Collaboration in Human-Robot Teams

- Unified theory of cognition for solving the whole problem
  - We already know how to be mobile, avoid collisions, etc
- Cognitive psychology
  - Study human-human collaboration
  - Determine important high-level cognitive skills
  - Build computational cognitive models of these skills
    - ACT-R, SOAR, EPIC, Polyscheme…
  - Use computational models as reasoning mechanism on robot for high-level cognition

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;
  - 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.

Cognitive Skills

- Appropriate knowledge representations
  - Spatial representation for spatial reasoning
  - Adapting representation to problem solving method
- Problem solving
  - Navigation routing with constraints (e.g., remaining hidden)
- Learning
  - Learning to recognize and anticipate others' behaviors
  - Learning characteristics of other's capabilities
- Vision
  - Object permanence and tracking (Cassimatis et al., 04)
  - Recognizing gestures
- Natural communication (Perzanowski et al., 01)
  - Multimodal communication, dialogue, etc.

Robot-Human Interaction

- The robot has to be able to communicate its intentions to the human
  - Output has to be easy to understand by humans
  - Robot has to be able to encode its intention
  - Interface has to keep human's attention

- Robot communication devices:
  - Screen-based
  - Robot speech
  - Robot "gestures"
  - Vibro/haptics

Human-Robot Interaction

- Robot Interfaces have to be easy to use
  - Robots have to be controllable by untrained users
  - Robots have to be able to interact not only with their owner but also with other people
- Robot interfaces have to be usable at the human's discretion
  - Human-robot interaction occurs on an irregular basis
    - Frequently the robot has to operate autonomously
    - Whenever user input is provided the robot has to react to it
- Interfaces have to be designed human-centric
  - The role of the robot is it to make the human's life easier and more comfortable (it is not just a tech toy)

Human-Robot Interaction

- Existing technologies
  - Simple voice recognition and speech synthesis
  - Gesture recognition systems
  - On-screen, text-based interaction
- Research challenges
  - How to convey robot intentions ?
  - How to infer user intent from visual observation (how can a robot imitate a human) ?
  - How to keep the attention of a human on the robot ?
  - How to integrate human input with autonomous operation ?

Human-Robot Interaction

- Field of study dedicated to **understanding, designing, and evaluating robotic systems for use by, or with, humans** [Goodrich 2007]

MIT
Media Lab

Rethink
Robotics

Osaka U.

NASA

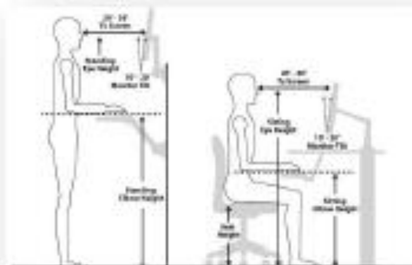
Human-Robot Interaction

- **1992**: 1st IEEE International Symposium on Robots and Human Interactive Communications (RO-MAN)
- **Late 1990's**: workshops and conference tracks on HRI at International Robotics Conferences (e.g. AAAI Symposia Series, ICRA, IROS, RSS, Human Factors, etc.)
- **2000**: 1st IEEE/RAS International Conference on Humanoid Robots (HUMANOIDS)
- **2004**: IEEE/RAS & IFRR summer school on Human-Robot Interaction
- **2006**: 1st ACM International Conference on Human-Robot Interaction (HRI)
- **2012**: inaugural volume of Journal of Human-Robot Interaction (jHRI)

Human-Robot Interaction

Related Fields

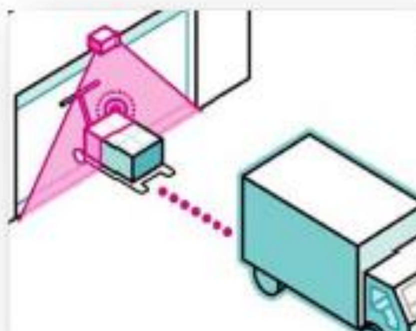
Telerobotics & Haptics



Human Factors

Air Traffic Control

Human-Computer Interaction



Operations Research

Machine Learning

Human-Robot Interaction

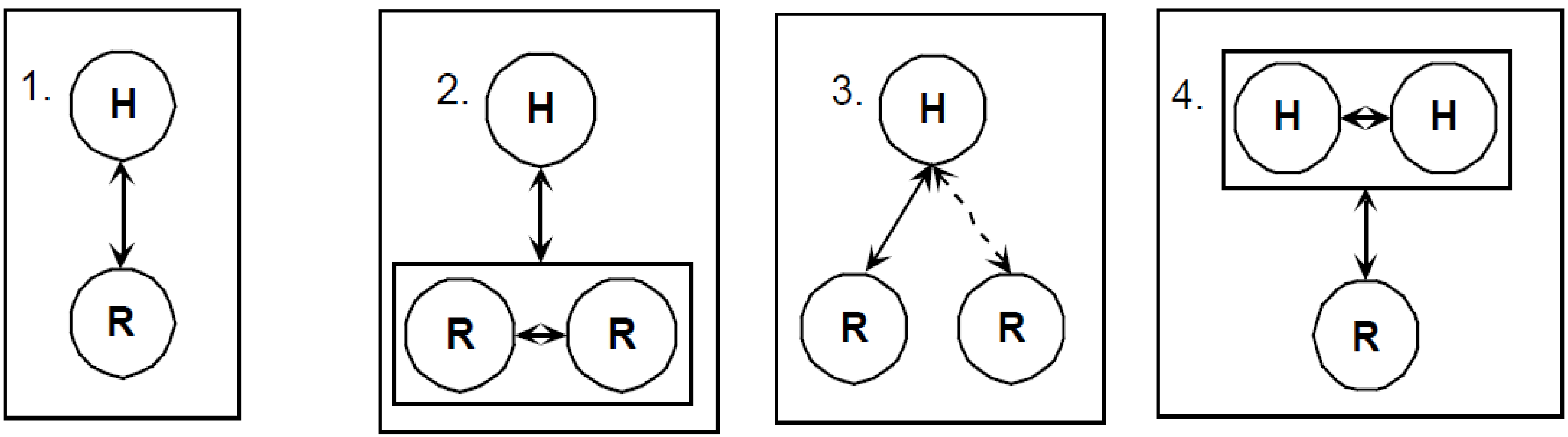
- Specific features:
  - Physical/cognitive interaction with embodied intelligence
  - Social relation between humans and robots
  - Complex, dynamic, unpredictable environment/human

Taxonomies in HRI

- Physical vs. cognitive
- Co-located vs. remote
- Team configurations [Yanco2002]

Taxonomies in HRI

- Team configurations [Yanco2002]

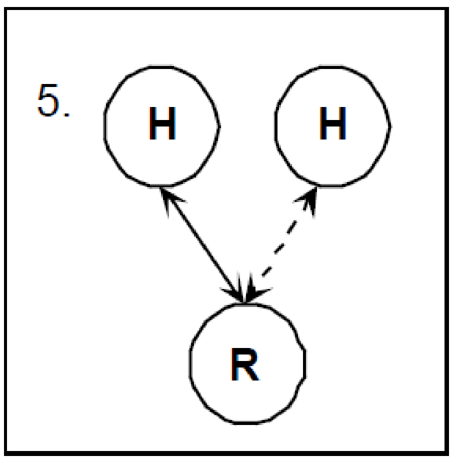


Human-Robot Human-Team Human-Robots Robot-Team

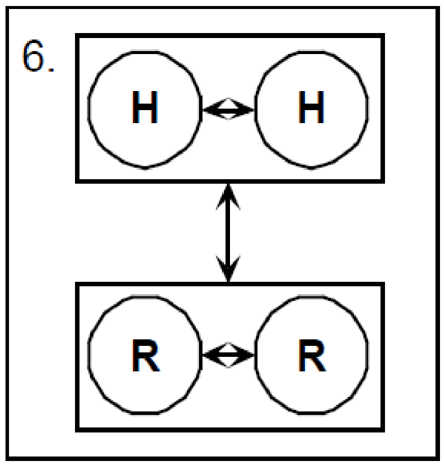
HUMAN-ROBOT-RATIO
ROBOT-TEAM-COMPOSITION
(homogeneous/heterogeneous)
ROBOT-MORPHOLOGY
(anthropomorphic,zoomorphic,functional)

Taxonomies in HRI

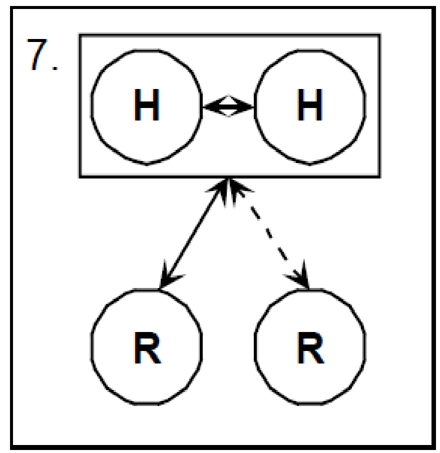
- Team configurations [Yanco2002]



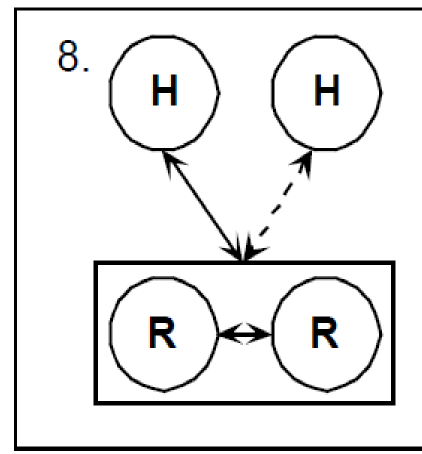
mHuman-1Robot



TeamR-TeamH



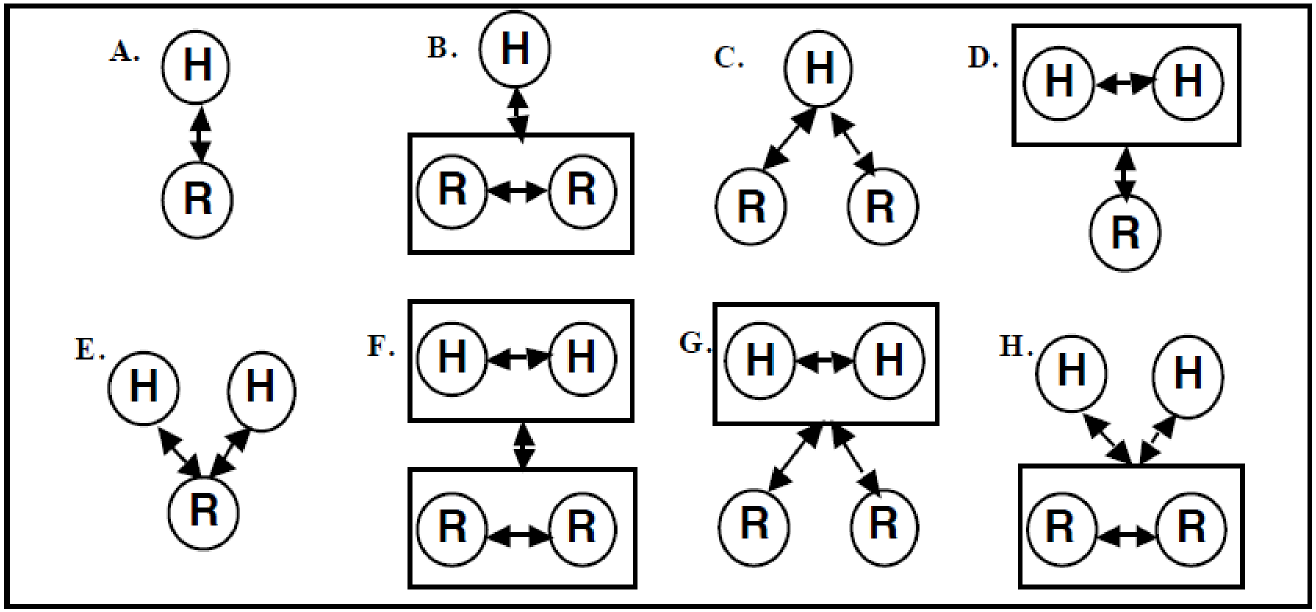
TeamH-mRobot



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

Taxonomies in HRI

- 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

Taxonomies in HRI

- Levels of Autonomy (LOA) [Sheridan 1978]
  1. Computer offers no assistance; **human does it all**
  2. Computer offers a **complete set of action alternatives**
  3. Computer narrows the selection down to a **few choices**
  4. Computer suggests a **single action**
  5. Computer executes that action **if human approves**
  6. Computer allows the human **limited time to veto** before automatic execution
  7. Computer **executes automatically** then necessarily informs the human
  8. Computer informs human after automatic execution **only if human asks**
  9. Computer informs human after execution **only if it decides to**
  10. Computer decides everything and acts autonomously, **ignoring the human**

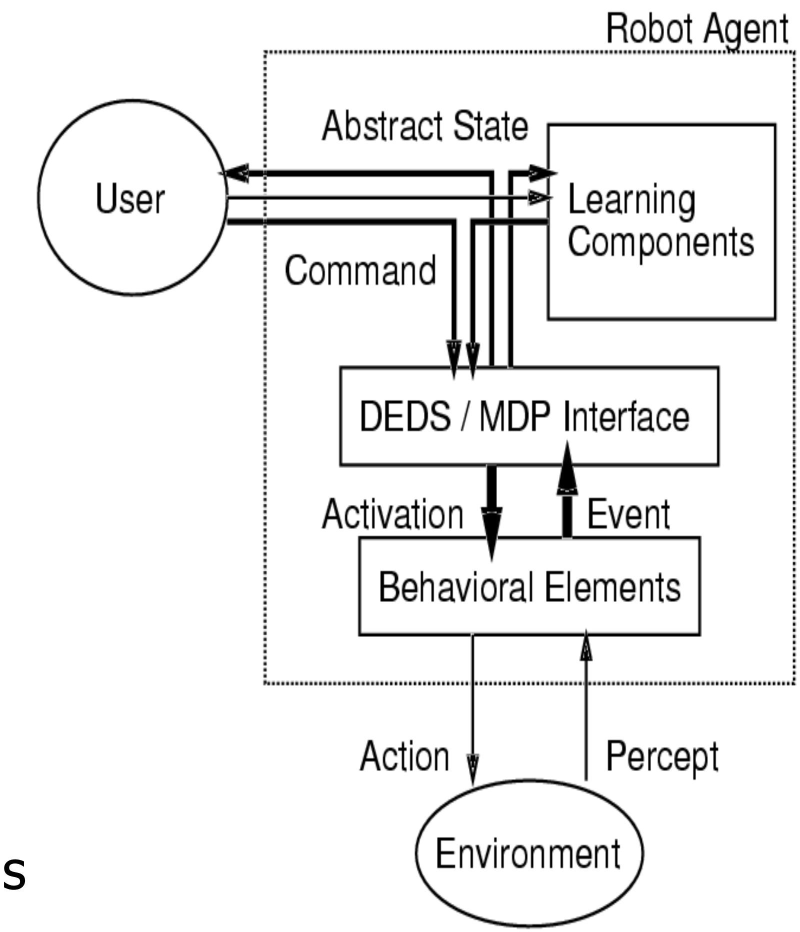
Taxonomies in HRI

Adjustable Autonomy

- The robot can operate at varying levels of autonomy

- Operational modes:
  - Autonomous operation
  - User operation / teleoperation
  - Behavioral programming
  - Following user instructions
  - Imitation

- Types of user commands:
  - Continuous, low-level instructions (teleoperation)
  - Goal specifications
  - Task demonstrations



Example System

Taxonomies in HRI

Multi-robot management

- Cognitive overload
- Relates to level of autonomy, nature of task, and mode of communication
- Fan-Out [Goodrich 2003]

Taxonomies in 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

Information Exchange

- Methods of communication:

Medium	Robot-Initiated	Human-Initiated
Visual	Visual displays (lights, GUI, VR, AR)	Gestures (hand, facial, body)
Audio	Natural language, non-speech audio	Natural language
Tactile	Haptics	Keyboard, mouse, gamepad, haptics

HRP-4C (NIST)

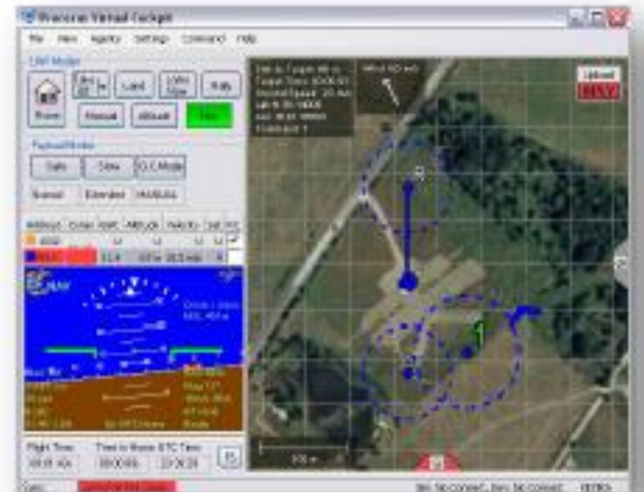
Tag-based Comm. with Aqua Robot (McGill U.)

Paro Theraputic Robot

Principles of Efficient HRI

1. Implicitly switch interface and autonomy modes
2. Let robot use natural human cues

© Raizlabs



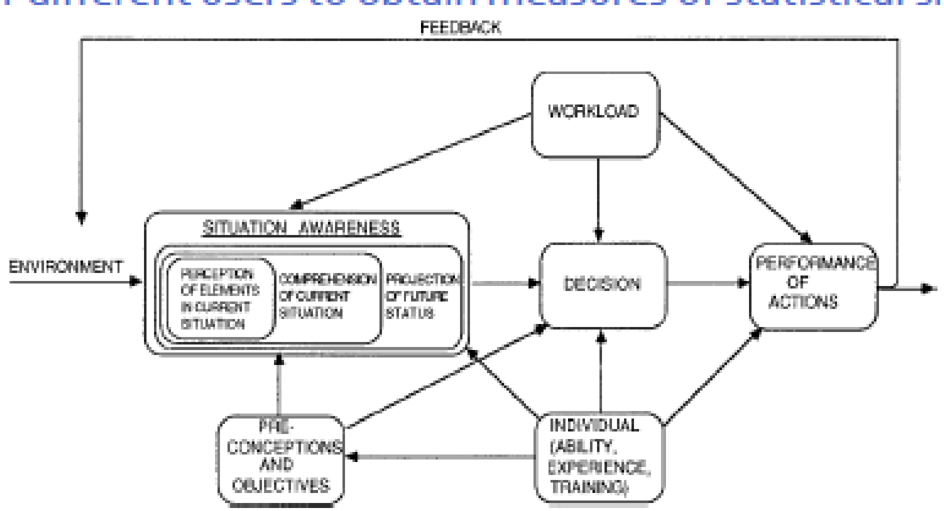
Virtual Cockpit v2
(Procerus / Lookheed Martin)

3. Manipulate the world instead of the robot
4. Manipulate relationship between the robot and the world
5. Let people manipulate presented information
6. Externalize memory
7. Help people manage attention

Situation Awareness

Situational Awareness Global Assessment Technique (SAGAT) [Endsley 1988]

- Evaluation procedure:
  - Assign task scenario to user and robot (in simulation)
  - At some **random point in time**, halt simulation and blank relevant displays
  - Administer **random subset** of questions about SA requirements
    - SA levels: immediate, intermediate, long-range
  - Compare real and perceived situation post-hoc, and report true/false %
  - Repeat for different users to obtain measures of statistical significance



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

Methodology - Embodiment

- Anthropomorphic
  - Many argue that to interact socially with people a robot should resemble a human
- Caricatured
  - Realism is not necessarily needed for believability.
- Functional
  - The embodiment should reflect the tasks it must perform.
- Zoomorphic
  - Most common are "pet" type robots
  - Human-creature relationships are simpler than human-human relationships
  - Easier to avoid the "uncanny valley“ in previous slide

Human-Oriented Perception

- To interact with humans in the real world, social robots must perceive the world the same way that humans do
- In particular, they must be able to track human features and interpret human communication
- Similar perception may require similar sensing

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

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:

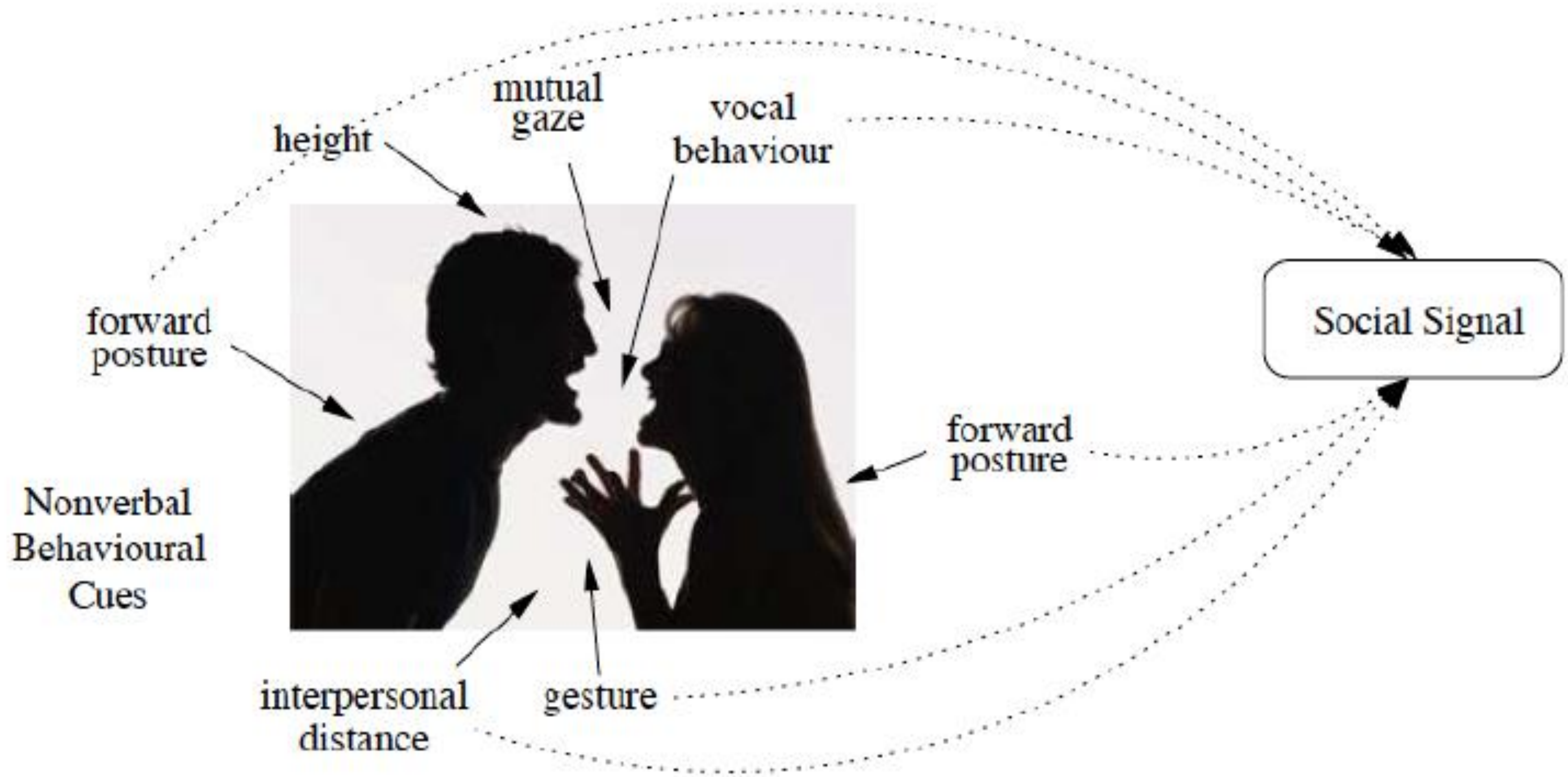
- How can robots determine the right emotion ?
- How can “emotions” be expressed by a robot ?

Social Interaction

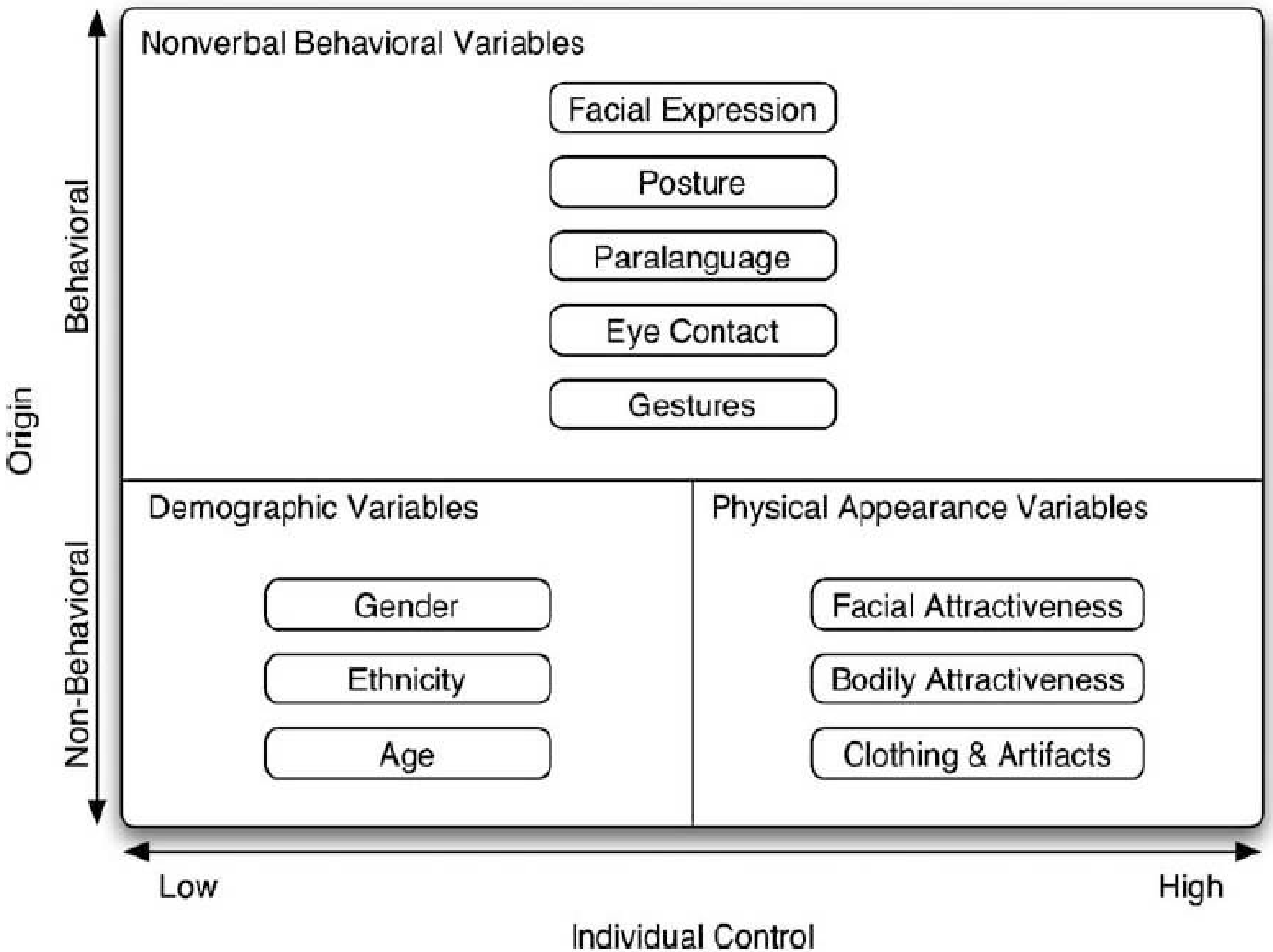
- In order to interact socially with a human, a robot must convey intentionality, that is, the human must believe that the robot has beliefs, desires, and intentions.

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

Social Signal



Social Signals



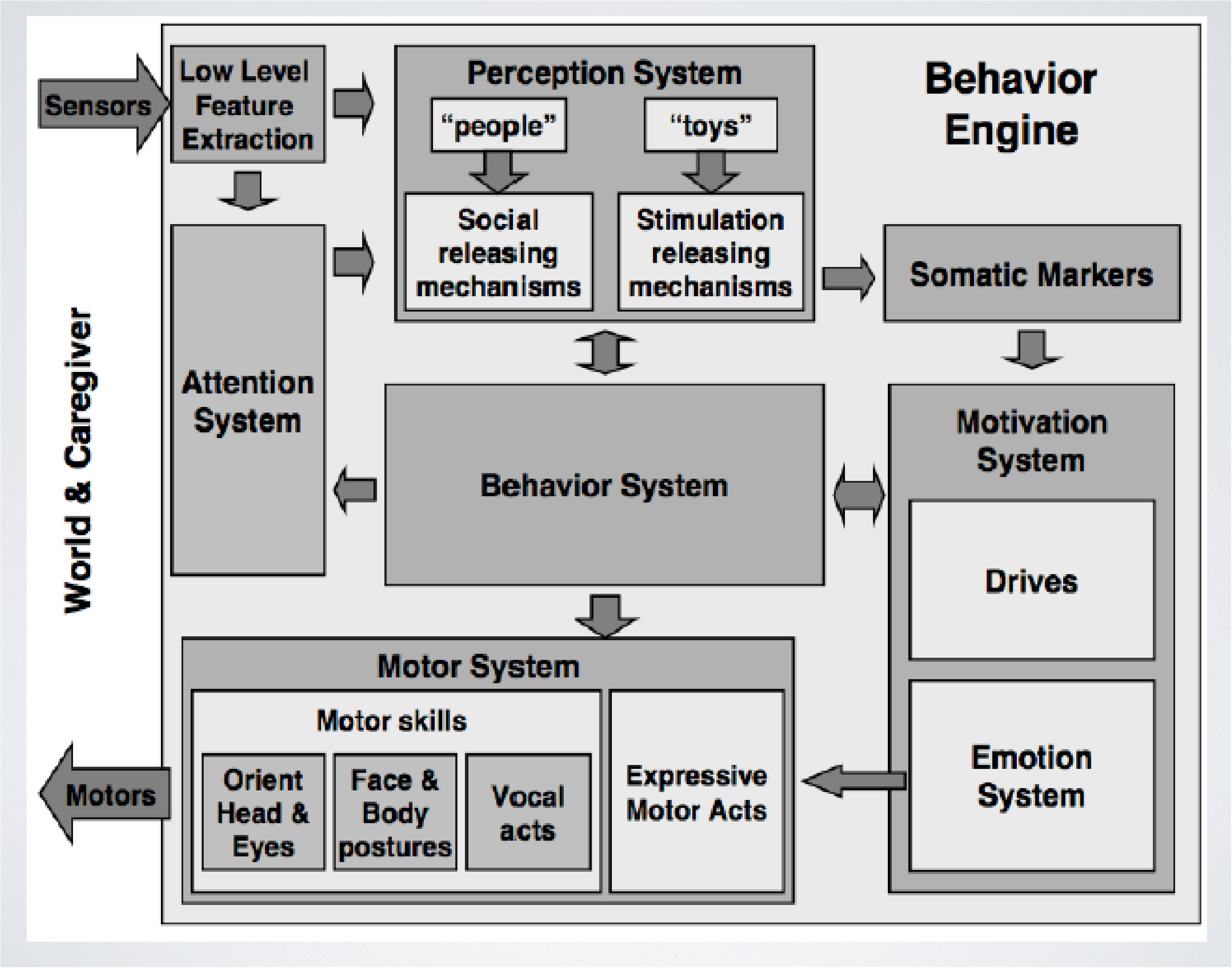
DeMeuse (1987)

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

Social Interaction



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

Attention

In robotics, attention is currently mostly used

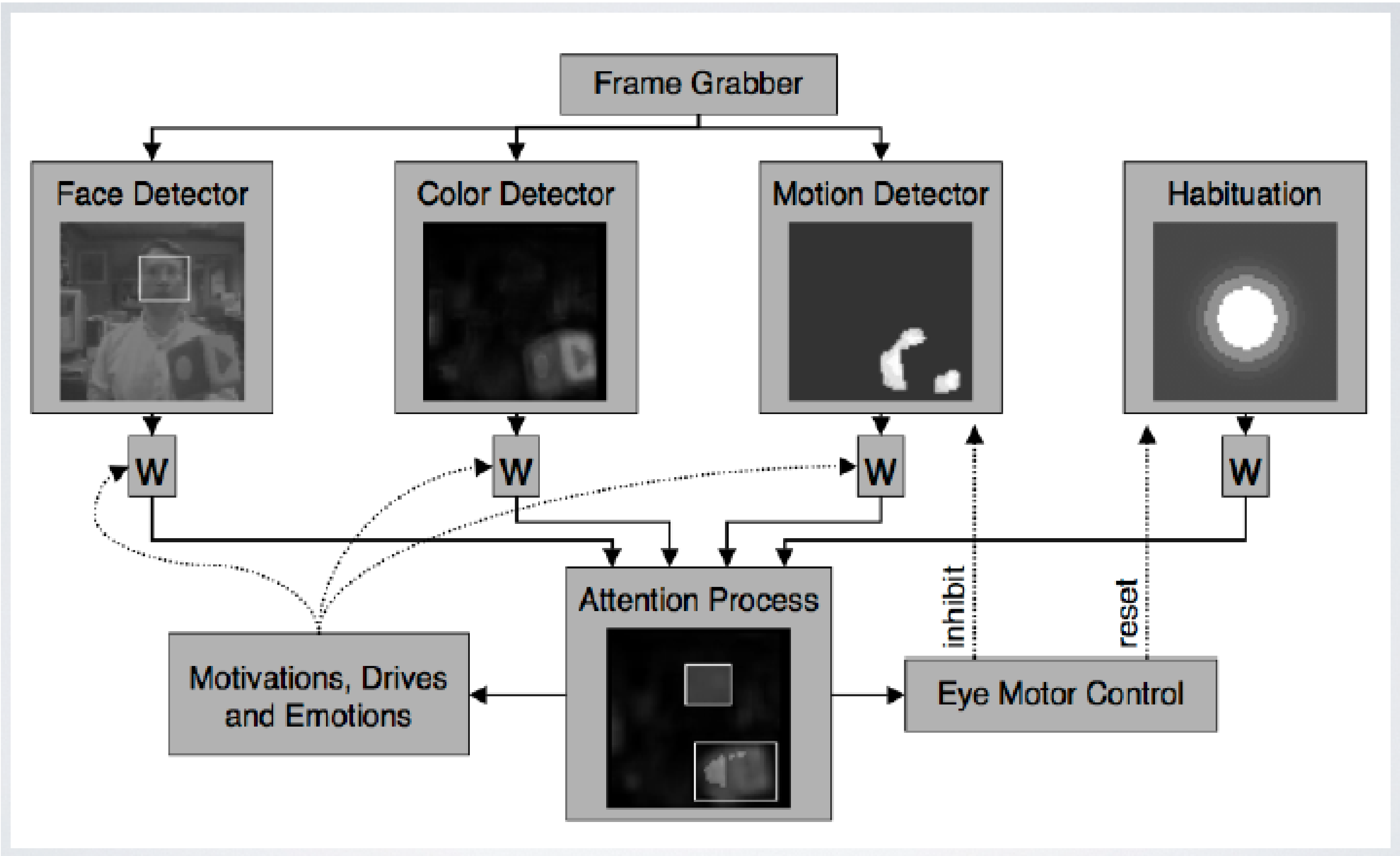
– as salient interest point detector

– as front-end for object recognition

– to guide robot action

– to establish joint focus of attention of human and machine (human-computer interaction, human-robot interaction)

Attention



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

Attention

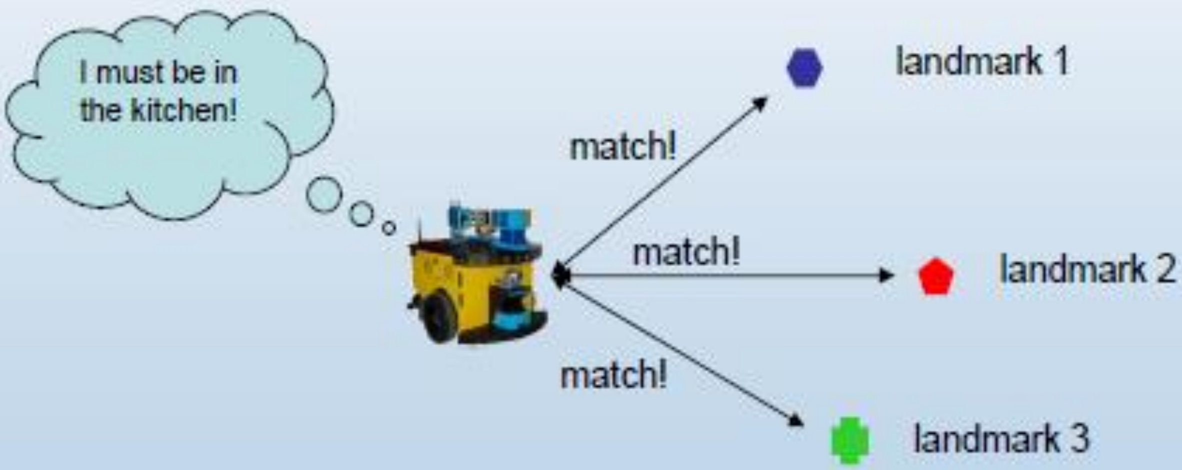
(Visual) attention is especially useful for autonomous robots, since they have similar requirements as humans:

- They can process only a fraction of the perceptual input in reasonable time (real-time processing wanted)
  => attention prioritizes
- They have physical constraints (one/few cameras for zooming and pan/tilt, one/few arms,…) and have to decide what to do next
  => attention supports decision making
- Many robots act in the same environments as humans and shall interact with them
  => joint/shared attention useful

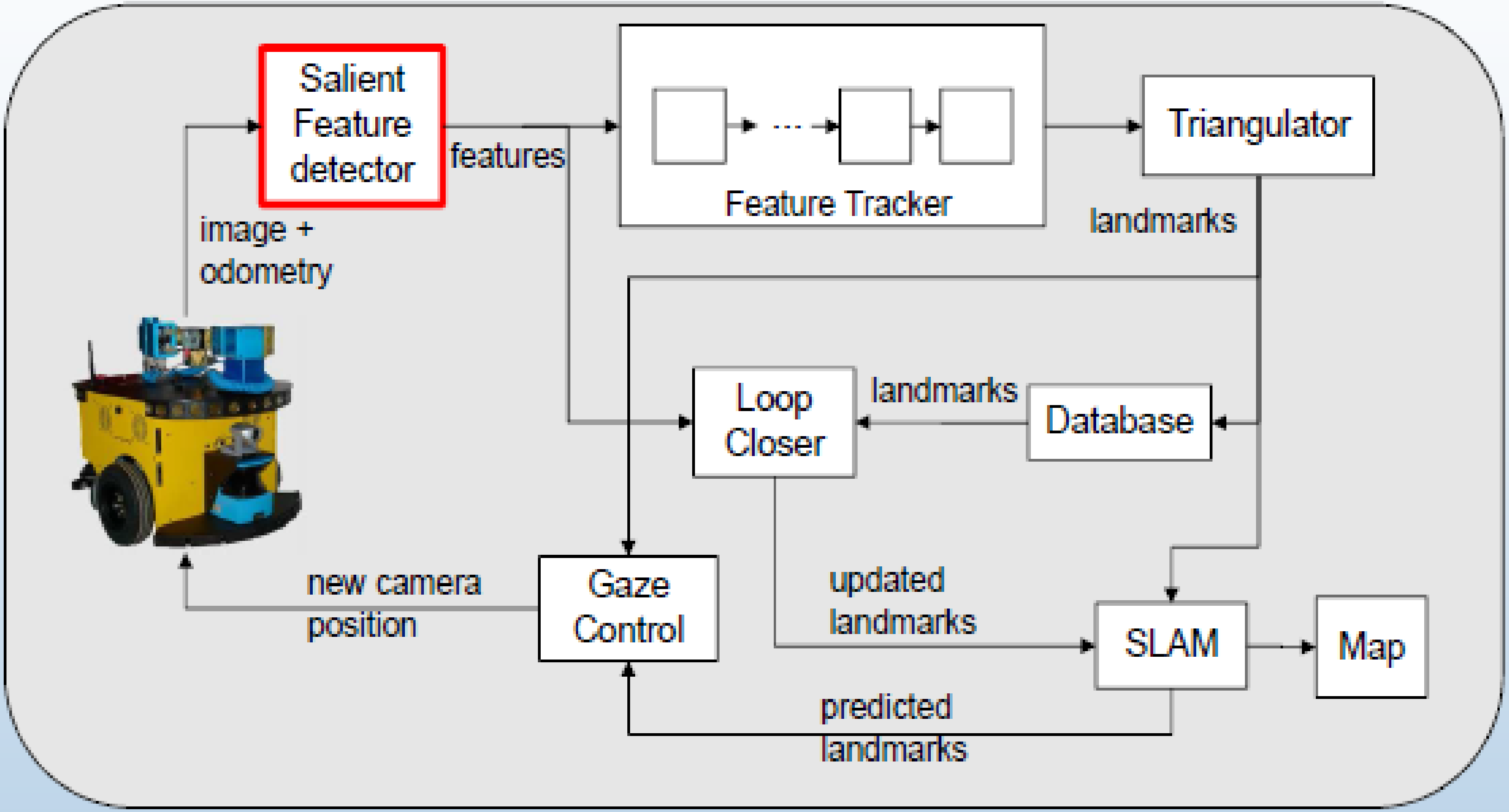
Attention

Visual localization: use landmarks to determine robot position

Salient landmarks: enable sparse landmark representation and quick and reliable redetection



Attention



Salient Feature detector
features
Feature Tracker
...
Triangulator
landmarks
image + odometry
Loop Closer
landmarks
Database
new camera position
Gaze Control
updated landmarks
SLAM
Map
predicted landmarks

Attention

Attention to guide robot action:

– active vision (control the camera and robot head)
– robot navigation (e.g. for finding a goal or following people)
– object manipulation (grasping, pushing, …)
– human-robot interaction

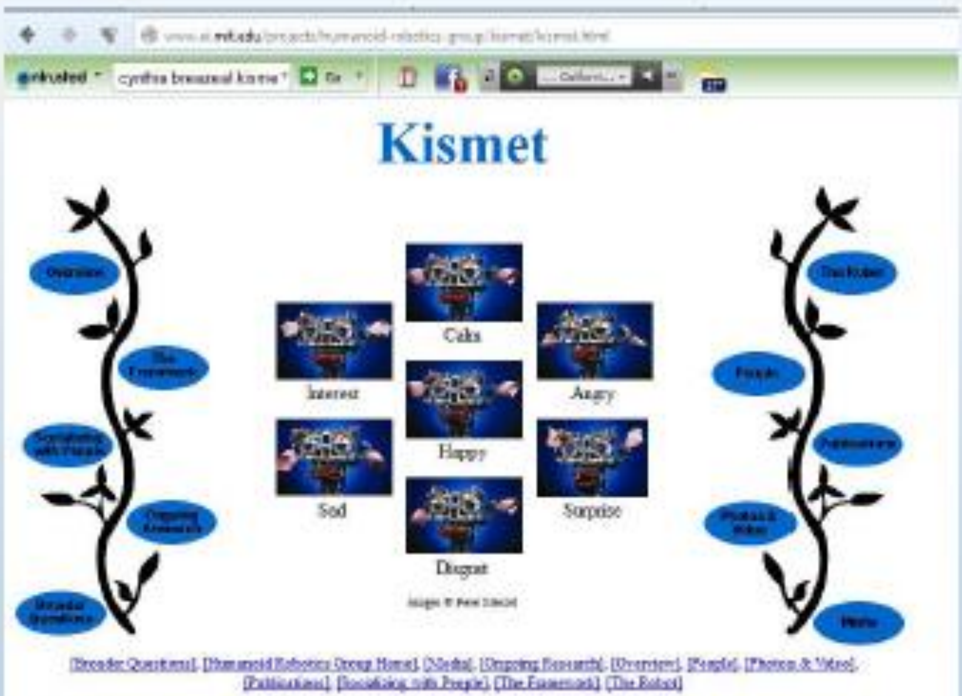
Attention

- Shared attention (also joint attention) means that an "agent" (human or artificial agent) shares the attention with another agent. I.e. they look at the same object, talk about the same object
- Shared attention is an important aspect of human-robot interaction
- Sharing attention mean to
  - Recognize what the partner is attending to (follow partner's gaze, interpret gestures and language, use context information, etc.) Bu and td attention can help here
  - Give feedback that you are paying attention to what the partner does (follow objects of interest with gaze, use mimics, sound etc. to express surprise, excitement, etc.)

Human-robot interaction at the university of Bielefeld

Attention

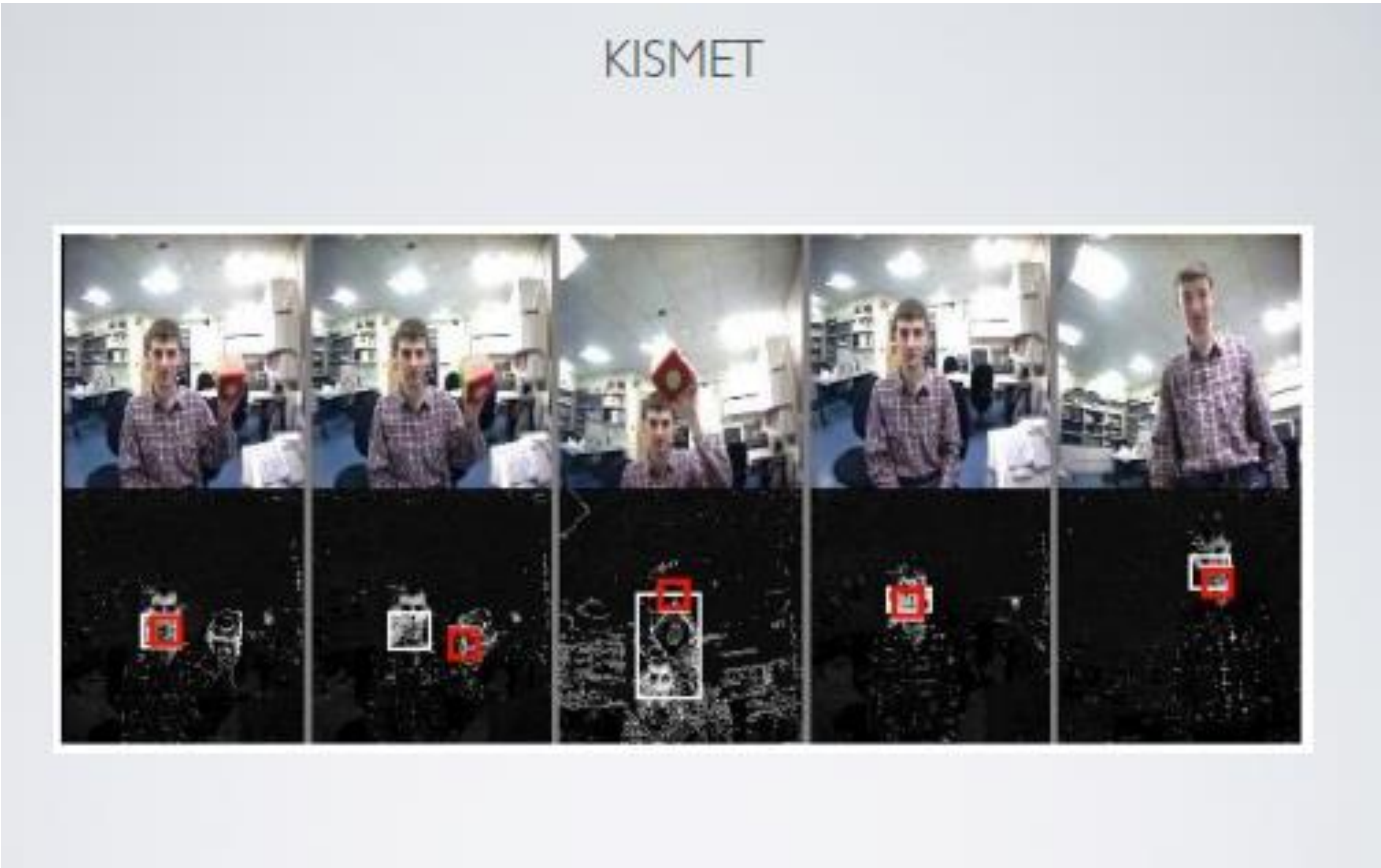
- Robot Kismet: one of the first robots with shared attention (group of Cynthia Breazeal at MIT)
- It interacts with humans in a natural and intuitive way



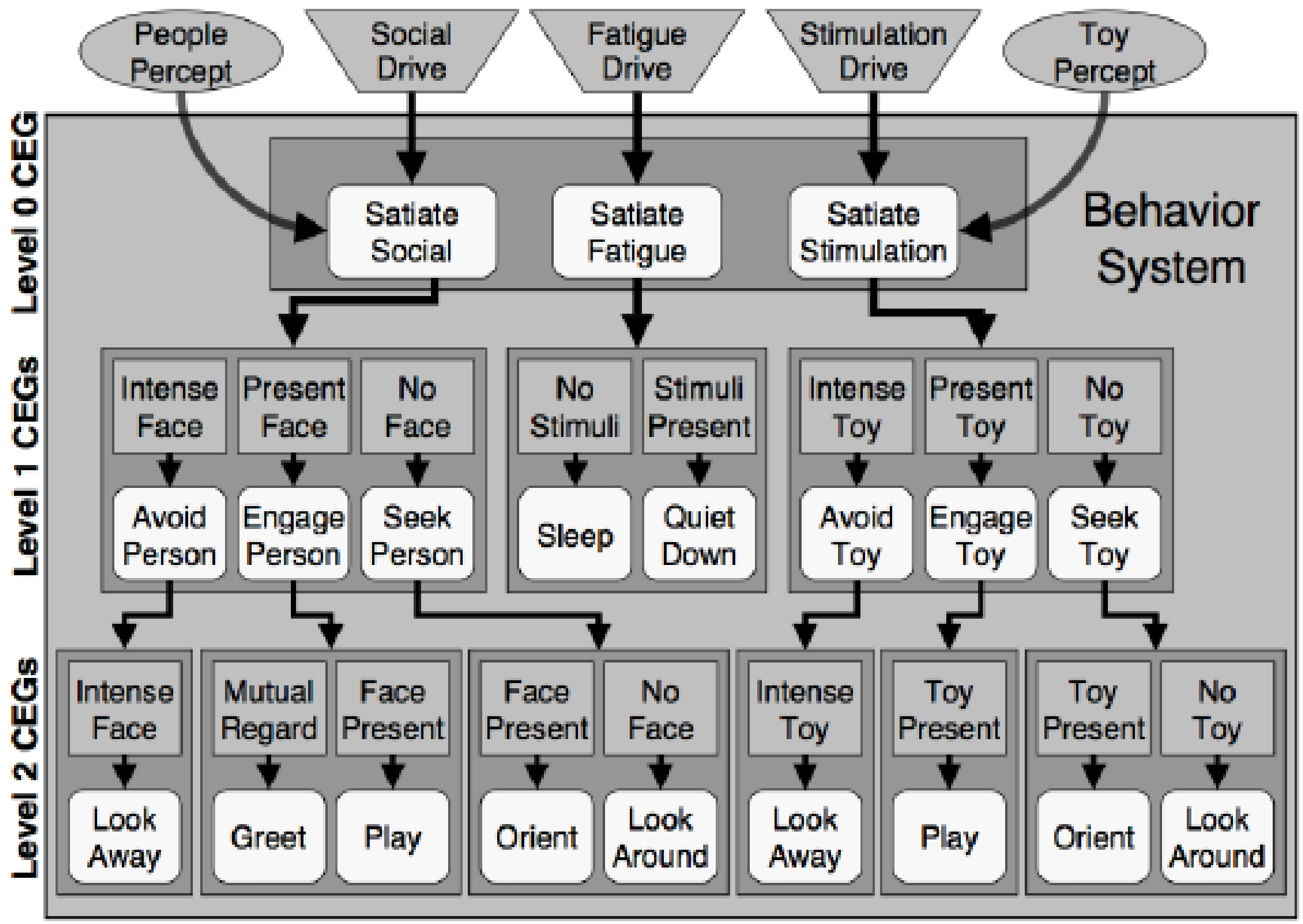
© Sam Ogden)

- http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html

Attention



Behaviors



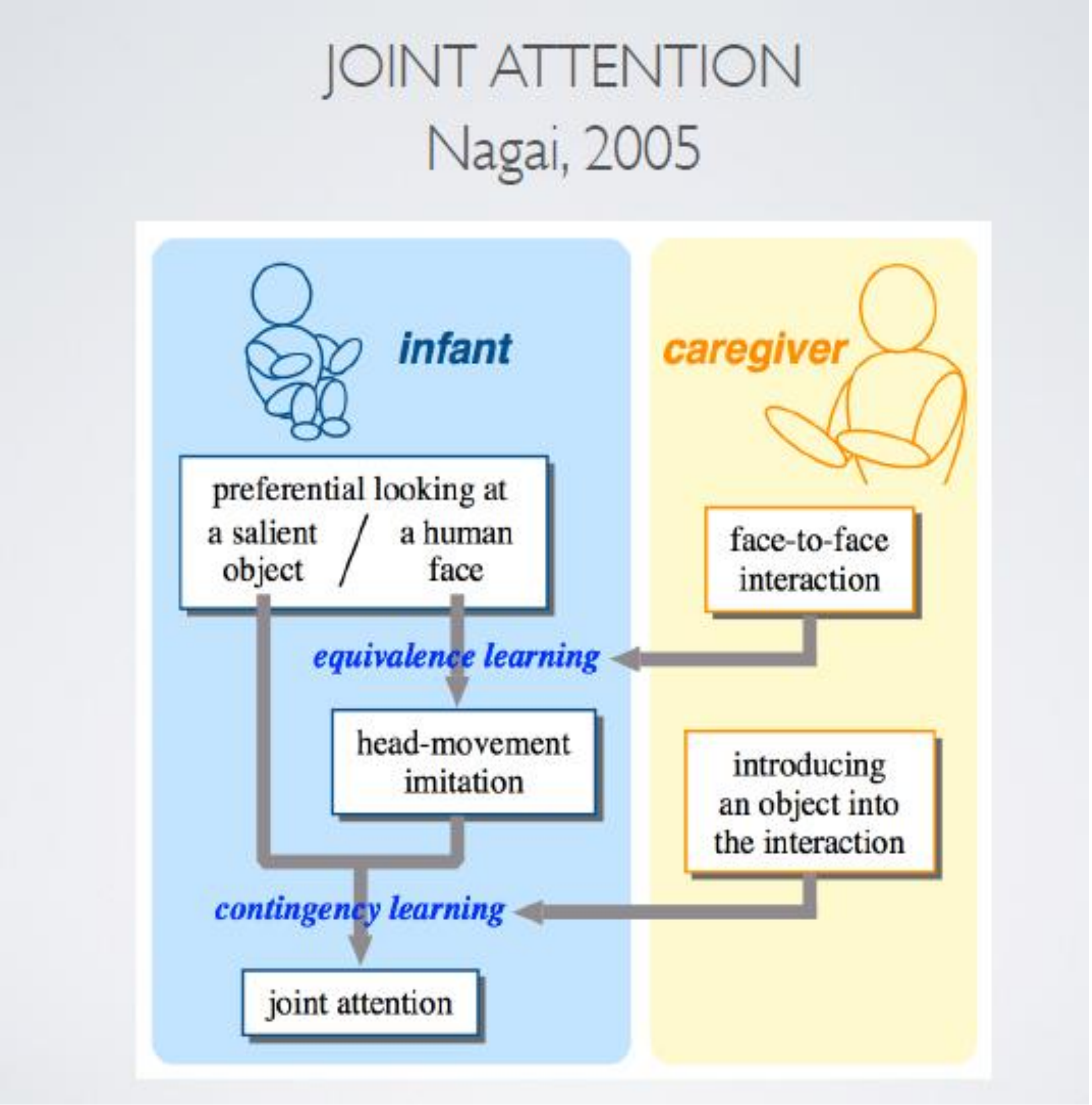
People Percept
Social Drive
Fatigue Drive
Stimulation Drive
Toy Percept
Level 0 CEG
Satiate Social
Satiate Fatigue
Satiate Stimulation
Behavior System
Level 1 CEGs
Intense Face
Present Face
No Face
Avoid Person
Engage Person
Seek Person
No Stimuli
Stimuli Present
Sleep
Quiet Down
Intense Toy
Present 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

Joint Attention

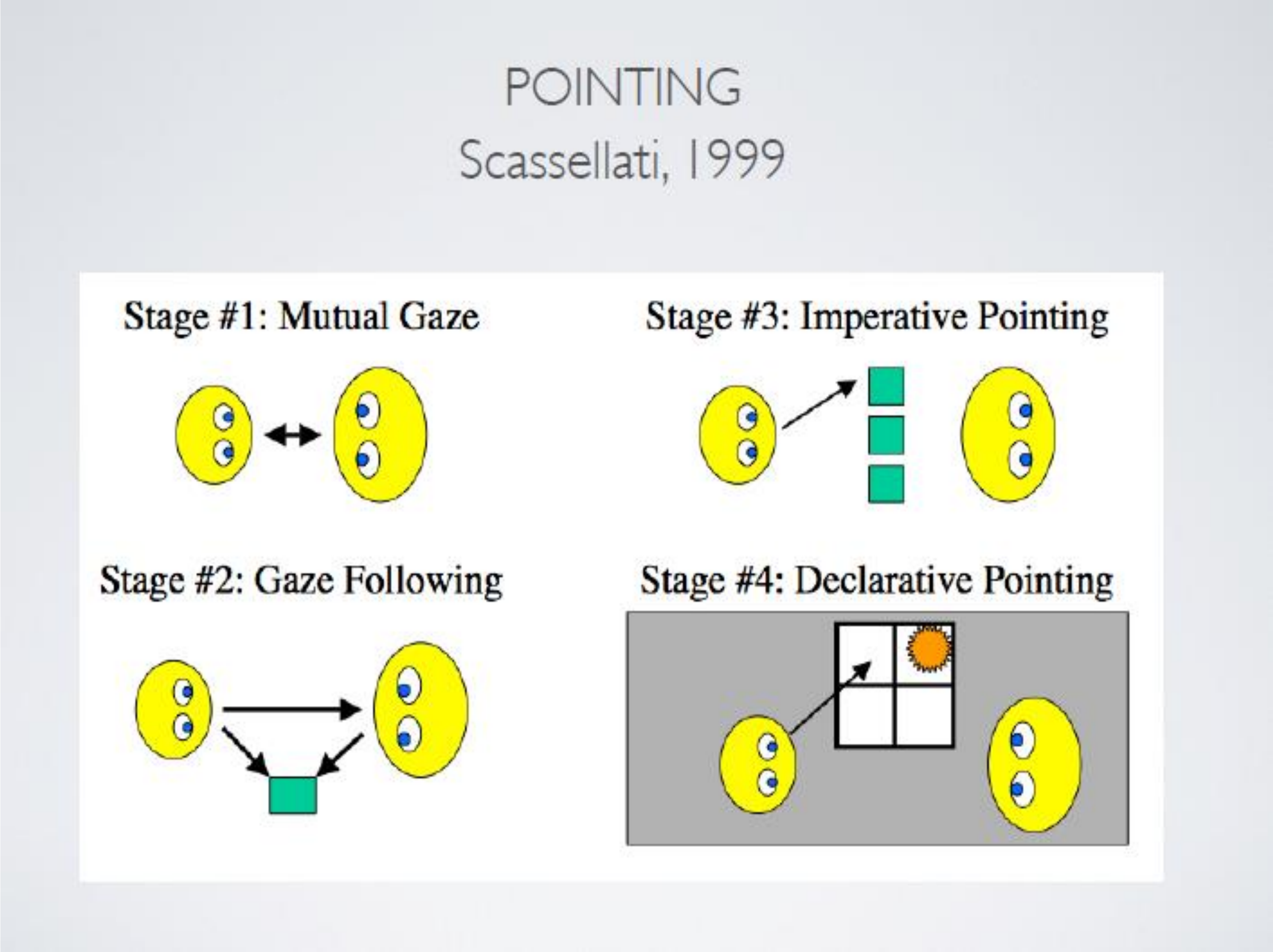
JOINT ATTENTION
Nagai, 2005

- an infant-like robot first has the experiences of visually tracking a human face based on the ability to preferentially look at salient visual stimuli
- imitation of head movements
- changes from tracking the human face to looking at an object which the human is looking at
- the robot learns joint attention behavior based on the contingency between the head movement and the object

Joint Attention



Joint Attention



Joint Attention

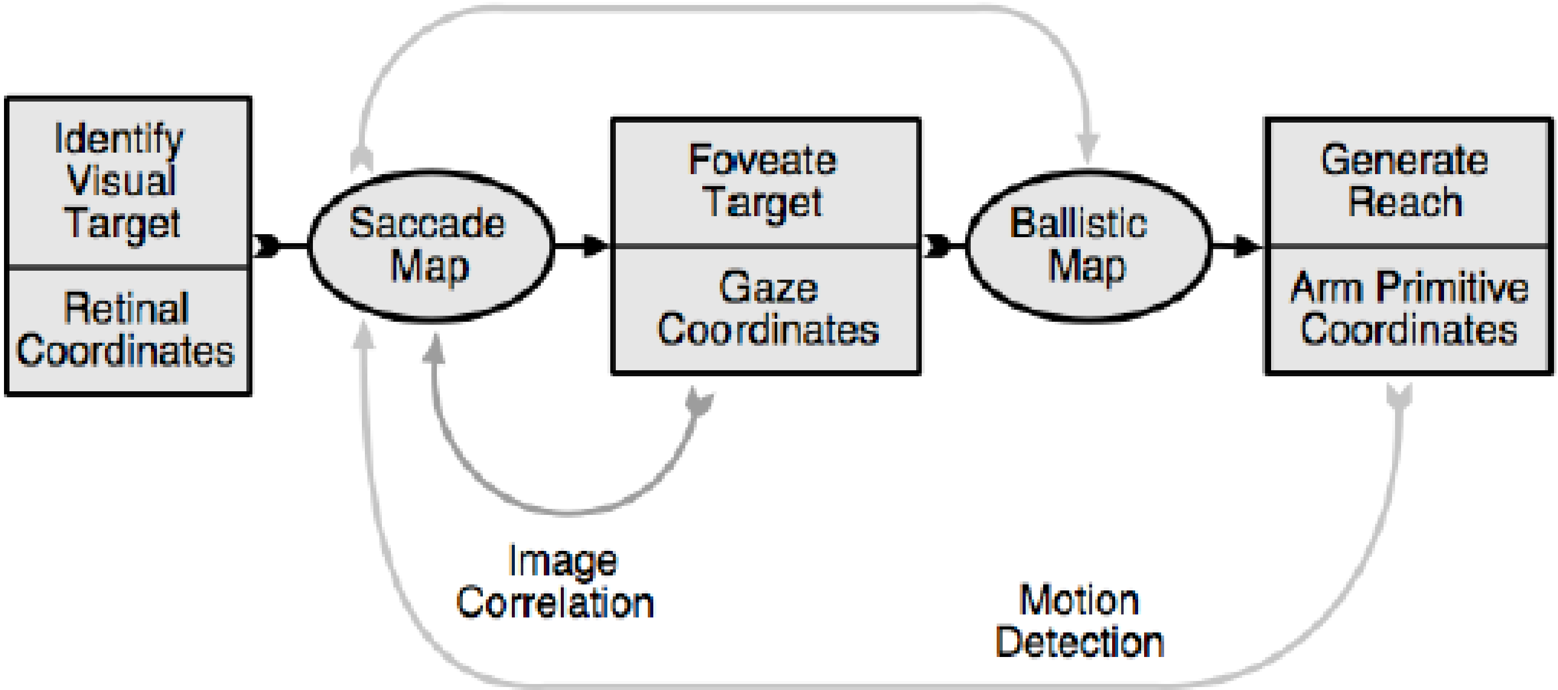


Fig. 9. Reaching to a visual target is the product of two subskills: foveating a target and generating a ballistic reach from that eye position. Image correlation can be used to train a saccade map which transforms retinal coordinates into gaze coordinates (eye positions). This saccade map can then be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

Joint Attention

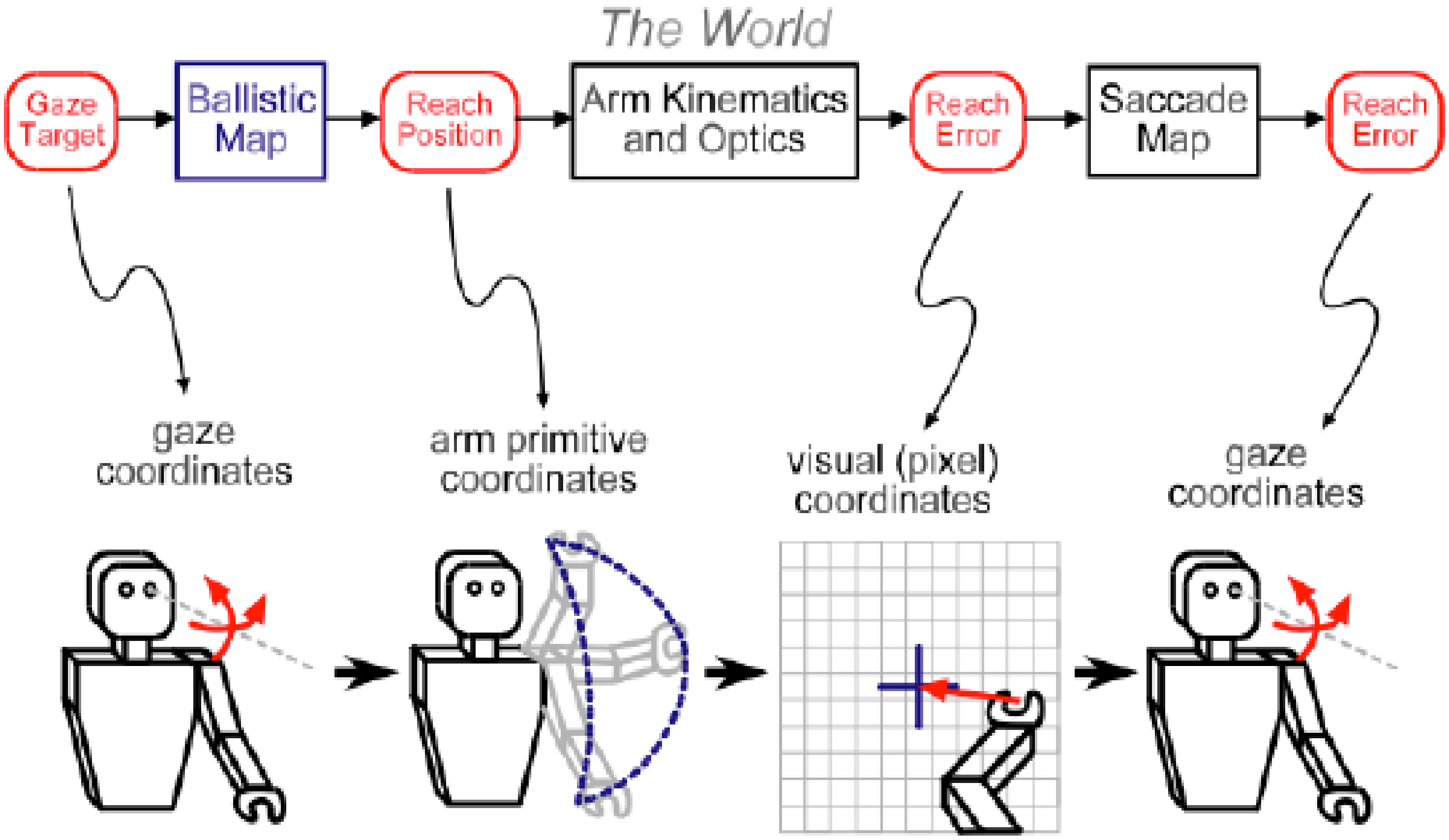


Fig. 10. Generation of error signals from a single reaching trial. Once a visual target is foveated, the gaze coordinates are transformed into a ballistic reach by the ballistic map. By observing the position of the moving hand, we can obtain a reaching error signal in image coordinates, which can be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

Social Interaction

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 ?

Human-Robot Interaction

- Robot Interfaces have to be easy to use
  - Robots have to be controllable by untrained users
  - Robots have to be able to interact not only with their owner but also with other people
- Robot interfaces have to be usable at the human's discretion
  - Human-robot interaction occurs on an irregular basis
    - Frequently the robot has to operate autonomously
    - Whenever user input is provided the robot has to react to it
- Interfaces have to be designed human-centric
  - The role of the robot is it to make the human's life easier and more comfortable (it is not just a tech toy)

Perspective taking
(Trafton et al., 2005)

- Analyzed a corpus of NASA training tapes
  - Space Station Mission 9A
  - Two astronauts working in full suits in neutral-buoyancy facility. Third, remote person participates.
  - Standard protocol analysis techniques; transcribed 8 hours of utterances and gestures (~4000 instances)
    - Use of spatial language (up, down, forward, in between, my left, etc) and commands
  - Research questions:
    - What frames of reference are used?
    - How often do people switch frames of reference?
    - How often do people take another person's perspective?

Perspective Taking and Changing Frames of Reference
(Trafton et al., 2005)

Bob, if you come straight down from where you are, uh, and uh kind of peek down under the rail on the nadir side, by your right hand, almost straight nadir, you should see the uh…

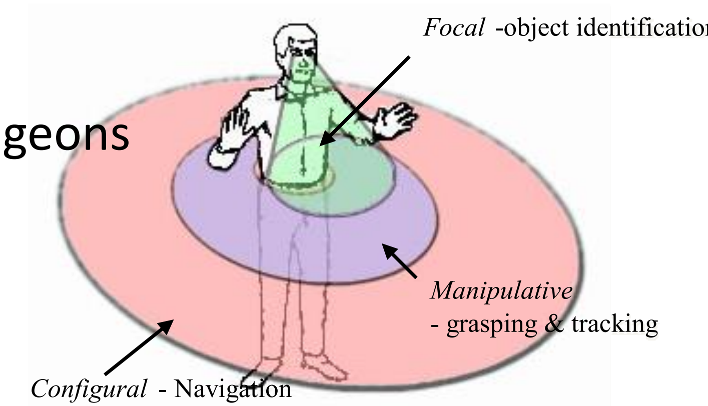
- Notice the mixing of perspectives: exocentric (down), object-centered (down under the rail), addressee-centered (right hand), and exocentric again (nadir) all in one instruction!
- Notice the “new” term developed collaboratively: mystery hand rail

Perspective taking in human interactions

- Ambiguous references that involve different spatial perspectives: (Clark, 96)
  - Principle of least effort (which implies least joint effort)
    - All things being equal, agents try to minimize the effort
  - Principle of joint salience
    - Solution that is the most salient, prominent, or conspicuous with respect to their current common ground.

Perspective Taking: Two systems

- ACT-R/S (Schunn & Harrison, 2001)
  - Perspective-taking system using ACT-R/S is described in Hiatt et al. 2003
    - Three Integrated VisuoSpatial buffers
    - Focal: Object ID; non-metric geon parts
    - Manipulative: grasping/tracking; metric geons
    - Configural: navigation; bounding boxes



- Polyscheme (Cassimatis)
  - Computational Cognitive Architecture where:
    - Mental Simulation is the primitive
    - Many AI methods are integrated
  - Perspective-taking using Polyscheme is described in Trafton et al., 2005