Delegating Responsibilities to Intelligent Robots

Gordana Dodig-Crnkovic
School of Innovation, Design and Engineering, Mälardalen University Sweden

http://www.idt.mdh.se/personal/gdc
Gordana Dodig-Crnkovic and Daniel Persson

*Sharing Moral Responsibility with Robots: A Pragmatic Approach.*

Tenth Scandinavian Conference on Artificial Intelligence SCAI 2008.
Volume 173, Frontiers in Artificial Intelligence and Applications.
Eds. A. Holst, P. Kreuger and P. Funk, 2008

---

NA-CAP@IU 2008

North American Computing and Philosophy Conference,
Indiana University, Bloomington, July 10-12, 2008

Gordana Dodig-Crnkovic and Daniel Persson

*Towards Trustworthy Intelligent Robots.*

*A Pragmatic Approach to Moral Responsibility*
Sharing Moral Responsibility With Robots: A Pragmatic Approach

Gonzalo Ondard-Cobos - gonzalo.ondard@dlk.imi.kth.se & Daniel Forsman - dfor@kth.se
School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden

1. Introduction

As artificial intelligence algorithms and systems become increasingly sophisticated, the question of who is responsible for decisions made by such systems arises. We propose an approach where responsibilities can be distributed among agents in a complex socio-technological system, where responsible (mental) agency is a matter of degree.

2. Roboethics

www.robocore.org

Roboethics is a newly established field of applied ethics that deals with the ethical implications of producing, programming, and interacting with artificially intelligent artifacts and systems such as
- Robots
- Softbots
- Direct neural interfaces
- Ubiquitous sensing systems

3. Views of responsibility

In the classical view, (moral) responsibility consists of intentions and consequences. An act that caused no evil and intended for no harm is not considered responsible for the evil.

This view of responsibility is focused on the individual agent. Applying this type of responsibility to artificial agents is controversial since there is doubt whether they are capable of internal states of intention.

Responsibility of an artifact

In the pragmatic (functional) view (moral) responsibility is seen as a social regulatory mechanism, aimed at ensuring the beneficial functioning of an agent as part of a larger socio-technological system.

4. Distribution of responsibility

The production and use of modern technological artifacts and systems require a new approach to the distribution and delegation of responsibilities.

Distribution of responsibility through social network of agents.

 Applying the pragmatic view of responsibility to such systems, we can see that an intelligent agent can be assigned responsibility for an action, but the legal and operational consequences of this responsibility are shared throughout the system.

5. Risks and responsibility

In many design processes, there are assumptions made that limited resources, products may not be used under conditions they have not been designed or tested for.

Safety-critical intelligent systems must be therefore be supported by socio-technological structures with different levels of organisational and physical barriers or up to access their safe functions under different conditions.

6. Conclusion

Intelligent systems are parts of socio-technological systems, where responsible (mental) agency is distributed, and a matter of degree based on functional relationships.

Practically, questions of responsibility to such systems may be addressed in the same way as in classical safety-critical systems.

A fundamental factor for establishing effective socio-technological systems of responsibility is education in professional ethics for engineers, and a broad, democratically based discussion on the production and use of intelligent systems.

Levels of operation, adapted from nuclear industry

1. Normal conditions
2. Abnormal situations
3. Design basis accident
4. Severe accident
Roboethics
www.roboethics.org

Information Ethics
Views of responsibility
Classical view

Causality + intention
Pragmatic view

Regulatory mechanism
Socio-technological network

Delegation & distribution
Responsible agents
“natural” & “artificial”
Safety barriers

- Physical
- Organizational
DEFENSE IN DEPTH

SEVERE ACCIDENT

DESIGN BASE ACCIDENT

ABNORMAL OPERATION

NORMAL CONDITION
Artifactual intelligence: behavior that in human would require intelligence.
Artifactual morality: behavior that in human would require morality.
RESPONSIBILITY ASCRITION

Artifacts ascribed artifactual responsibility for a task in planning operations
Proactive approach
- prediction,
- prevention,
- preparedness
Legal aspects

Analysis of technosocial system done on case by case basis.
CONVERGENCE

HUMAN AGENCY

ARTIFACTUAL AGENCY
Wide range, long-standing consequences of the deployment of intelligent systems in human societies must be discussed on a democratic basis as the intelligent systems have a potential of radically transforming the future of humanity.
Education in Ethics for Engineers

Education in professional ethics for engineers is a fundamental factor for building a socio-technological system of responsibility.
Swedish Computer Science and Engineering education follows in many respects the ACM/IEEE Computing Curriculum

http://www.computer.org/education/cc2001/index.htm
Computing Curricula 2001, ACM/IEEE

- Social context of computing
- Methods and tools of analysis of ethical argument
- Professional and ethical responsibilities
- Risks and liabilities of safety-critical systems
- Intellectual property
- Privacy and civil liberties
- Social implications of the Internet
- Computer crime
- Philosophical foundations of ethics
“All products of technology present some potential dangers, and thus engineering is an inherently risky activity. In order to underscore this fact and help in exploring its ethical implications, we suggest that engineering should be viewed as an experimental process. It is not, of course, an experiment conducted solely in a laboratory under controlled conditions. Rather, it is an experiment on a social scale involving human subjects.”

A Framework for Ethical Decision Making

Responsible ethical decision making implies ability to:

– recognize a moral issue
– get the facts
– evaluate the alternative actions from various moral perspectives
– make a decision
– act
– reflect on the results of the decision afterwards.
The Goal of Ethics for Engineers

The most important goal is to develop the ethical autonomy, i.e. the skill and the habit of thinking rationally and critically about the ethical issues.
Why is Professional Ethics Important for Computer Scientists and Engineers?

Because the Professional Ethics shall be a part of education for every socially important profession, as an important component of the concept of professionalism!
## Professional Ethics in Science and Engineering Course at MDH

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>METHODS AND TOOLS OF ANALYSIS OF ETHICAL ARGUMENT. Philosophical Foundations of Ethics. Ethical Relativism, Absolutism and Pluralism</td>
</tr>
</tbody>
</table>

**Teacher and examiner:** [Gordana Dodig-Crnkovic](mailto:gordana.dodig-crnkovic@mdh.se)
| L3   | The Ethics of Conscience  
The Ethics of Duty  
The Ethics of Respect |
|------|---------------------------|-----------------------------|-----------------------------|
| L4   | The Ethics of Consequences: Utilitarianism  
The Ethics of Rights  
The Ethics of Justice |
| L5   | The Ethics of Character  
Ethics and Gender |
| L6   | GUEST LECTURE AI AND ETHICS |
| L7  | PRIVACY AND CIVIL LIBERTIES  
|     | In-class activity: CASE STUDIES |
| L8  | ENVIRONMENTAL ETHICS  
|     | In-class activity: CASE STUDIES |
| L9  | GUEST LECTURE ETHICS OF CARE  |
| E1  | PROFESSIONAL AND ETHICAL RESPONSIBILITIES  
|     | CODES OF ETHICS  
|     | WHISTLE BLOWING  
|     | In-class activity: CASE STUDIES  
|     | Social context of profession of engineer and scientist/researcher |
| L10 | RISKS IN TECHNOLOGY AND SCIENCE  
Risks and liabilities of safety-critical systems  
PRECAUTIONARY PRINCIPLE |
| L11 | GUEST LECTURE BY MONIKA EIBORN |
| E2  | INTELLECTUAL PROPERTY  
In-class activity: CASE STUDIES |
| L12 | In-class activity: ORAL PRESENTATIONS |
| E3  | COURSE WRAP-UP  
TAKE-HOME EXAM |

http://www.idt.mdh.se/kurser/cd5590/ Professional Ethics Course, MDH