Humans Interacting with Robots: For the Worse or For the Better?

A Reflection on Ethical Issues

Raja Chatila

LAAS-CNRS

Toulouse, France

Raja.Chatila@laas.fr

Journée GDR Robotique et Ethique
Paris, 26-05-2010

Ethics

- Ethics is about morality and respect of values such as safety, freedom, dignity, independence, privacy, …

- Researcher/Engineer ethics: design of machines intrinsically respecting those values

Ethics and Robotics

- Robotics:
  - Operational Autonomy (sensing and motion):
    ➔ operational morality (by design…)

  - Decisional Autonomy (interpretation, decision-making, planning): ➔ decisional morality

- Complexity of perception and decision-making processes and situations. How to make decisions respecting moral values?

Focus on Robots Interacting With People

- Abilities: **physical** and **cognitive** interaction with people
- Requirements:
  - Safety
  - Dignity
  - Comfort
  - Acceptability
  - Legibility
  - Predictability
  - Dependability
  - Adaptation to users

HRI should address all these problems

Some Ethical Concerns in Robotics and HRI

- Robot behavior in presence of humans;

- Responsibilities in cooperative human-machine deliberation and action;

- Human behavioral ethics toward robots (and implications towards humans);

- Individual and societal impact of human-machine cognitive and affective bonds;
- Impact on Human identity and nature.

Awareness of robots about humans

- Detecting and understanding human actions

- Models of humans, physical, behavior, preferences

- Decisional Interactions

Understanding Humans; Human Models

- Physiology
  - Geometrical model (shape, constraints, workspace)
  - Kinematic model (speed, acceleration, limits)
  - Perceptual capacities, field of view
- Physical activity
  - Postures
  - Gestures
  - Actions
- Psychology
  - Social spaces
  - Comfort
  - Acceptability
  - Moods, Expressions, intentions and emotions
- User-Specific adaptation
  - Physical capacities
  - Personal habits, recurrent behaviors
  - Preferences

Detecting, tracking People and Gestures
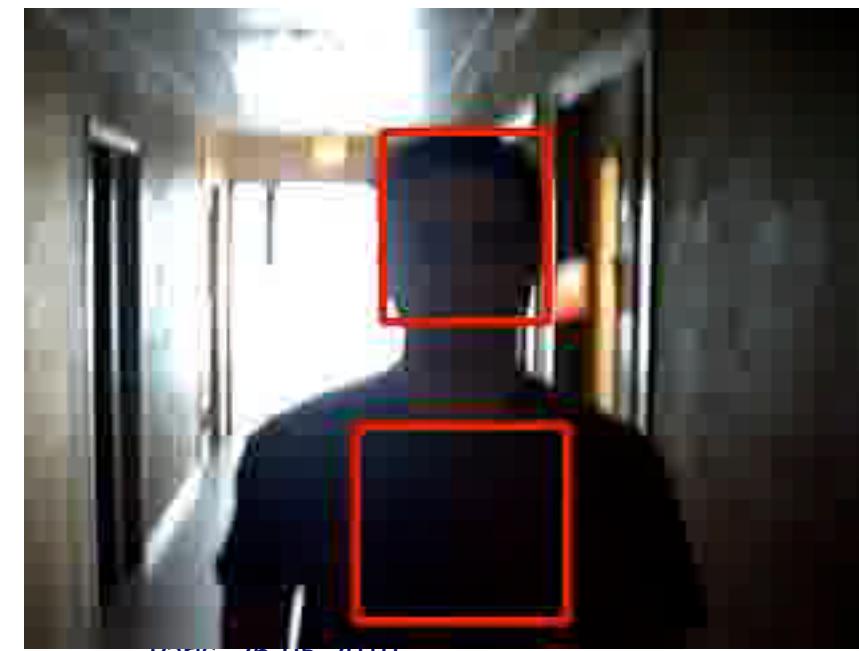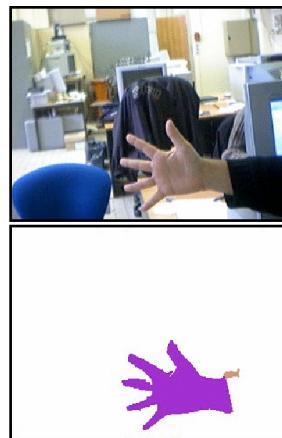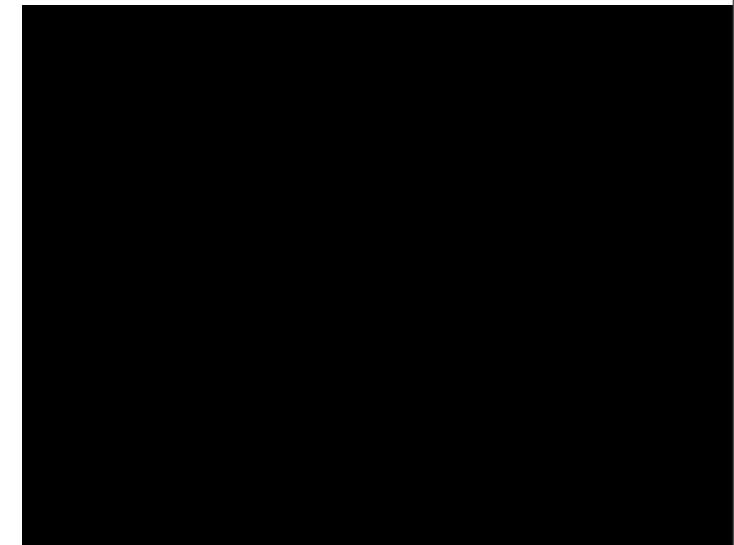
Intrinsic Physical Safety

III. Impact Experiments LWRIII - Human

Intrinsic Physical Safety

$$\text{HIC}_{36} = \max_{(\Delta t)} \left\{ (\Delta t) \left(\frac{1}{\Delta t} \int_{t_1}^{t_2} \|\ddot{\mathbf{x}}_H\|_2 \, dt \right)^{\left(\frac{5}{2}\right)} \right\} \leq 650$$

$$\Delta t = t_2 - t_1 \leq \Delta t_{\max} = 36\text{ms}.$$

Injury Level ← **HIC Level**

Injury Indicator   **Head Injury Criterion**

Very high

20% AIS ≥ 3

High

Medium

Low

5% AIS ≥ 3

Very low

1000

serious, but not life threatening

650
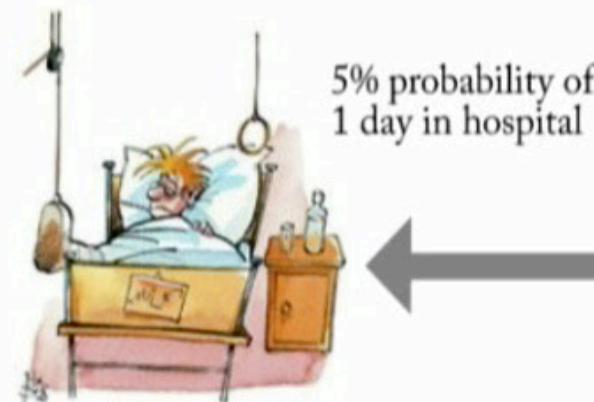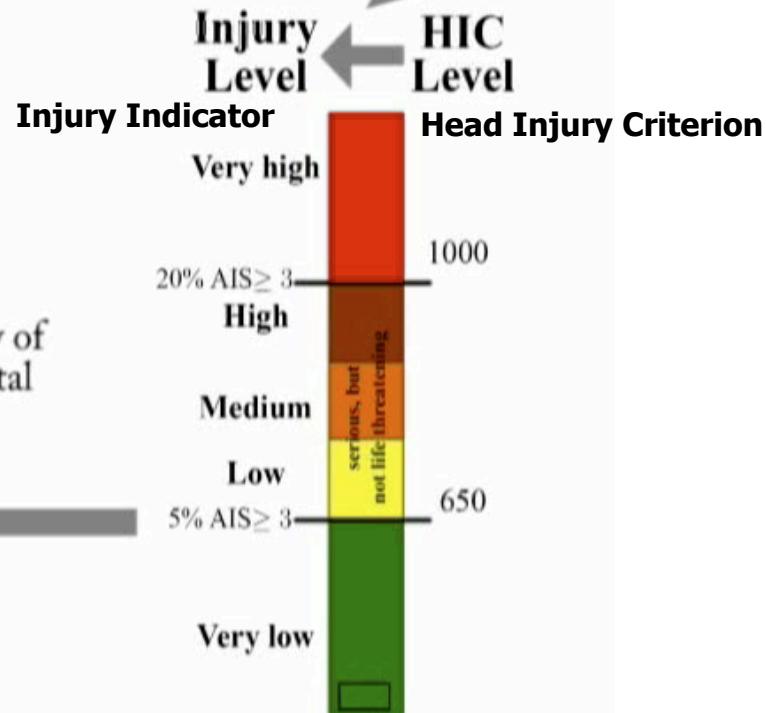
5% probability of 1 day in hospital

Interacting closely with People
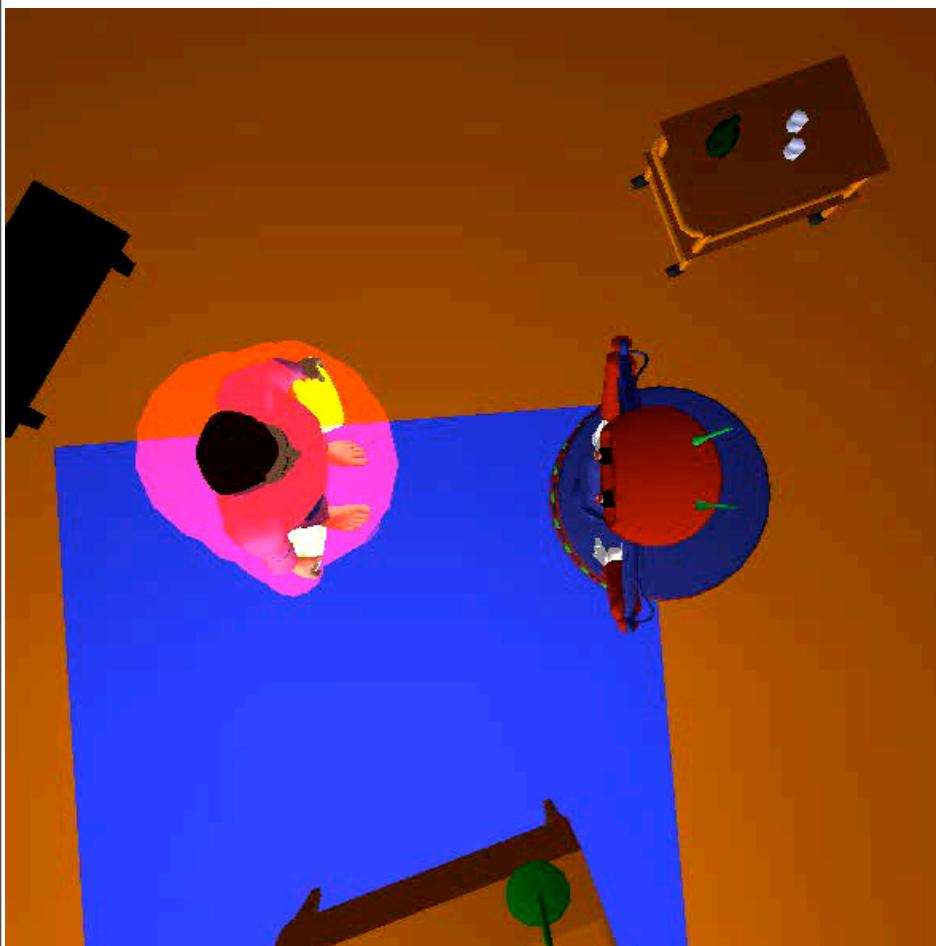
Safety

Comfort

Detecting, Recognizing and Understanding People

Cogniron - The Cognitive Robot Companion

FP6-002020

-

Tutors' Face Recognition
LAAS - CNRS

Detecting, Recognizing and Understanding People

Cogniron - The Cognitive Robot Companion

FP6-002020

-

Tutors' Face Recognition
LAAS - CNRS

Motion Analysis for Activity Recognition

- Based on articulated body pose tracking

- Extraction of >300 features from body motion

- Activity scenarios

- Relevant features per activity

- Statistical learning

Distance of reference points

Joint angles

Motion direction

Motion periodicity and planarity
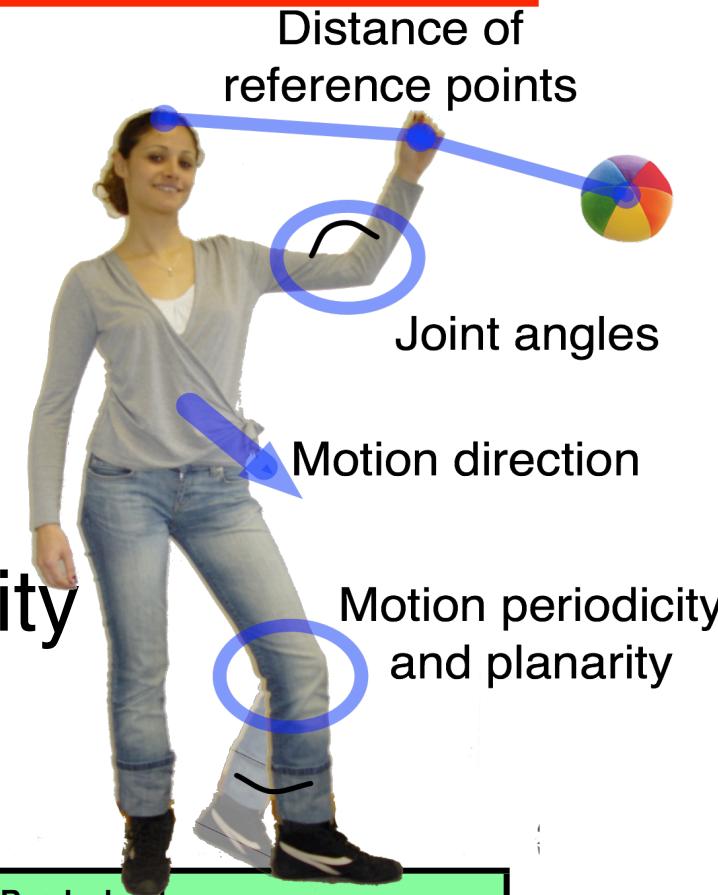
Standing	Hold out object, right hand	Read a book
Absolute angle left thigh	Absolute angle right forearm	Distance hand - book
Absolute angle right thigh	Absolute angle right upper arm	Distance left hand - right hand
Angle left knee	Distance hand - object	Angle left upper arm - torso
Angle right knee	Absolute angle right lower leg	Angle left thigh - torso

Understanding Human Actions

Humans' Safety and Preferences

- Human field of view, preferences, motion capacities
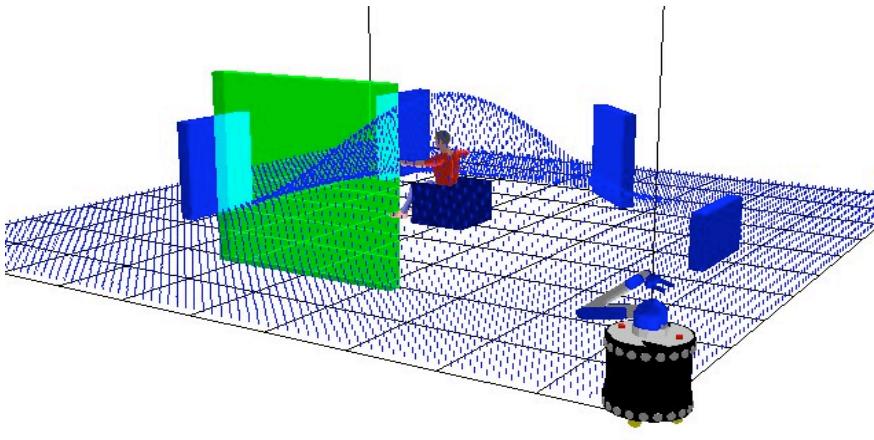
User studies at U. Hertfordshire

Humans' Safety and Preferences

- Human field of view, preferences, motion capacities
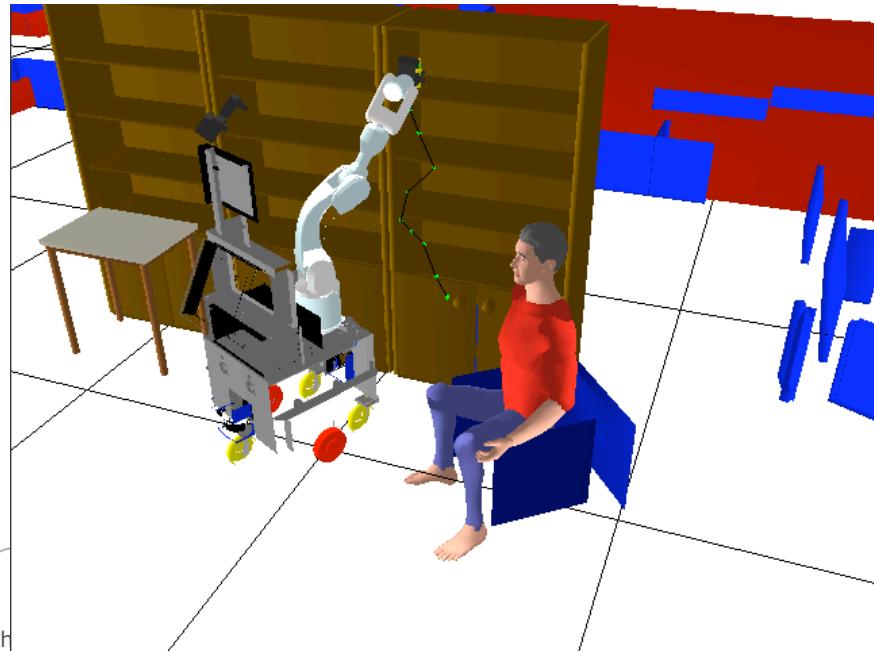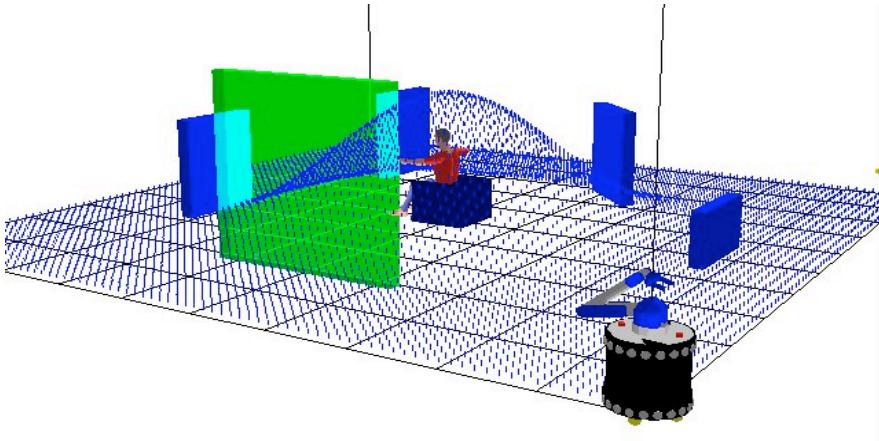
User studies at U. Hertfordshire
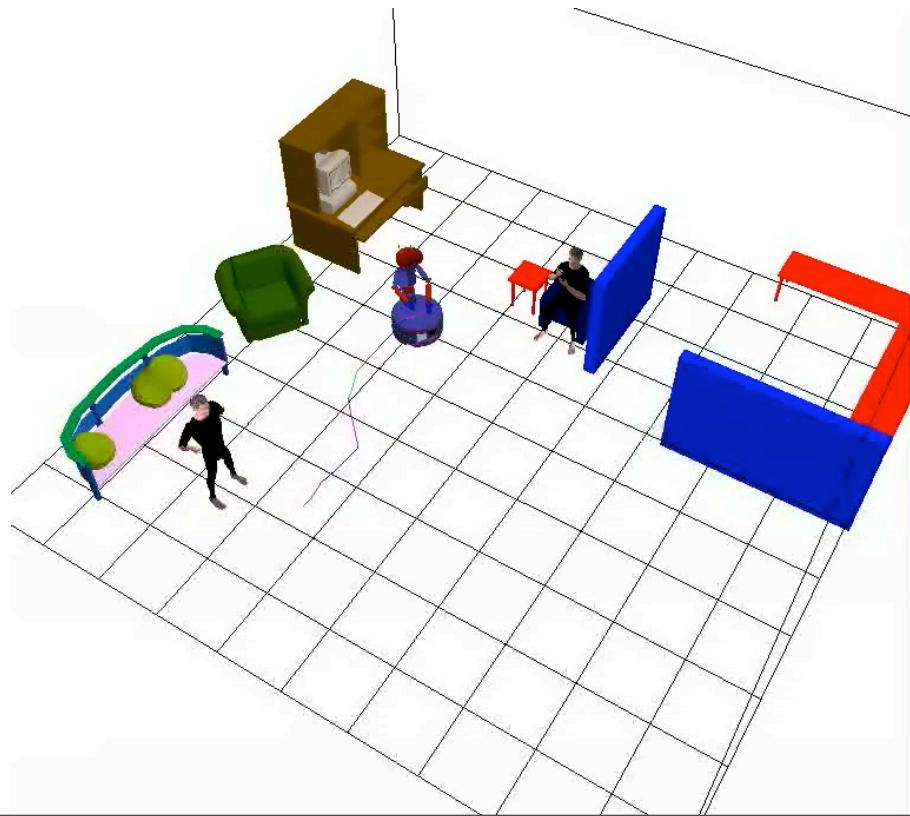
Human-Robot Decisional Interaction

- Joint Intention theory (Cohen-Levesque), Teamwork, Joint Activities (Clark)

- Joint Human-Robot goals: Human robot interaction for task achievement

- Interaction Agents represent humans (state, abilities, preferences...)

- Perspective taking
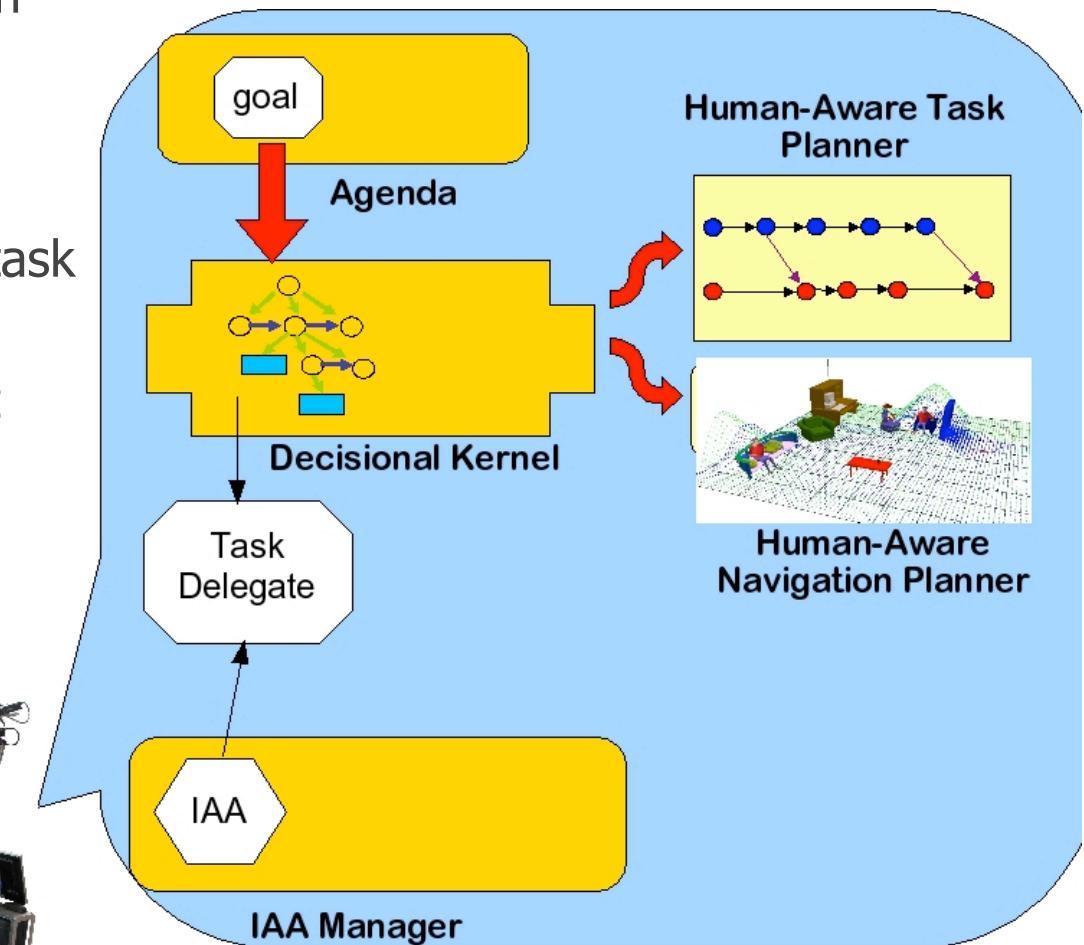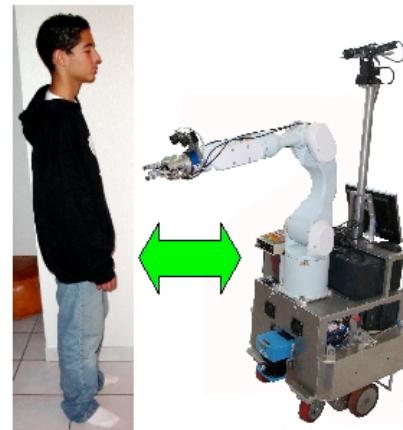
Instantiate IAAs

Establishing common goals, Verifying and reacting to the commitment level of the human partner

Acceptability:
Does appearance matter?

HRP2 AIST

Toyota

Asimo (Honda)
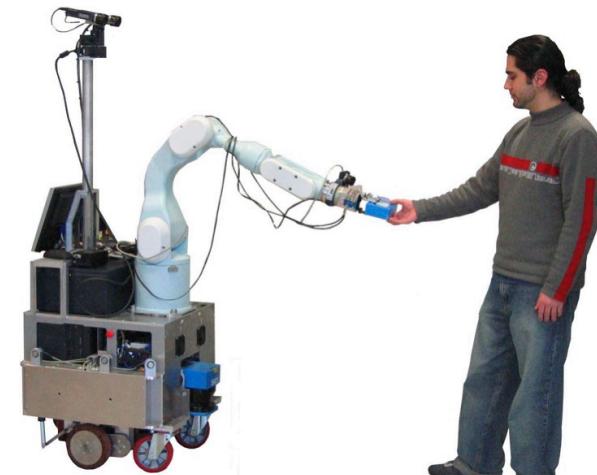
LAAS
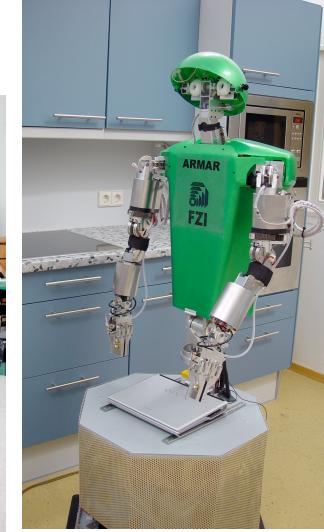
thique

U. Karlsruhe

Shape, Relationships and Ethics

Human behavior with robots (and impact on humans?)

Shape, Relationships and Ethics

U. Osaka

**Human behavior with robots (and impact on humans?)**

Transforming humans

- ## Human augmentation
 - Implants
  - BCIs
  - Exoskeletons
  - Prostheses

SUAW -
DEMAR

Cybernics - HAL

18
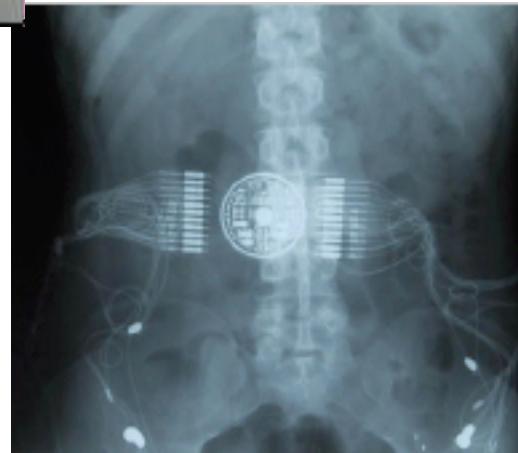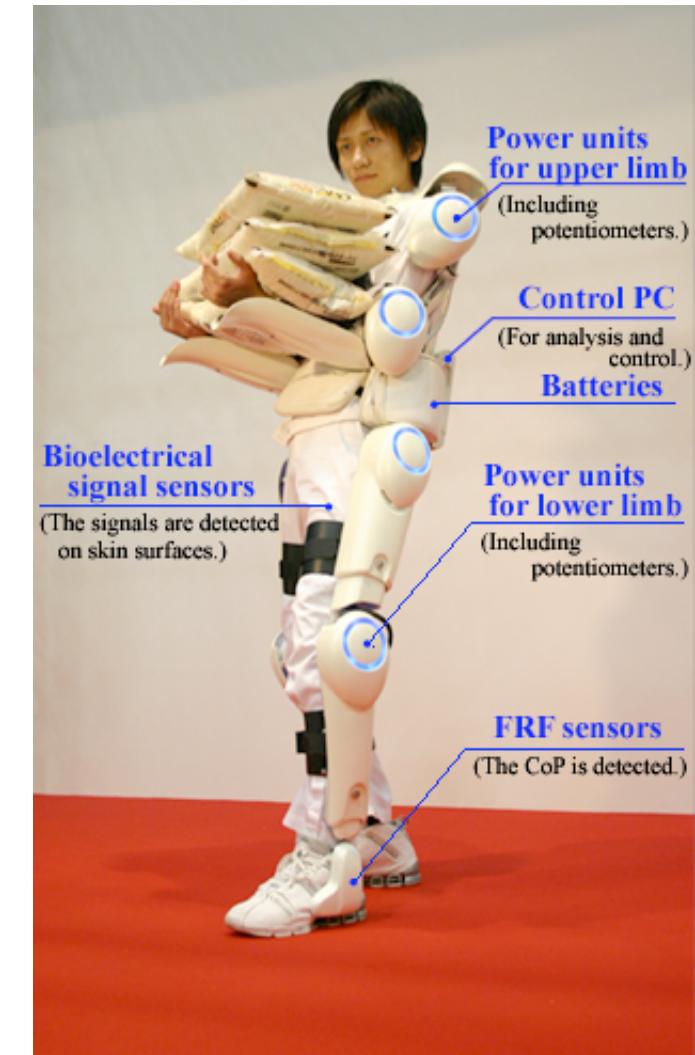
Brain Computer Interfaces

BCI
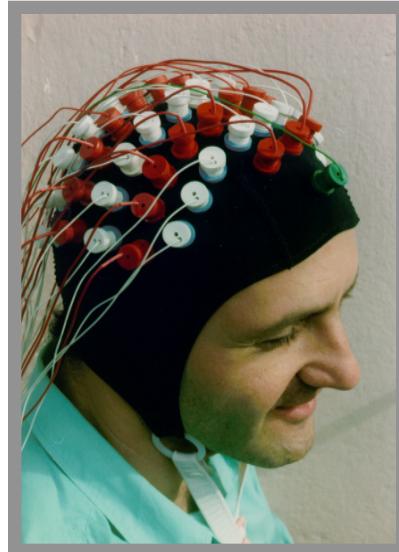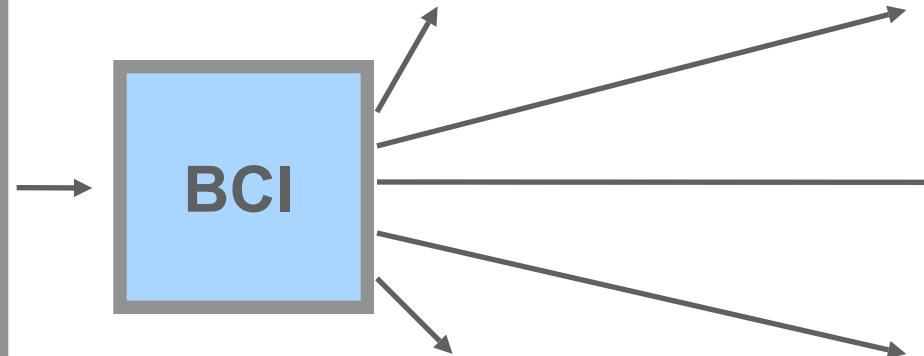
Leeb et al., Computational Intelligence and
Neuroscience, 2007 (doi:10.1155/2007/79642)

Some Conclusions

- The Better? Robots interacting with humans provide:

  - Service to humans
  - Independence of humans (more autonomy)
  - Safety
  - Aids for human disabilities
  - Improve human performance

Some Conclusions

- ## The Worse?

  - Information disclosure, privacy: robots can be connected and interact with the outside world.
  - Unwanted / unsolicited actions and initiatives
  - Human transformation (transhumanism)
  - Moral borders

Precautionary Principle Measures?

- Intrinsically safe physical design and behavior.

- Implement intrinsically safe cognitive behavior and code of conduct.

- Limited knowledge:
  - acquisition/persistence,
  - interpretation,
  - transmission.

- Limited initiative and action capabilities