Social Robots: Research Challenges & Ethical Issues

Carme Torras @ Universiteit Leiden
My standpoint

Affiliation: Institut de Robòtica i Informàtica Industrial

28 Doctors

37 PhD students

10 technicians

14 Support staff

>10 Robots
Outline

1a. Industrial Robotics  ➔ Assistive Robotics

1b. Our assistive projects: CLOTHILDE, I-Dress, Socrates

2a. Ethical and social implications

2b. Education/debate through Science Fiction

Research challenges

Roboethics
Rise of service robots and co-workers
Assistive robots in human environments
Domestic robots
Domestic robot
Humanoid robot
Humanoid robot
Research challenges in Robotic AI

Industrial robots  →  Assistive robots

Industrial robots:
- Programmed by experts
- Caged
- Rigid objects
- Accurate
- Fixed sequences
- Non-interactive

Assistive robots:
- Easy to instruct by non-experts
- Intrinsically safe for people
- Able to perceive and manipulate non-rigid objects
- Tolerant to noisy perceptions and inaccurate actions
- Capable of goal-directed execution
- Collaborating with people
Outline

1a. Industrial Robotics → Assistive Robotics

1b. Our assistive projects: CLOTHILDE, I-Dress, Socrates

2a. Ethic and social implications → Roboethics

2b. Education/debate through Science Fiction

Research challenges for AI

Robotic AI meets the Humanities
Research Team and Lab
CLOTHILDE

CLOTH manipulation Learning from DEMonstrations

housekeeping and hospital logistics

automation in the clothing and internet business

increasing the autonomy of the elderly and disabled
Theory of cloth manipulation based on computational topology

Cloth Perception  Manipulation and Planning  Learning from Demonstration
Usability
Easy instruction by non-experts

Exhaustive **programming** taking into account all situations

Learning from **demonstrations** + **reinforcement** learning
Safety

Control based on a model of robot dynamics
Capable of goal-directed execution
From associations to understanding

Associative learning (perception → motion)

Attaching **semantics** to perceptions of objects and situations
Reasoning about **functionalities** and goals
Collaborating with people & customizable
Outline

1a. Industrial Robotics → Assistive Robotics

1b. Our assistive projects: CLOTHILDE, I-Dress, Socrates

2a. Ethical and social implications

2b. Education/debate through Science Fiction
Ethical and social implications
Assistive robots

Issues shared with other technologies:
- incidence on the job market
- legal liability
- privacy
- digital gap
- ...

New issues in entering the domains of:
- decision making
- feelings & relationships
- human enhancement

Roboethics: subfield of applied ethics studying both the positive and negative implications of robotics for individuals and society.

1. Human ethics applied to robotics
2. Codes of ethics embedded in the robots themselves ("machine ethics")
Ethical and social implications

Two types of human-robot ties:

**Physical**
Exoskeletons, artificial arms and hands, sensorized dresses, telepresence, artificial retinas, implants... robotic prostheses extend the body and potentially the autonomy of people with disabilities.

**Relational**
Living with artificial nannies, learning from robotic teachers, sharing work, leisure and affection with humanoids... will enhance the intellect and social habits of people? Will help develop new capabilities?
Ethical and social implications
Prosthesis / Human-machine interfaces

Some distinctions:

- Alleviate disability / Improve capacities / New capacities

Third arm
Infrared vision
Ethical and social implications
Prosthesis / Human-machine interfaces

Some distinctions:
- Alleviate disability / Improve capacities / New capacities
- No connection / with connection to the nervous system (peripheral / central)
Ethical and social implications
Prosthesis / Human-machine interfaces

Some distinctions:
- Alleviate disability / Improve capacities / New capacities
- No connection / with connection to the nervous system (peripheral / central)
- Individual repercussion ([brain plasticity](#) / psychological) / social

Second thumb
Ethical and social implications
Relational ties: domestic robot... and pet!
Roboethics
Regulations and standards


• RoboLaw project “Regulating Emerging Robotic Technologies in Europe: Robotics facing Law and Ethics”: deliverable D6.2 - Guidelines on Regulating Robotics (2014)

• euRobotics project “The European Robotics Coordination Action”: deliverable D3.2.1- Ethical Legal and Societal issues in Robotics (2012)

2016-18: IEEE Standards Association opened a draft on Ethically Aligned Design to public discussion.
Roboethics
Education and dissemination

The ACM/IEEE Computer Science Curricula consists of 18 knowledge areas, one of which is: “Social Issues and Professional Practice” that includes courses related to Ethics in Technology.
Outline

1a. Industrial Robotics  →  Assistive Robotics

1b. Our assistive projects: CLOTHILDE, I-Dress, Socrates

2a. Ethical and social implications

2b. Education/debate through Science Fiction
Role of Science Fiction
Ethics education in Computer Science and Engineering

1. Anticipate possible future scenarios

“What SF stories can do better than almost anything else is to provide not just an idea for some specific technical innovation, but also to supply a coherent picture of that innovation being integrated into a society, into an economy, and into people’s lives.”

[Neal Stephenson, 2011]

2. Engage technology students

“Using fiction to teach ethics allows students to safely discuss and reason about difficult and emotionally charged issues without making the discussion personal.”

[Judy Goldsmith, 2018]
Roboethics

Education/dissemination based on Science Fiction

When trying to establish an ethical debate, disseminate concepts to the general public, or teach a course on roboethics, classic science fiction stories are often used to exemplify possible future conflicting situations.
Modern Science Fiction related to Roboethics

“It is the relationships that we have constructed which in turn shape us”

Robert C. Solomon
“The Passions”
Modern Science Fiction related to Roboethics

https://mitpress.mit.edu/books/vestigial-heart
Modern Science Fiction related to Roboethics

Four items:

• A **novel** about a future society in which people rely on personal-assistant robots to navigate daily life.

• An **appendix** with 24 ethics questions raised by the novel, as well as hints to trigger a debate.

• An **online teacher's guide** for 6-8 sessions on "Ethics in Social Robotics and AI" following the chapters in the novel and including scholarly references for further reading.

• A **100-slide presentation** that teachers can use and extend as desired.

[https://mitpress.mit.edu/books/vestigial-heart](https://mitpress.mit.edu/books/vestigial-heart)
### Teaching materials

**Ethics in Social Robotics**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.</td>
<td>Overview and background</td>
</tr>
<tr>
<td>1.</td>
<td><em>Designing</em> the “perfect” assistant</td>
</tr>
<tr>
<td>2.</td>
<td>Robot <em>appearance</em> and emotion</td>
</tr>
<tr>
<td>3.</td>
<td>Robots in the <em>workplace</em></td>
</tr>
<tr>
<td>4.</td>
<td>Robots in <em>education</em></td>
</tr>
<tr>
<td>5.</td>
<td>Human-robot <em>interaction</em> and human dignity</td>
</tr>
<tr>
<td>6.</td>
<td>Social <em>responsibility</em> and robot morality</td>
</tr>
<tr>
<td>7.</td>
<td>Bibliography and initiatives to follow up</td>
</tr>
</tbody>
</table>
0.2. Ethical theories relevant to robotics

- **Utilitarianism or Consequentialism** — maximizing the number of people that enjoy the highest beneficial outcomes
- **Deontologism** — acting only according to maxims that could become universal laws
- **Virtue ethics** — relying on the moral character of virtuous individuals
- **Social justice** — all human beings deserve to be treated equally and there must be a firm justification in case of mistreatment
- **Common goods** — living in a community places constraints on the individual
- **Religious ethics** — norms come from a spiritual authority
- **Information ethics** — policies and codes for governing the creation, organization, dissemination, and use of information

[Sullins 2015]
0.2. Ethical theories relevant to robotics: Hybrid ethics

Since no single theory is appropriate for addressing all ethical issues arising in the design and use of robots, we take a pragmatic option:

Hybrid ethics combines:

- Top-down theories — those applying rational principles to derive norms, and
- Bottom-up theories — those inferring general guidelines from specific situations.

[Wallach and Allen 2018]
Teaching materials

Ethics in Social Robotics

0. Overview and background
1. Designing the “perfect” assistant

2. **Robot appearance and emotion**
3. Robots in the workplace
4. Robots in education
5. Human-robot interaction and human dignity
6. Social responsibility and robot morality
7. Bibliography and initiatives to follow up
2. Robot appearance and emotion

2.1. Highlights from *The Vestigial Heart*

2.2. Ethical Background and Discussion:
   - Four questions
   - Hints for a debate on each question

2.3. Revisiting Issues

2.4. Scholarly References for Further Reading
2.1. Highlights from *The Vestigial Heart*

Chapters 9/12 - Celia

As a birthday present, Lu [adoptive mother] gave me a robot. [...] it has a kind of head with no nose, mouth or ears, it just has two cameras, and a screen embedded in its chest. It’s called ROBbie.

[...] Celia, touched by the words, looks for his eyes: no friend had ever sworn their loyalty so convincingly, but two black holes bring her back down to earth. Though not entirely.

[...] she watches the robot out of the corner of her eye and it pleases her to see his dignified posture, gently swinging his strong, shiny arms. It feels good to walk along beside him, she feels protected, she can trust him. What does it matter that he doesn’t have eyes, people don’t look at each other anymore anyway.”

At the Disasters stand, Leo is puzzled by a realistic mechanical baby. [...] What woman could resist the charm of a baby that smiles when she coos at it, that she can cuddle at will while watching her favorite program, that recognizes her voice and crawls along behind her, flattering her with sweet noises? Well no sir, the product didn’t take off, almost certainly because it’s too much like the real thing, déjà vu.
Robot appearance and emotion

2.2. Questions

2.A - How does robot appearance influence public acceptance?

2.B - What are the advantages and dangers of robots simulating emotions?

2.C - Have you heard of/experienced the “uncanny valley” effect?

2.D - Should emotional attachment to robots be encouraged?
Robot appearance and emotion
Debate stemming from Chapters 9, 10 and 12 in the novel

Appearance strongly influences people’s attitude towards robots: the more anthropomorphic the robot, the more positive and empathetic the human response. [Riek et al. 2009]

However, the relation doesn’t grow unlimited; on the contrary, a point is reached where excessive similarity of the robot to a human causes distress and provokes a sudden repulsion; this is known as the “uncanny valley” effect.

Celia feels attached to her robot ROBbie because of its loyal, trustworthy and predictable behavior, which is enforced by its undeceiving machine appearance.

Leo realizes that a too-close similarity to a human being can doom a robot product.
Robot appearance and emotion

2.3. Revisiting issues

“I [Celia] have been more shocked by the kids, and even some things Lu does, than by ROBbie. For the robot, everything follows a series of rules, it’ll never surprise me with anything inappropriate.”

Transparency (1.C): Celia likes that ROBbie has a more predictable behavior than her classmates and her adoptive mother, since it has to follow rules and can’t shock her with nonsense.

“It feels good to walk along beside him [ROBbie], she feels protected, she can trust him.”

Trust (1.A): Celia feels protected by her robot, which she sees as a faithful companion that she can trust.
0. Overview and background
1. Designing the “perfect” assistant
2. Robot appearance and emotion
3. Robots in the workplace
4. Robots in education
5. Human-robot interaction and human dignity
6. Social responsibility and robot morality
7. Bibliography and initiatives to follow up
Designing the “perfect” assistant

1.2. Ethical background and discussion

Virtues attributed to the Victorian lady’s companion:

• able to distinguish its owner from other people, animals and things
• able to recognize its owner’s emotions and intentions
• behaving in a predictable and dependable way
• protective and supportive of the user in the handling of information and communication with other people
• polite but firm in the owner’s interest
• having a model of its own capabilities
• operationally reliable and requiring neither much effort from the owner to use nor special maintenance

[Peltu and Wilks 2008]
Amazing future perspectives

What role will the human and the robot play in this “pas à deux” in which we are irremissibly engaged?

Carme Torras
torras@iri.upc.edu