

# An Analytic Approach to the Philosophy of Robots: Dennett's Stances and Design for Values

July 7, 2018

Pieter E. Vermaas  
Delft University of Technology  
p.e.vermaas@tudelft.nl

# Aim

1. To present an analysis of robots using tools from analytic philosophy of technology
2. To introduce design for values
3. To give some challenges for designing robots for values

1. To present an analysis of robots using tools from analytic philosophy of technology

# Analysis of robots

The fundamental question:

- What is a robot?

Two fundamental answers:

1. a technical artefact designed for carrying out complex actions automatically
2. an entity with some form of consciousness as reasoning on the basis of observations, goals and intentions

# Analysis of robots

The fundamental question:

- What is a robot?

Two fundamental answers:

1. a technical artefact designed for carrying out complex actions automatically

Philosophy of technology

2. an entity with some form of consciousness as reasoning on the basis of observations, goals and intentions

Philosophy of mind

# Quick Philosophy of Mind: Dennett's three stances

## The physical stance:

predict behavior with physics, chemistry, ..



## The design stance

predict behavior with functions and goals



## The intentional stance

predict behavior by assuming that

- an entity as conscious and rational, and
- it has certain observations, thoughts and aims



# Epistemology and not ontology

Ontological issues whether entities are actually designed or really conscious are side-stepped

We apply a specific stance to a particular entity if that enables us to predict its behavior **effectively and efficiently**

- For applying the intentional stance to Heidegger, we do first determine whether Heidegger is rational, has observations, thoughts and aims.
- If application of this stance leads efficiently to good predictions of Heidegger, then we describe Heidegger as an intentional entity



# An analytic answer to the consciousness of robots

The ontological issues whether robots are actually conscious can be side-stepped

We describe robots as conscious if that description efficiently predicts their behavior



# Towards an analytic answer to the design of robots

Robots are technical artefacts that are designed by engineers and used by users

For bringing these elements into the analysis, the intentional stance should be applied to them too

I focus on the designing engineer



# The intentional intentional stance

1. take two entity as conscious and rational
2. assume that they both have certain observations, thoughts and aims
3. presuppose that the one (the robot) is designed by the other (the engineer)



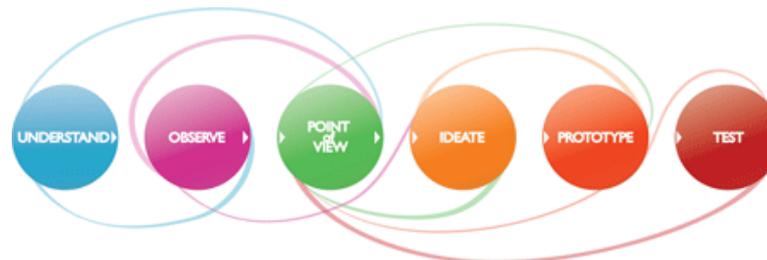
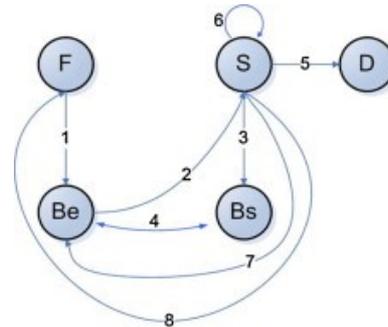
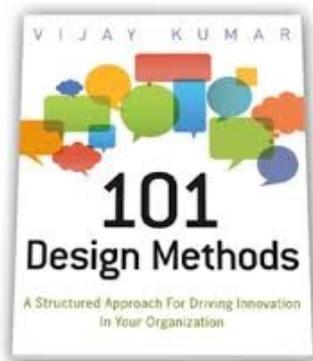
2. To introduce design for values

# Less quick philosophy of technology: design

Study the way engineers design for

- Understanding what it means that robots are designed
- Giving recommendations how they should be designed

But there are some problems



**Frame Creation  
steps:**

**Archaeology**

**Paradox**

**Context**

**Field**

**Themes**

**Frames**

**Futures**

**Transformation**

**Integration**

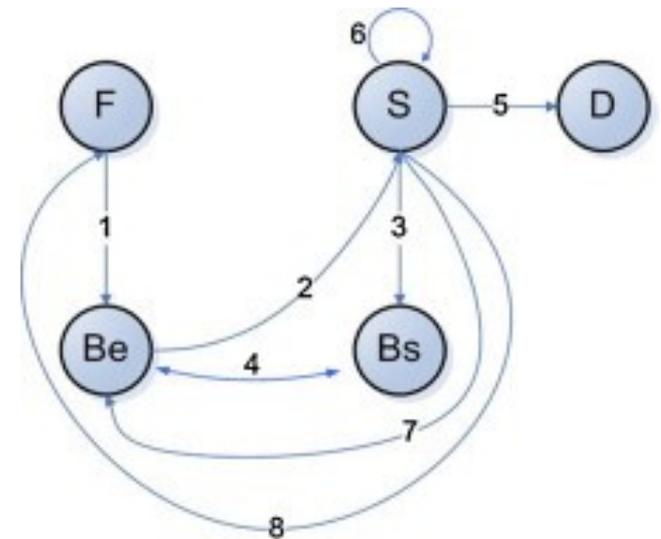
# Engineering design/critical philosophy of technology

## Design

- Translating functional requirements into descriptions of technical artefacts that meet those requirements

## Philosophy of technology

- Arguing that engineers fail



# Engineering design/critical philosophy of technology

## Design

- Translating functional requirements into descriptions of technical artefacts that meet those requirements

## Philosophy of technology

- User dynamics
- Technological determinism
- Technological fixes



# Engineering design/critical philosophy of technology

## Design

- Translating functional requirements into descriptions of technical artefacts that meet those requirements

## Philosophy of technology

- Disaster



# Constructive turn: adapting technology

In engineering some values are actively incorporated:

- Design for safety
- Design for sustainability
- Inclusive design
- Participatory design
- Social design



Batya Friedman's [Value-sensitive design](#):

- Adjust technology by incorporating our moral and social values in the design process

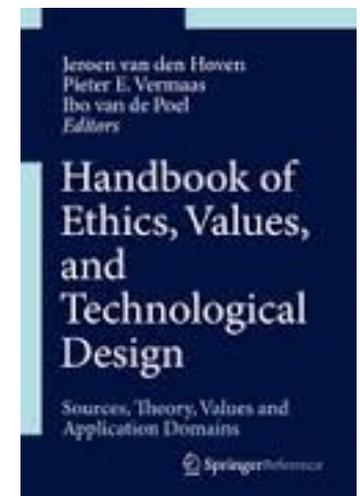


Jeroen van den Hoven: Developing the means to actively incorporate values in engineering design

1. for mitigating societal rejection of technologies
2. for shaping technology with our values

Targets:

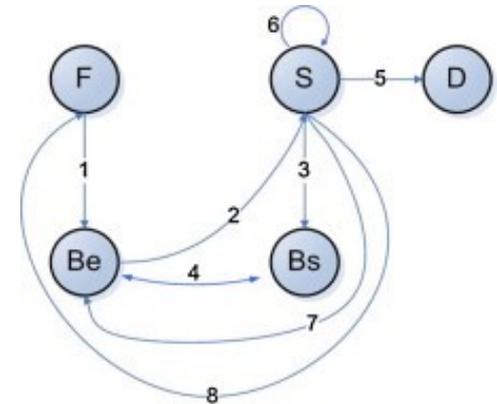
- collecting practices of design for values
- creating methods for design for values
- ethics labs
- enabling auditing



# From engineering design to innovation design

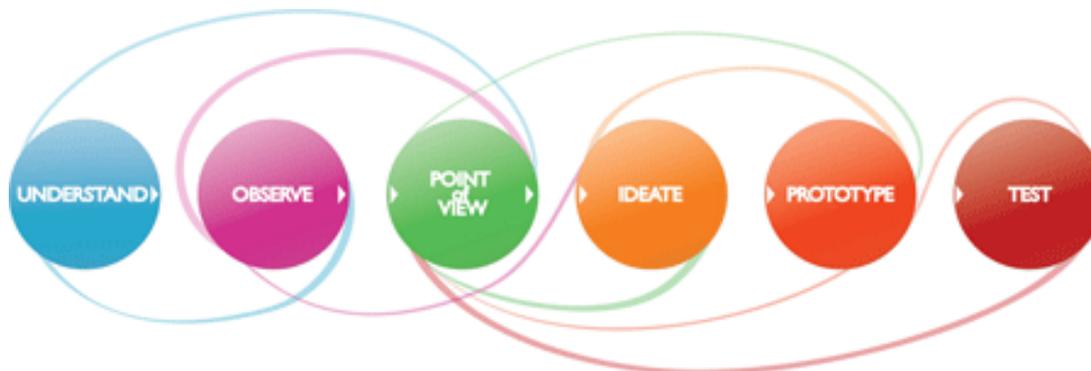
## Engineering design

- Translating functional requirements into descriptions of technical artefacts that meet those requirements



## Innovation design

- Analysing tasks for finding solution directions



**Frame Creation  
steps:**

**Archaeology**

**Paradox**

**Context**

**Field**

**Themes**

**Frames**

**Futures**

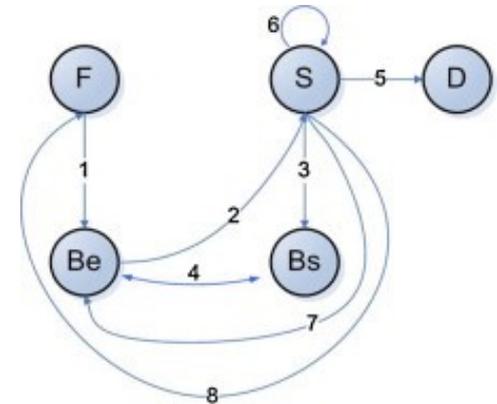
**Transformation**

**Integration**

# Innovation design/helpful philosophy of technology

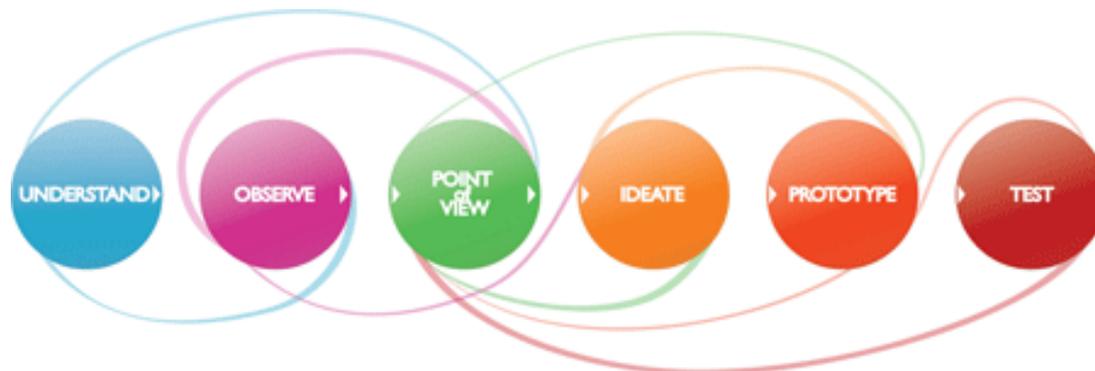
## Engineering design

- Translating functional **and value** requirements into technical artefacts that meet those requirements



## Innovation design

- Analysing tasks such as incorporating values for finding solution directions



**Frame Creation  
steps:**

**Archaeology**

**Paradox**

**Context**

**Field**

**Themes**

**Frames**

**Futures**

**Transformation**

**Integration**

3. To give some challenges for designing robots for values

# Designing robots for values

Two general challenges:

1. which values?
2. how to design for them?



# Which values?

Examples of values mentioned at our workshop:

1. Human autonomy: robots should not deskill human
2. No harm by androids to cognitive development of infants
3. No harm by robots for humanitarian aid

Analyses at our workshop called for analyzing the impact that the existence of robots can have on our world, and a normative response to it, focusing on, e.g.

1. Human vulnerability
2. Love and sex

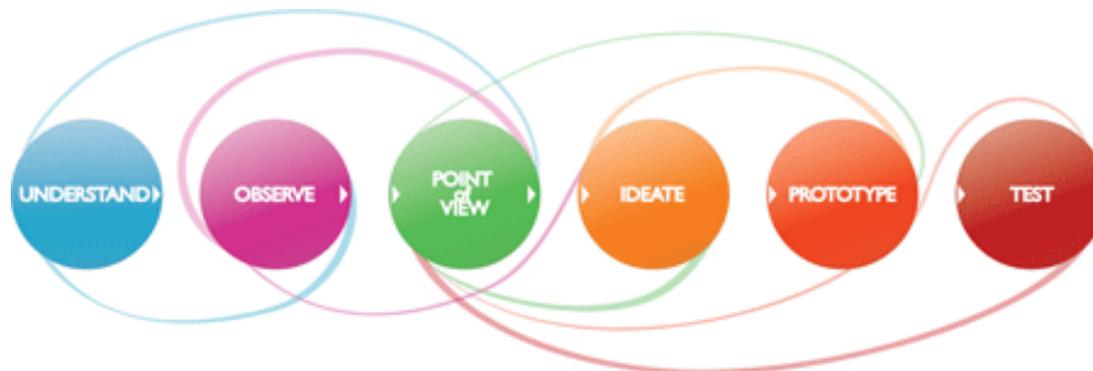
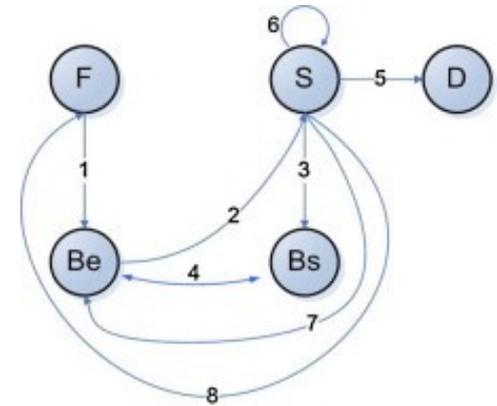
# How to design for these values

## Engineering design

- Translating the values into design requirements

## Innovation design

- Analysing the values for finding solution directions



**Frame Creation  
steps:**

**Archaeology**

**Paradox**

**Context**

**Field**

**Themes**

**Frames**

**Futures**

**Transformation**

**Integration**

# Conclusions

1. I presented an analysis of robots with Dennett's intentional stance and design methods
2. I introduced design for values as an effort in philosophy of technology to develop means to incorporate moral and societal values into technologies such as robots
3. The challenge becomes finding the values we hold for robots and finding ways to incorporate them in robots

