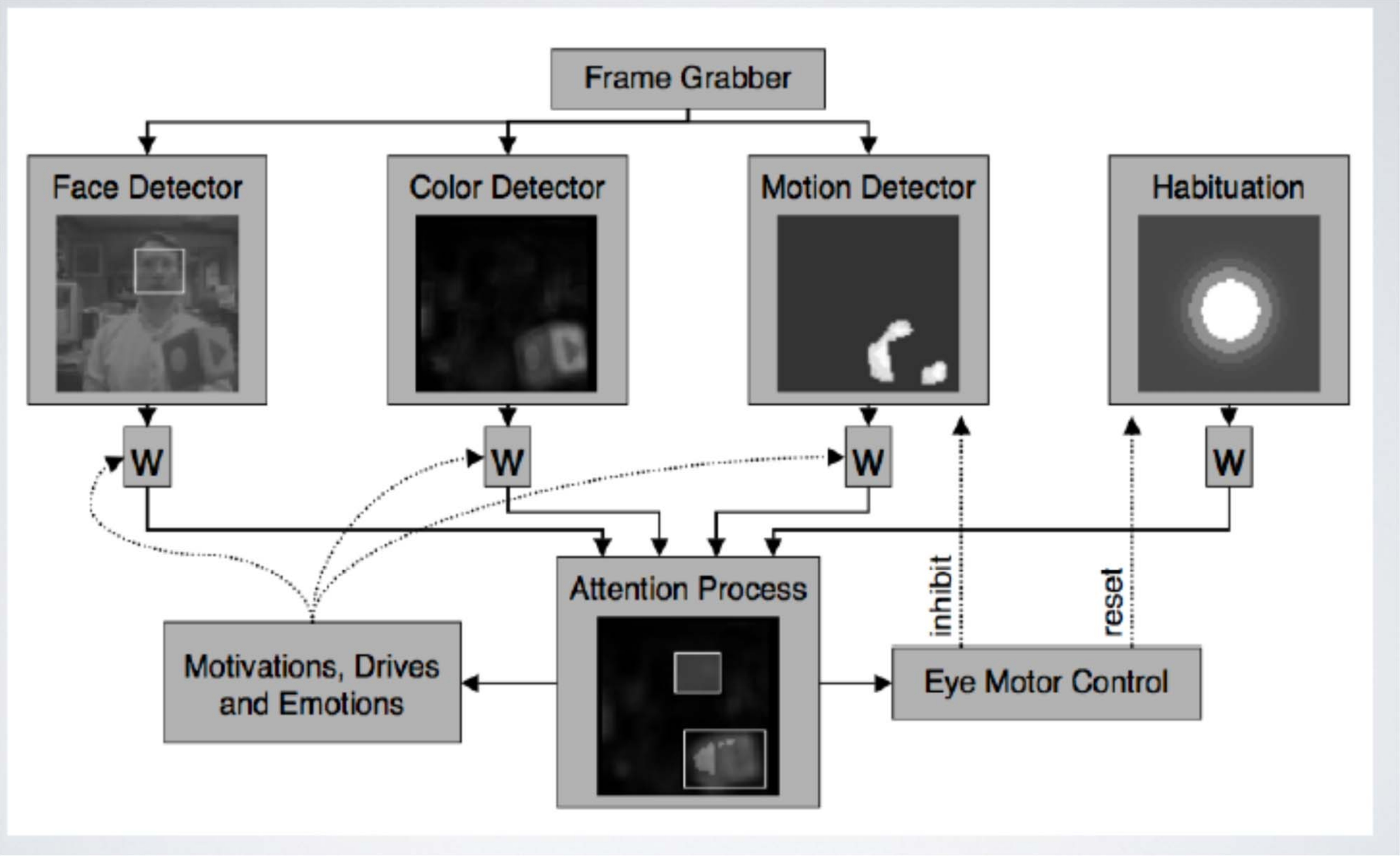
C. Breazeal, Designing Sociable Robots, MIT Press. 2002

Visual Attention



Visual Attention

Inspired by [Wolfe 1994]

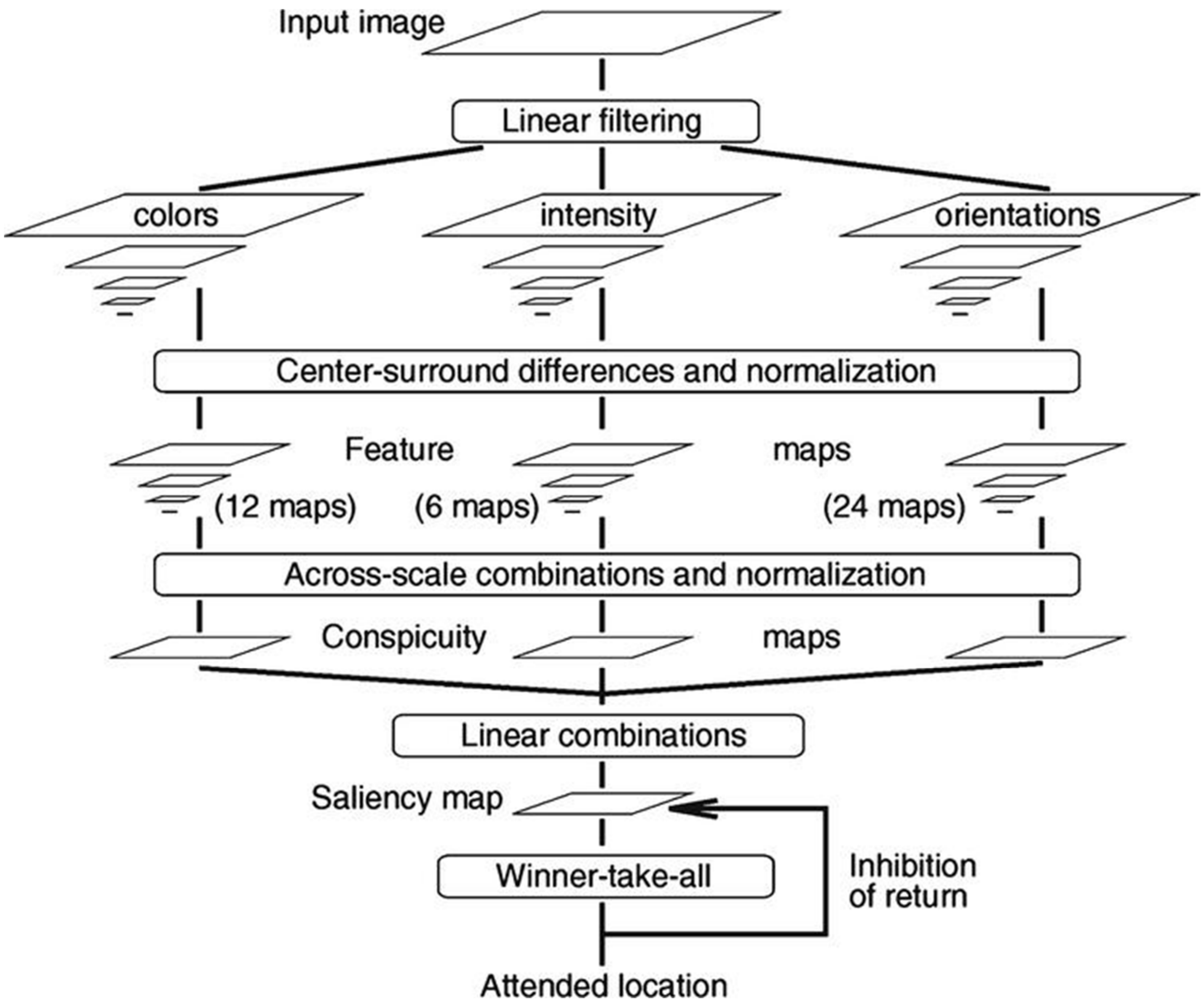


Attention

- Attenzione meccanismo fondamentale per gestire le risorse fisiche e cognitive:
    - Attenzione visiva, spaziale, acustica, etc.
    - Attenzione divisa, attenzione esecutiva, etc.

- Nella comunicazione umana (animale):
    - Monitoraggio dell’attenzione (dove guarda? cosa intende?)
    - Manipolazione dell’attenzione (indicare, evocare, guardare)

Visual Attention



Saliency

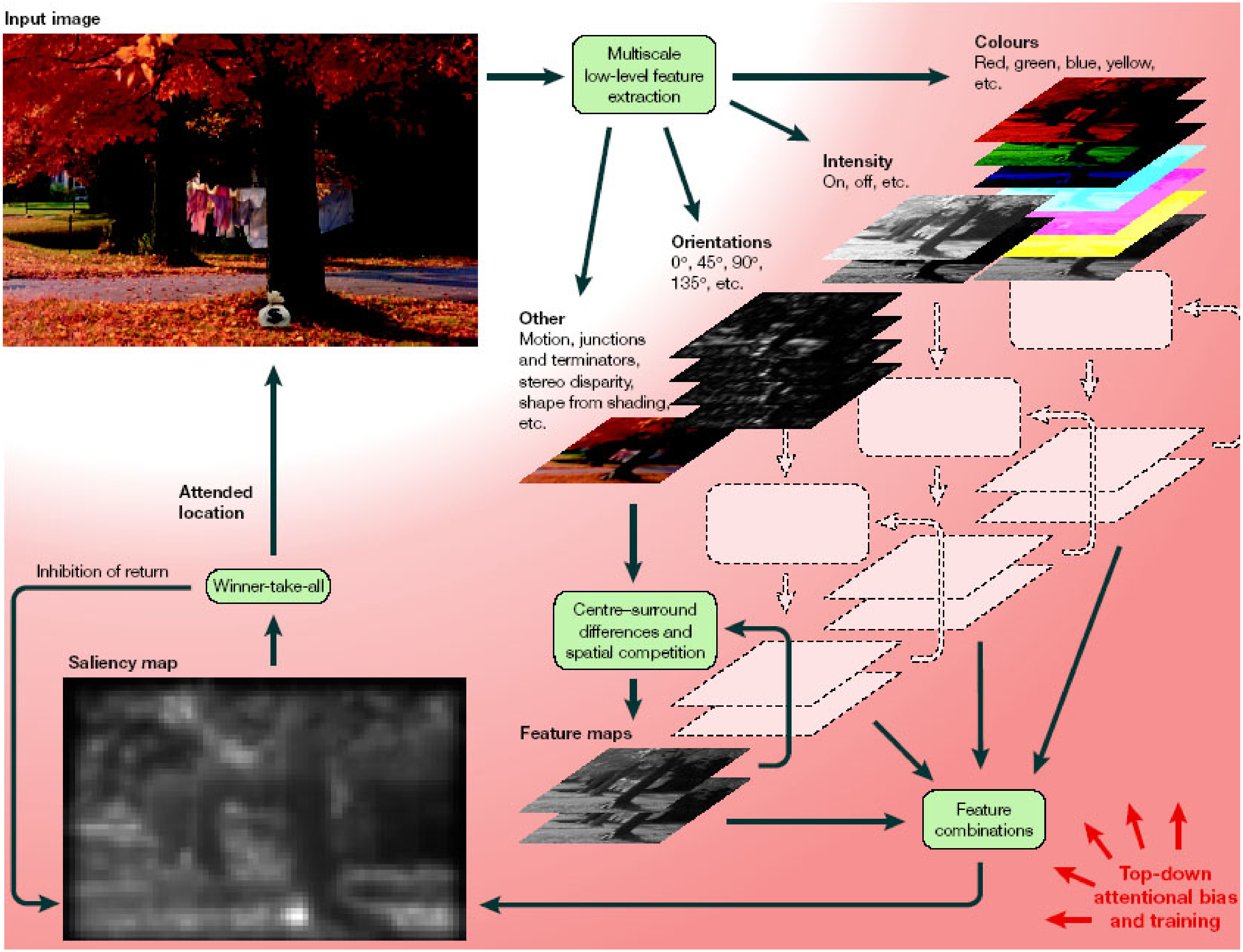
- Something is said to be **salient** if it **stands out**
- **E.g. road signs should have high saliency**

Saliency

- Trying to model visual attention
- Find locations of **Focus of Attention** in an image
- Use the idea of saliency as a basis for a model
- For primates focus of attention directed from:
  - **Bottom-up: rapid, saliency driven, task-independent**
  - Top-down: slower, task dependent

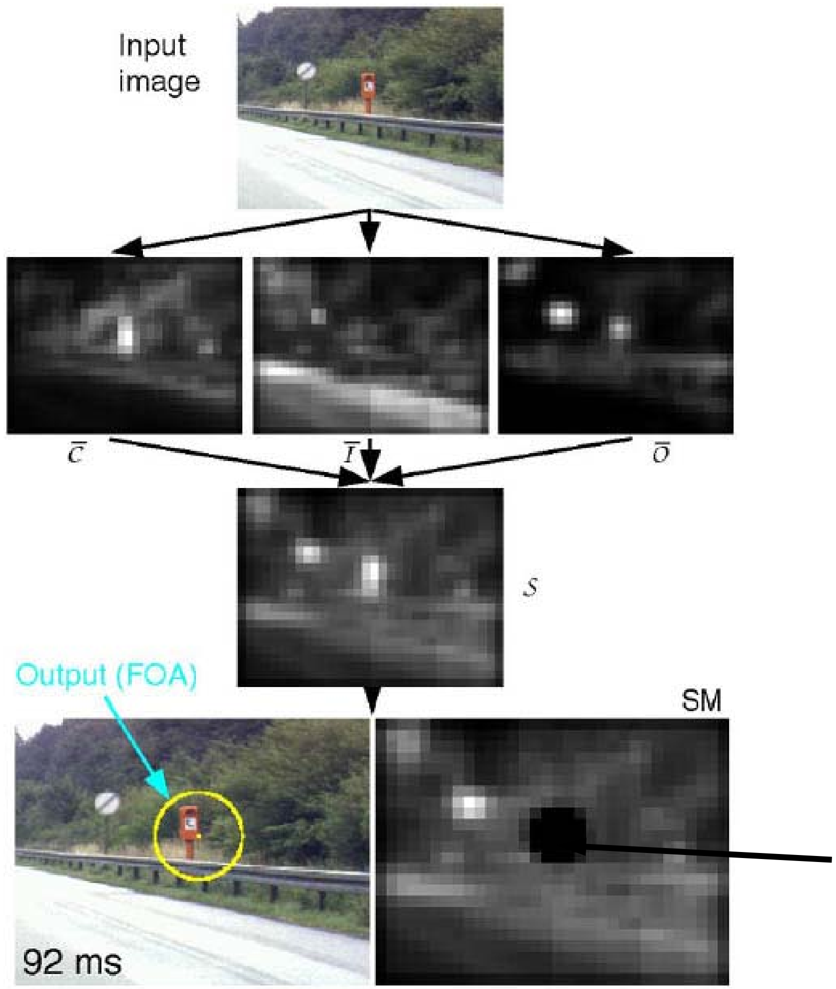
Saliency Map

- Purpose: represent saliency at all locations with a scalar quantity
- Feature maps combined into three "conspicuity maps"
  - Intensity (I)
  - Color (C)
  - Orientation (O)
- Before they are combined they need to be normalized



Input image
Multiscale low-level feature extraction
Colours
Red, green, blue, yellow, etc.
Intensity
On, off, etc.
Orientations
0°, 45°, 90°, 135°, etc.
Other
Motion, junctions and terminators, stereo disparity, shape from shading, etc.
Attended location
Inhibition of return
Winner-take-all
Centre–surround differences and spatial competition
Saliency map
Feature maps
Feature combinations
Top-down attentional bias and training

Example



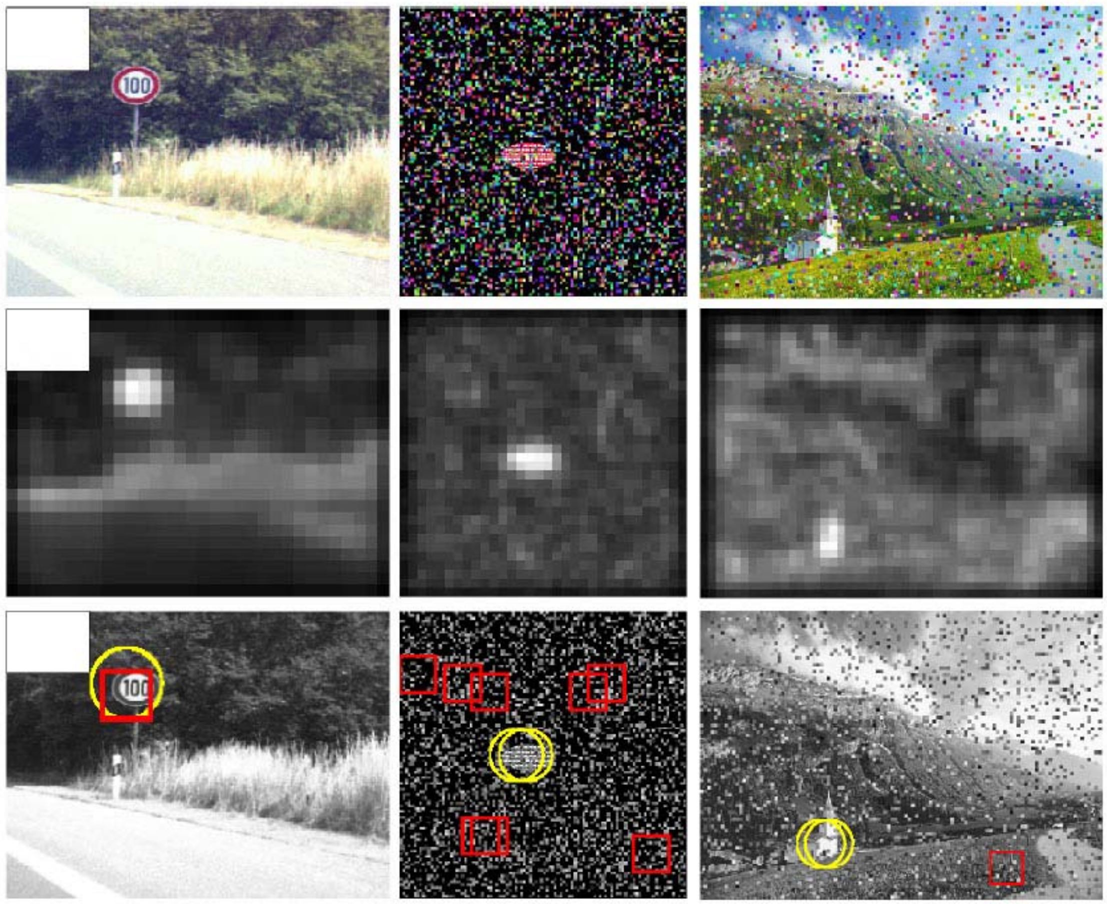
Inhibition of return

Results

Image

Saliency Map

High saliency Locations (yellow circles)



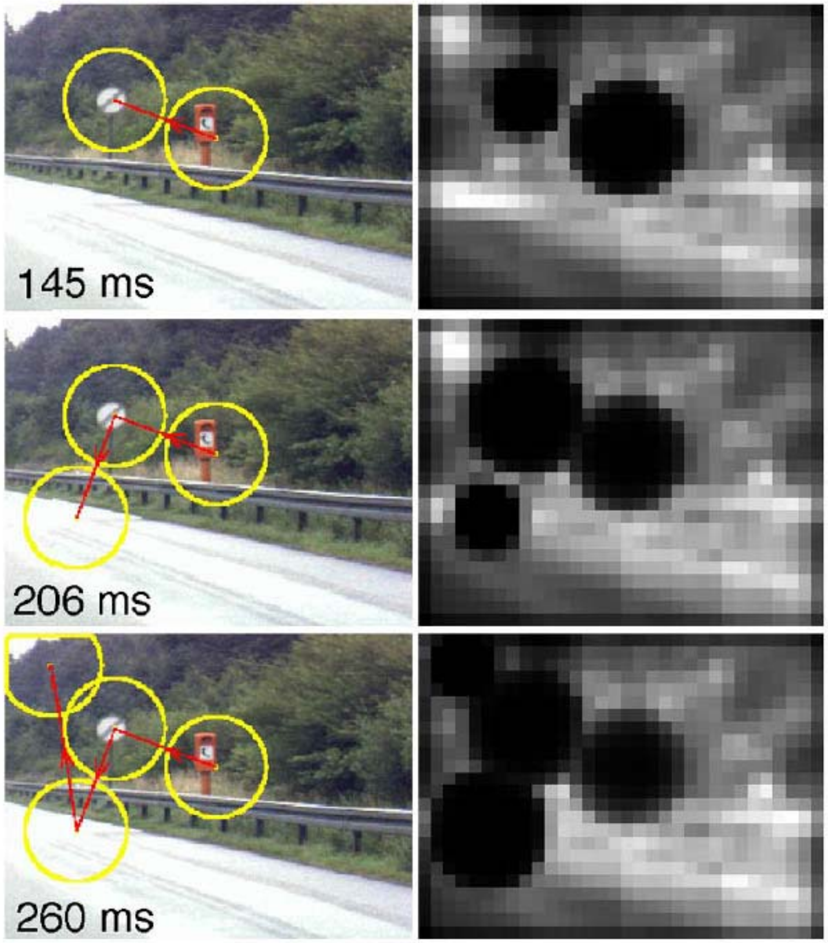
(1) (2) (3)

Results of the Model

- Only considering “Bottom-up”
  - → task-independent

Shifting Attention

- Using 2D "winner-take-all" neural network
- FOA shifts every 30-70 ms

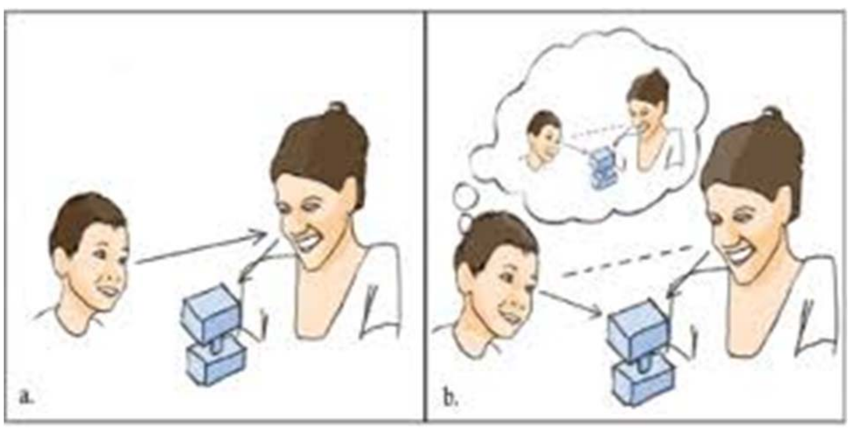


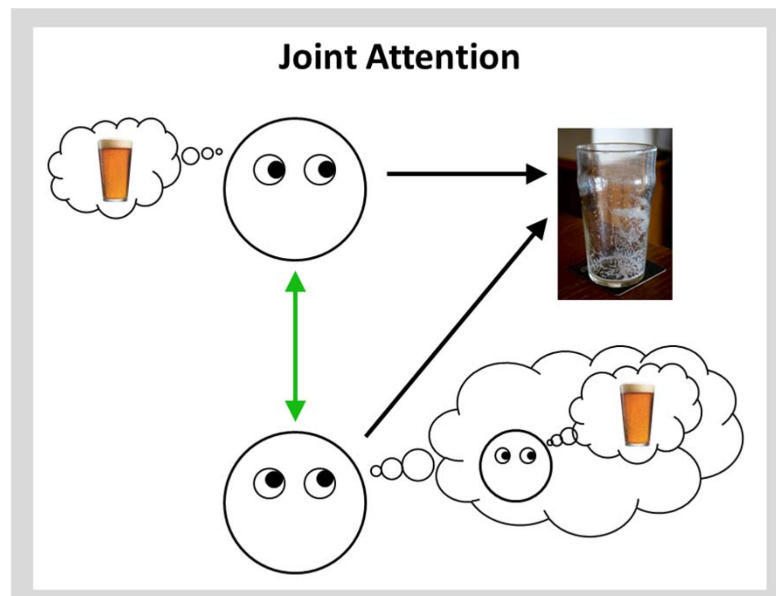
References

– Itti, Koch, and Niebur: “A Model of Saliency-Based Visual Attention for Rapid Scene Analysis” IEEE PAMI Vol. 20, No. 11, November (1998)
– Itti, Koch: “Computational Modeling of Visual Attention”, Nature Reviews – Neuroscience Vol. 2 (2001)
– Parkhurst, Law, Niebur: “Modeling the role of salience in the allocation of overt visual attention”, Vision Research 42 (2002)

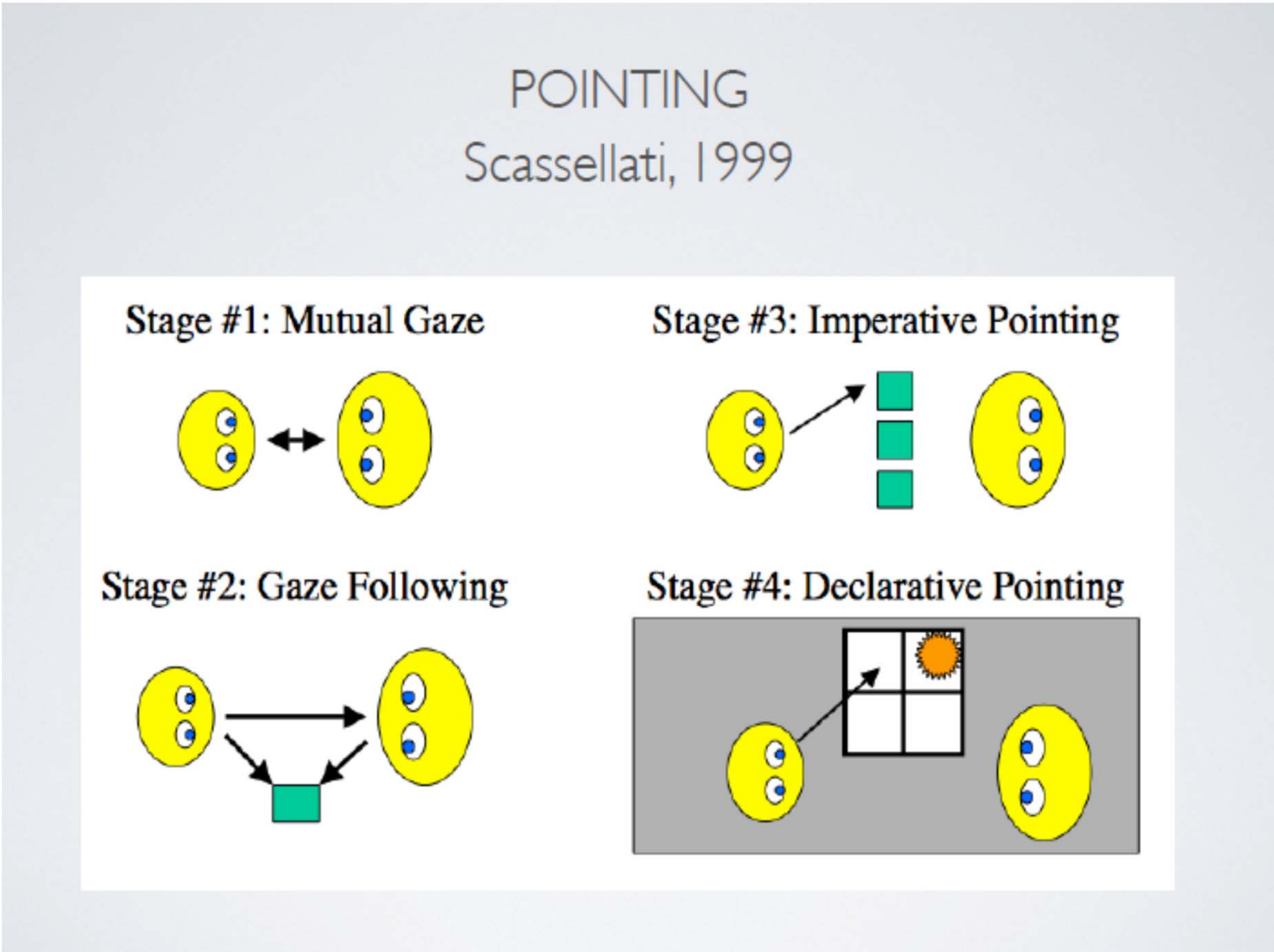
Joint Attention

- Attenzione condivisa (joint attention):
  - Focalizzazione condivisa dell’attenzione
    - Seguire lo sguardo, condividere il target, sapere che è condiviso





Joint Attention



Joint Attention

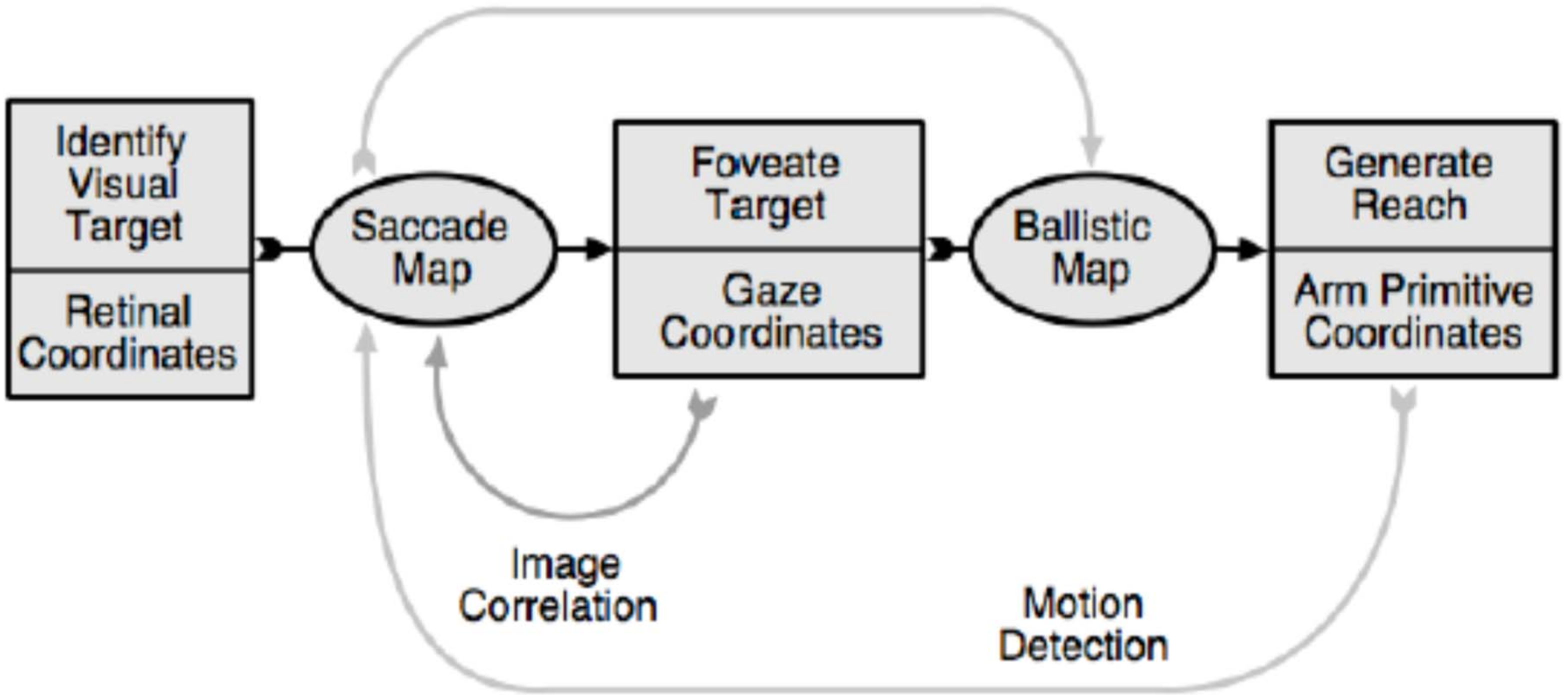


Fig. 9. Reaching to a visual target is the product of two subskills: foveating a target and generating a ballistic reach from that eye position. Image correlation can be used to train a saccade map which transforms retinal coordinates into gaze coordinates (eye positions). This saccade map can then be used in conjunction with motion detection to train a ballistic map which transforms gaze coordinates into a ballistic reach.

Joint Attention

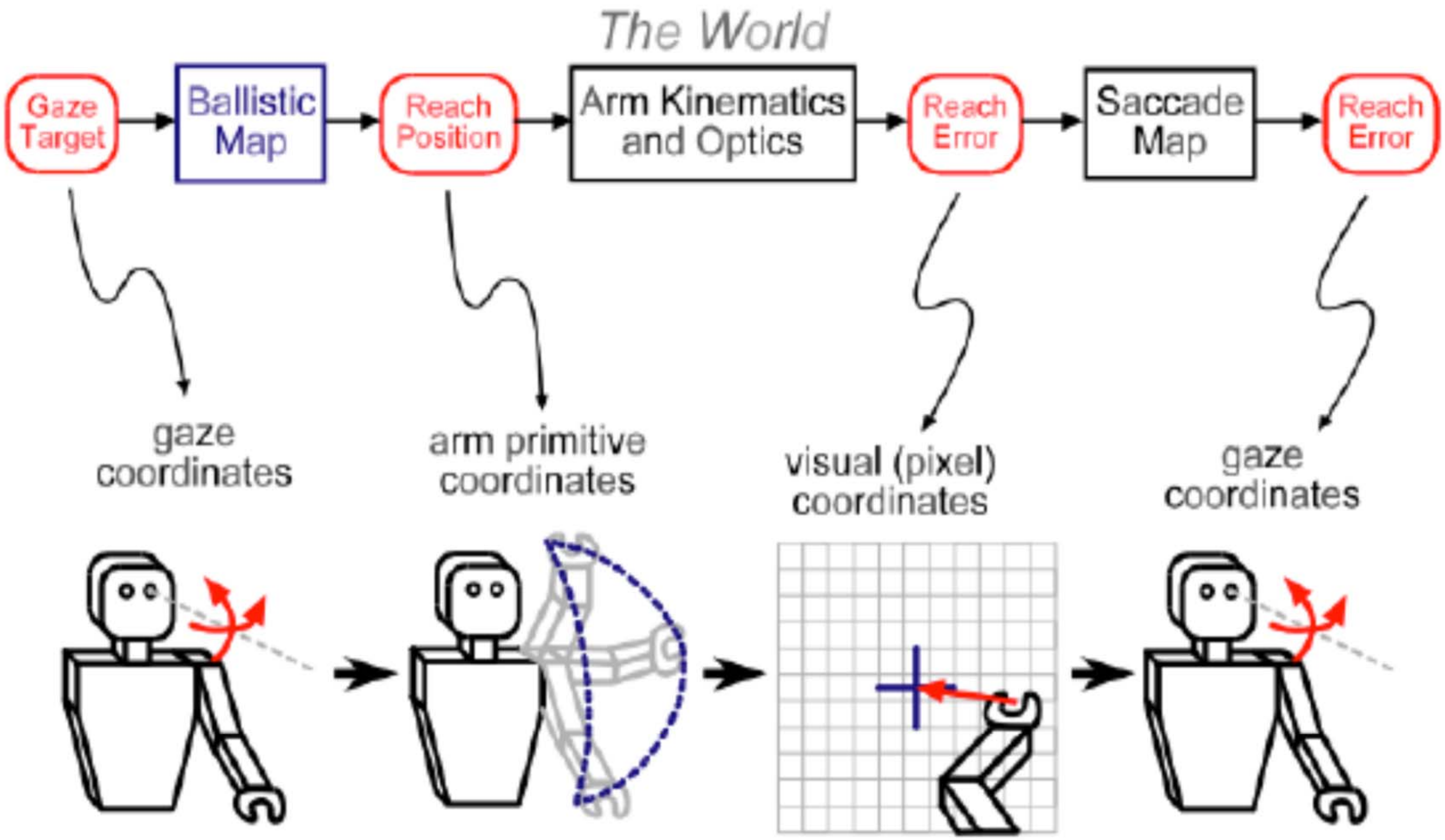


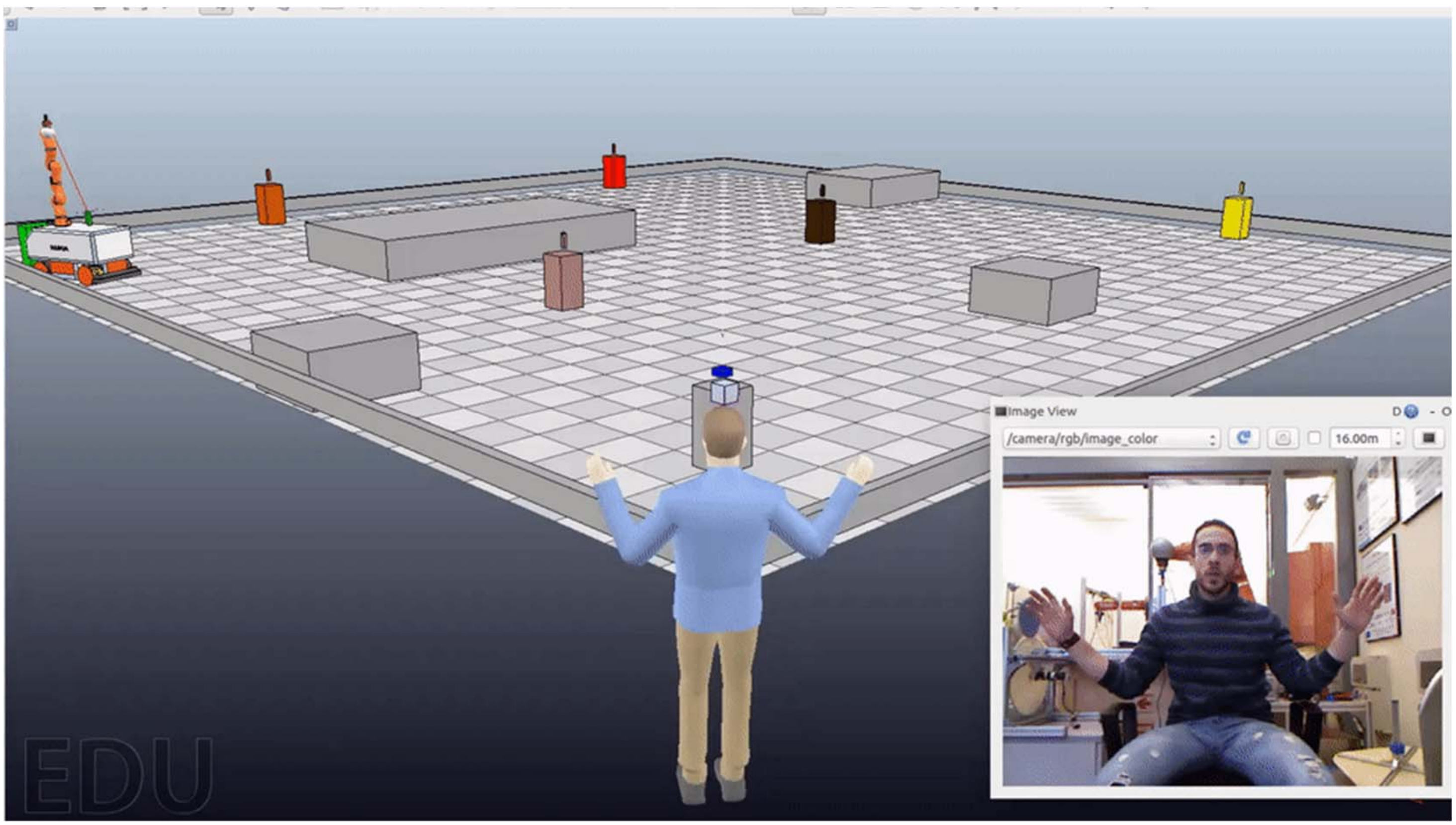
Fig. 10. Generation of error signals from a single reaching trial. Once a visual target is foveated, the gaze coordinates are transformed into a ballistic reach by the ballistic map. By observing the position of the moving hand, we can obtain a reaching error signal in image coordinates, which can be converted back into gaze coordinates using the saccade map.

Attention

- Attention manipolation:
  - Gestures, gaze, cues, etc.

Attention

- Attention manipolation:
  - Gestures, gaze, cues, etc.



[Caccavale et al. 2016]

Social Interaction and Emotions

Advantages:

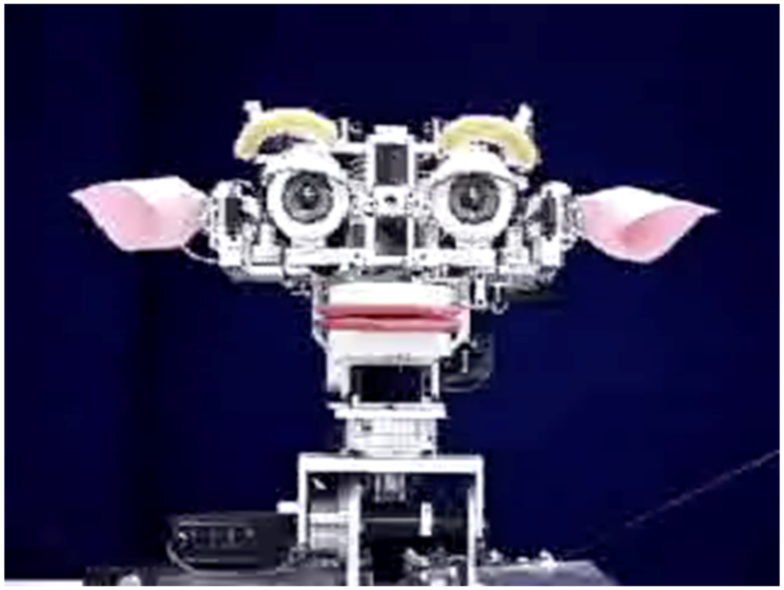
- Robots that look human and that show “emotions” can make interactions more “natural”
  - Humans tend to focus more attention on people than on objects
  - Humans tend to be more forgiving when a mistake is made if it looks “human”
- Robots showing “emotions” can modify the way in which humans interact with them

Problems:

- How can robots determine the right emotion ?
- How can “emotions” be expressed by a robot ?

Social Interaction and Emotions

- Riconoscimento ed esternazione di emozioni
  - Facilita la comunicazione e aumenta l'empatia
  - Espressioni facciali, postura, velocità, colore, etc.

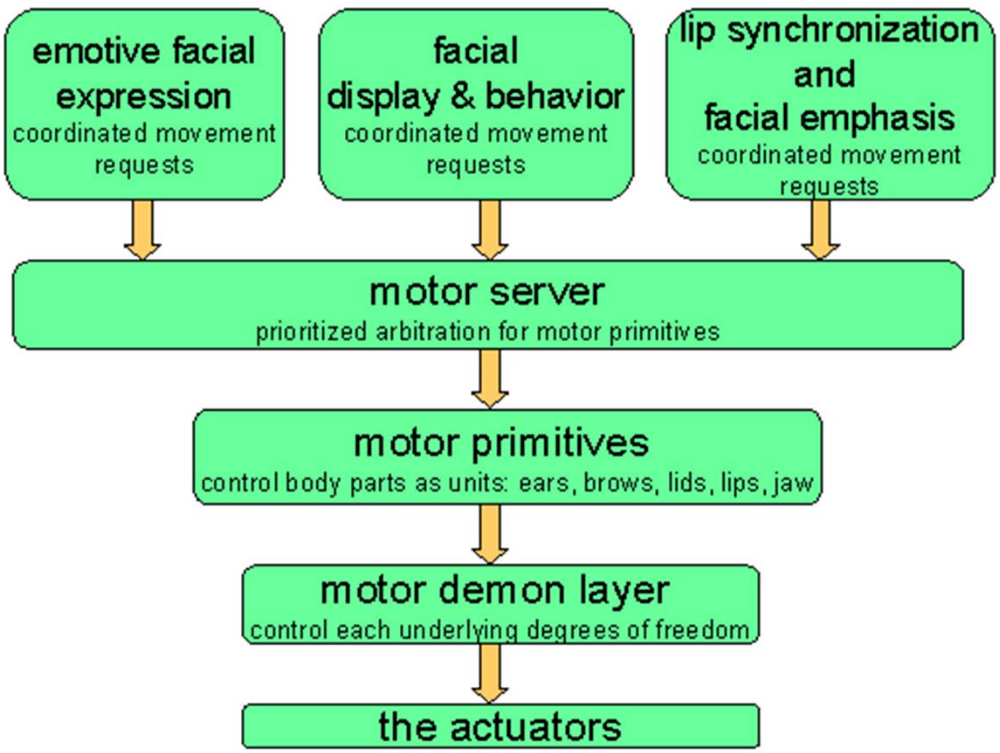


KISMET

COZMO

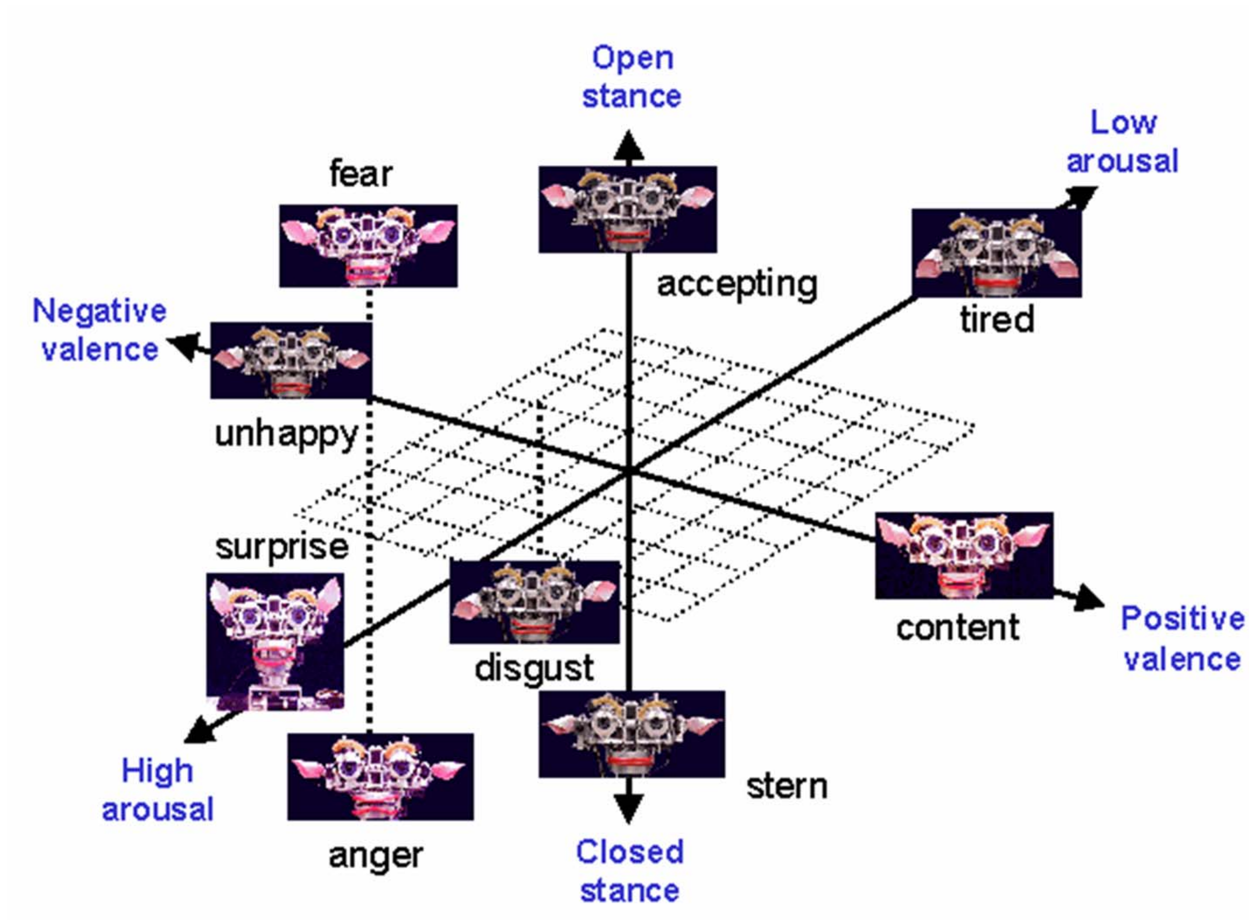
Social Interaction and Emotions

- Riconoscimento ed esternazione di emozioni
  - Facilita la comunicazione e aumenta l'empatia
  - Espressioni facciali, postura, velocità, colore, etc.



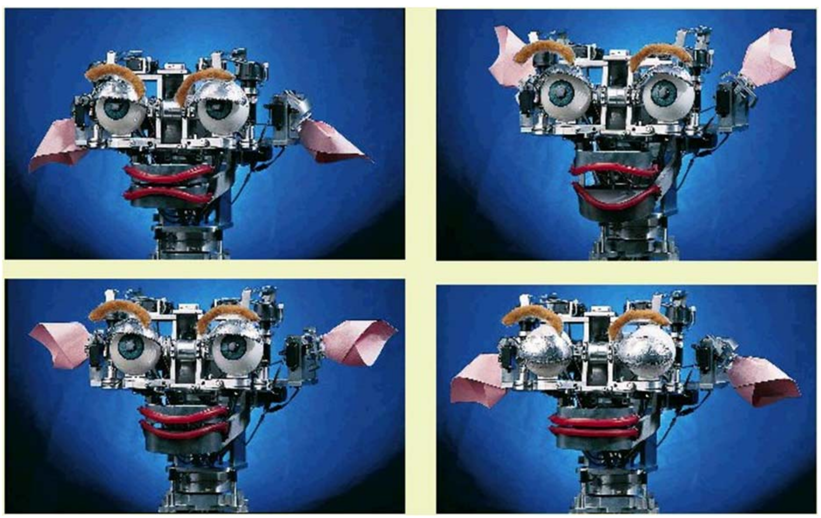
Social Interaction and Emotions

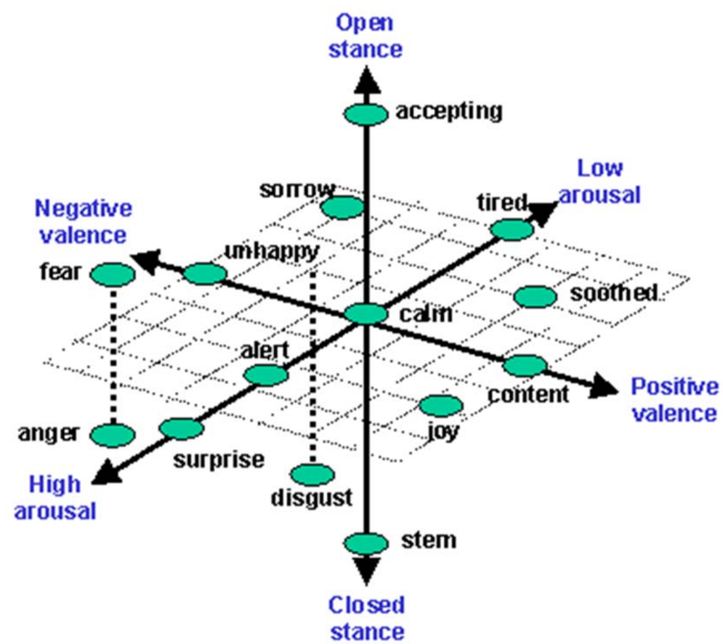
- Riconoscimento ed esternazione di emozioni



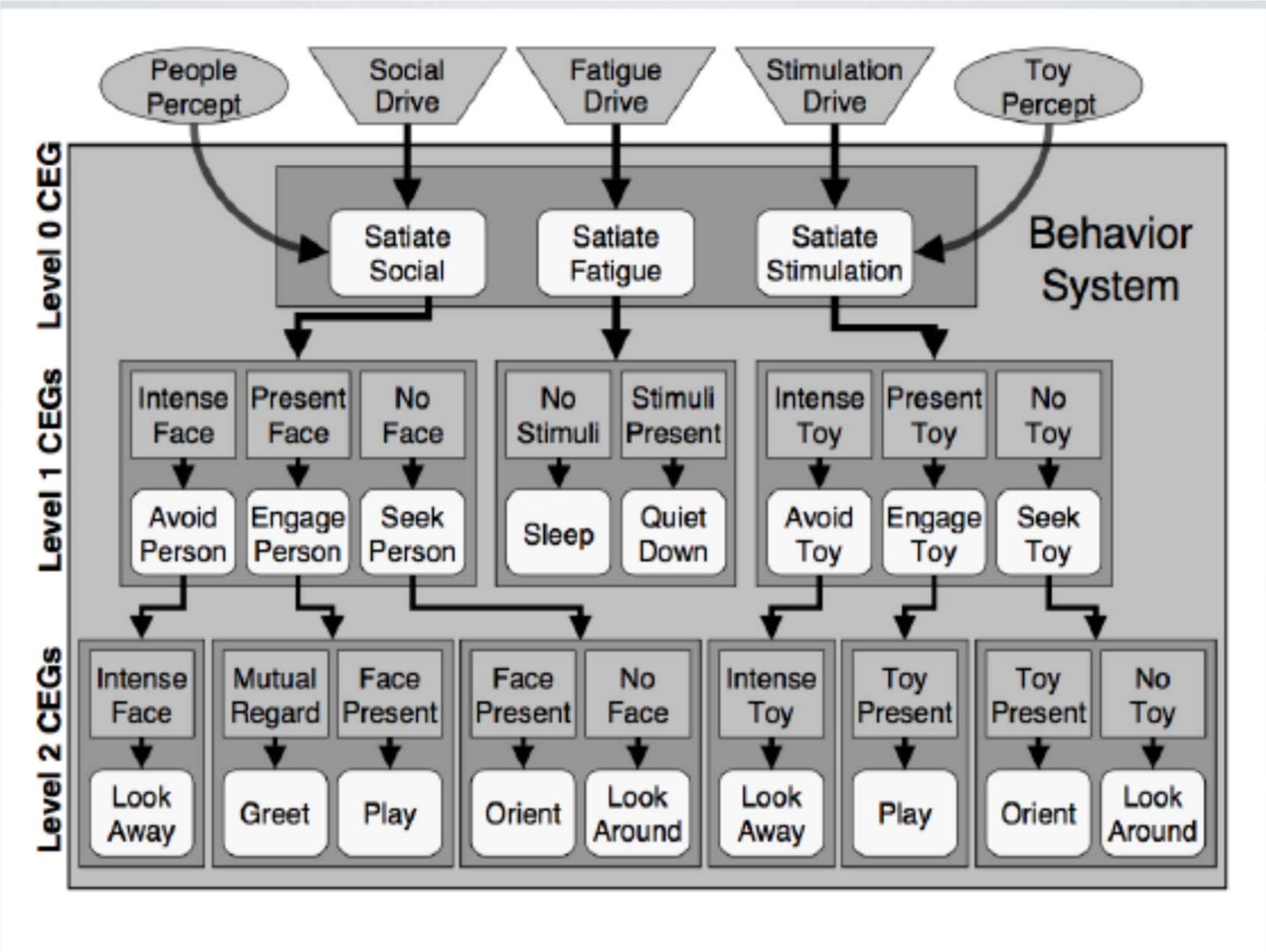
Emotions

- Riconoscimento ed esternazione di emozioni
  - Facilita la comunicazione e aumenta l'empatia
  - Espressioni facciali, postura, velocità, colore, etc.





Behaviors


People Percept
Social Drive
Fatigue Drive
Stimulation Drive
Toy Percept
Level 0 CEG
Satiate Social
Satiate Fatigue
Satiate Stimulation
Behavior System
Level 1 CEGs
Intense Face
Present Face
No Face
Avoid Person
Engage Person
Seek Person
No Stimuli
Stimuli Present
Sleep
Quiet Down
Intense Toy
Present Toy
No Toy
Avoid Toy
Engage Toy
Seek Toy
Level 2 CEGs
Intense Face
Look Away
Mutual Regard
Greet
Face Present
Play
Face Present
Orient
No Face
Look Around
Intense Toy
Look Away
Toy Present
Play
Toy Present
Orient
No Toy
Look Around

Summary

- Cognitive Robotics
  - Goal, methods, applications
  - Human-Robot Interaction:
    - Service Robotics, Collaborative Robots, Social Robotics

- Approaches:
  - Enactive Robotics, Developmental Robotics, Neuro-robotics
    - Phylogenesis, Ontogenesis, Learning
  - Cognitive Architectures:
    - Symbolic, Hybrid, Emergent
    - Functional, Bio-Palusible
  - Cognitive Control:
    - Behavior orchestration, attention model, cognitive workload
  - Communication and Interaction:
    - Intention recognition, activity recognition, multimodal interaction