

REAL AI IS SOCIAL AI

FRANK DIGNUM



UMEÅ UNIVERSITY



Universiteit Utrecht

Agenda

- What is **AI**?
- Being **social**
- Modeling **social concepts**
- Consequences **for AI research**

Artificial Intelligence

Intelligence:

the ability to perceive or infer information, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context

Artificial Intelligence:

the **research area** that studies human intelligence through the design of computer systems that emulate (parts of) human intelligence.

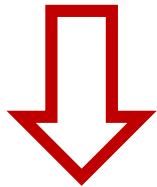
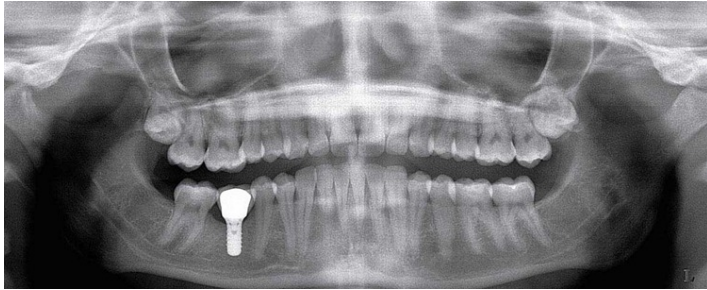
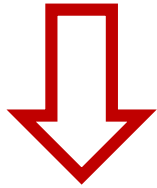
Artificial Intelligence (warning for fake AI)

Techniques used in sub areas are part of AI
research

BUT NOT

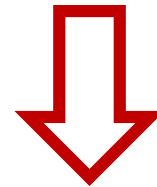
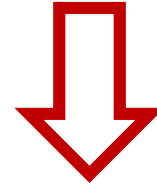
All systems that use these techniques are AI
systems!

Does this person have cavities in her teeth?



Yes/No

Where and how can we best receive refugees?



?

Main message

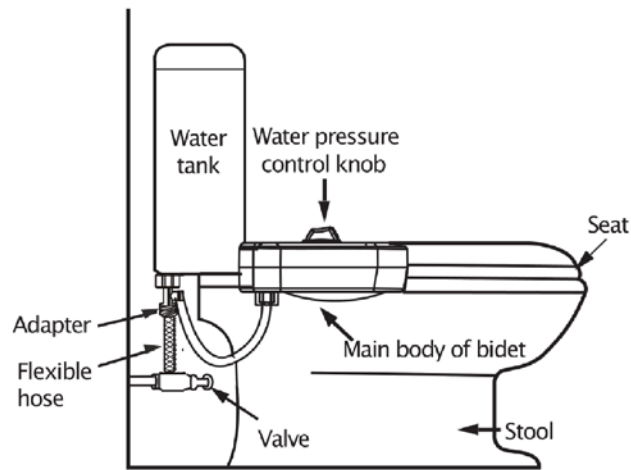
Intelligence is social



**AI should become
social**

Intelligence **Is social**

Who knows how a toilet works?



We only need to know how to interact with it
And to know someone who can repair it when it is broken

**Umntu
Ngumuntu
Ngabantu**

**A PERSON IS A PERSON
THROUGH OTHER PERSONS**

AI: from tool to partner



AI in a SOCIAL context

- Optimal decision → Accepted decision
- Black box decision → Explainable decision
- One shot decision → Repeated decisions
- Action → Interaction
- Direct effect → Long term social effect
- ...

predict and adapt intelligently to social behaviour



incorporate Sociality in AI systems

Sociality according to Social Science

1. **Individualists** (Weber, beginning 1900's):
individuals only, social arises from individual behaviors
2. **Socialists** (Durkheim, 1910's):
"natural" norms/roles determine individual behavior
3. **Textualists** (Habermas, 1960's):
conceptualization and language determine our social reality
4. **Social Practice** (Latour/Reckwitz, 1990's): social reality is shaped by practices, the process is central
5. **Social Persons** (John Mbiti, 1975):
I am because we are, and since we are, therefore I am;
Persons are shaped through their interactions with other persons

Sociality according to Agents Community

Individualists: AAMAS, Game Theory

Socialists: COIN, Social Simulation

Culturalists: Agent Communication

Social Practice: Social Simulation

Social Persons:

Socio-Cognitive Systems?

Social structures and rules

- Formal social structures:
 - Institutions, Organizations, Nations,...
- Informal social structures:
 - Teams, Groups, Families, Friends,...

Social rules are described in terms of:

- Roles
- Social Practices
- Conventions
- Norms
- Values
- Culture
- ...

Model and implement social reality

Social
Reality



=



What is silly putty?



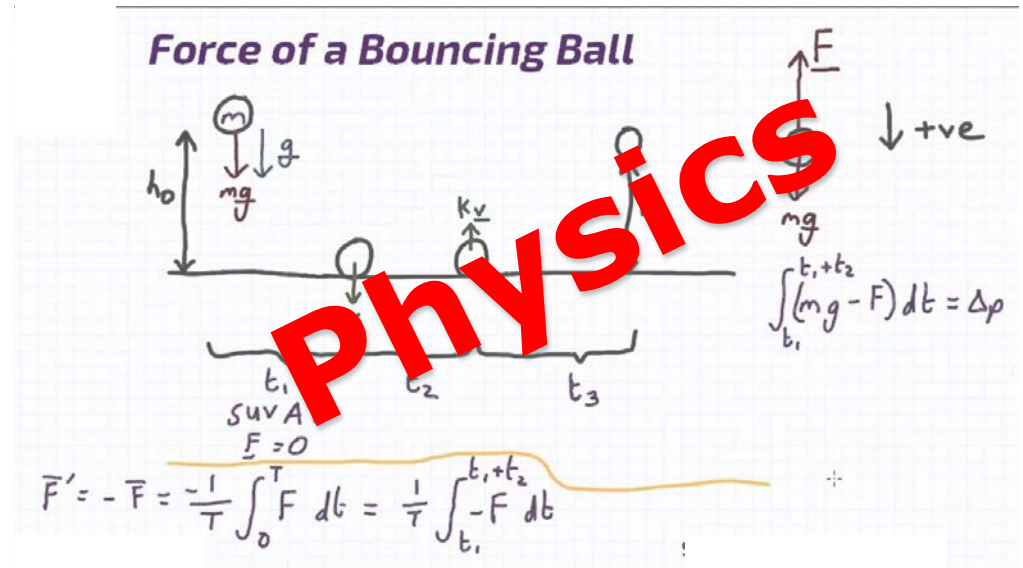
=

65% [dimethylsiloxane](#)
([hydroxy](#)-terminated polymers with [boric acid](#)),
17% [silica](#) (crystalline quartz),
9% Thixatrol ST (a [polyol](#) derivative),
4% [polydimethylsiloxane](#),
1% [dodecylmethyl cyclopentasiloxane](#),
1% [glycerine](#),
1% [titanium dioxide](#).

What is silly putty?



=



What is silly putty?



=



What is silly putty?



=

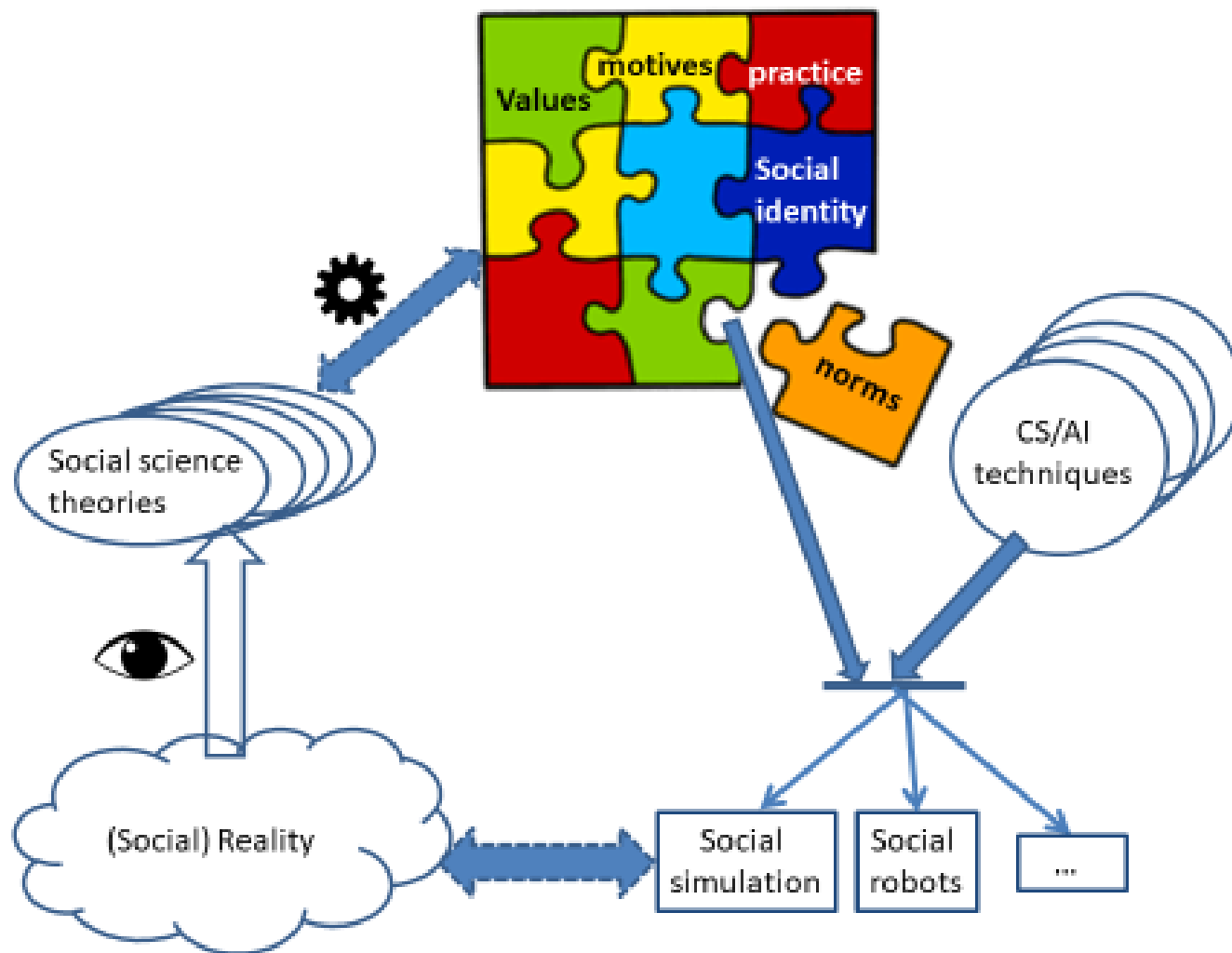


The danger of computer science

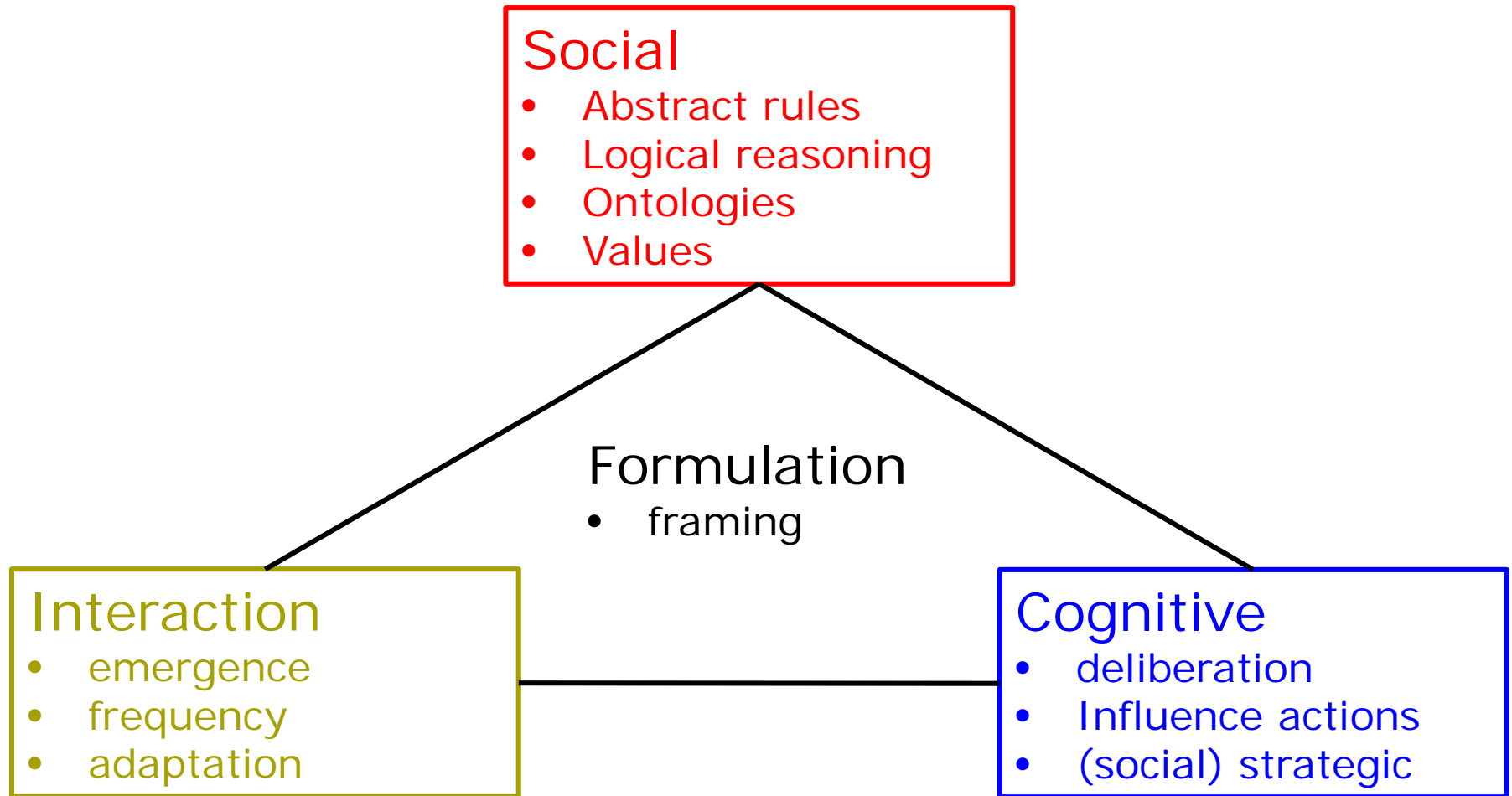
- Start with a technique
- Explain a phenomenon in terms of that technique
- **It fits,**
- because everything is simplified until it fits!

- Logic → Axioms, consistency,...
- Game theory → Utility, strategy,...
- Petri-Nets → Liveness, deadlock,...
- Bayesian Networks → Priors, influence, probability,...
- Neural Networks → classification,...
- Social simulations → emergence,...
- Complex systems → networks, feedback loops,...
- Linear programming → optimal solution,...
- ...

Attempt with socio-cognitive systems



Social concepts have three perspectives



Example: norms

Social

- Abstract: It is forbidden to discriminate
- Paradoxes: $F(\text{paint}(\text{fence}, \text{white}))$ but if done $O(\text{paint}(\text{gate}, \text{white}))$
- $\text{Person}(\text{male}) > \text{Person}(\text{female})$ counts_as discriminate

Formulation

- Forbidden to run in corridor

Interaction

- Social norms emerge from interactions
- When is a pattern a norm?
- Adaptation?

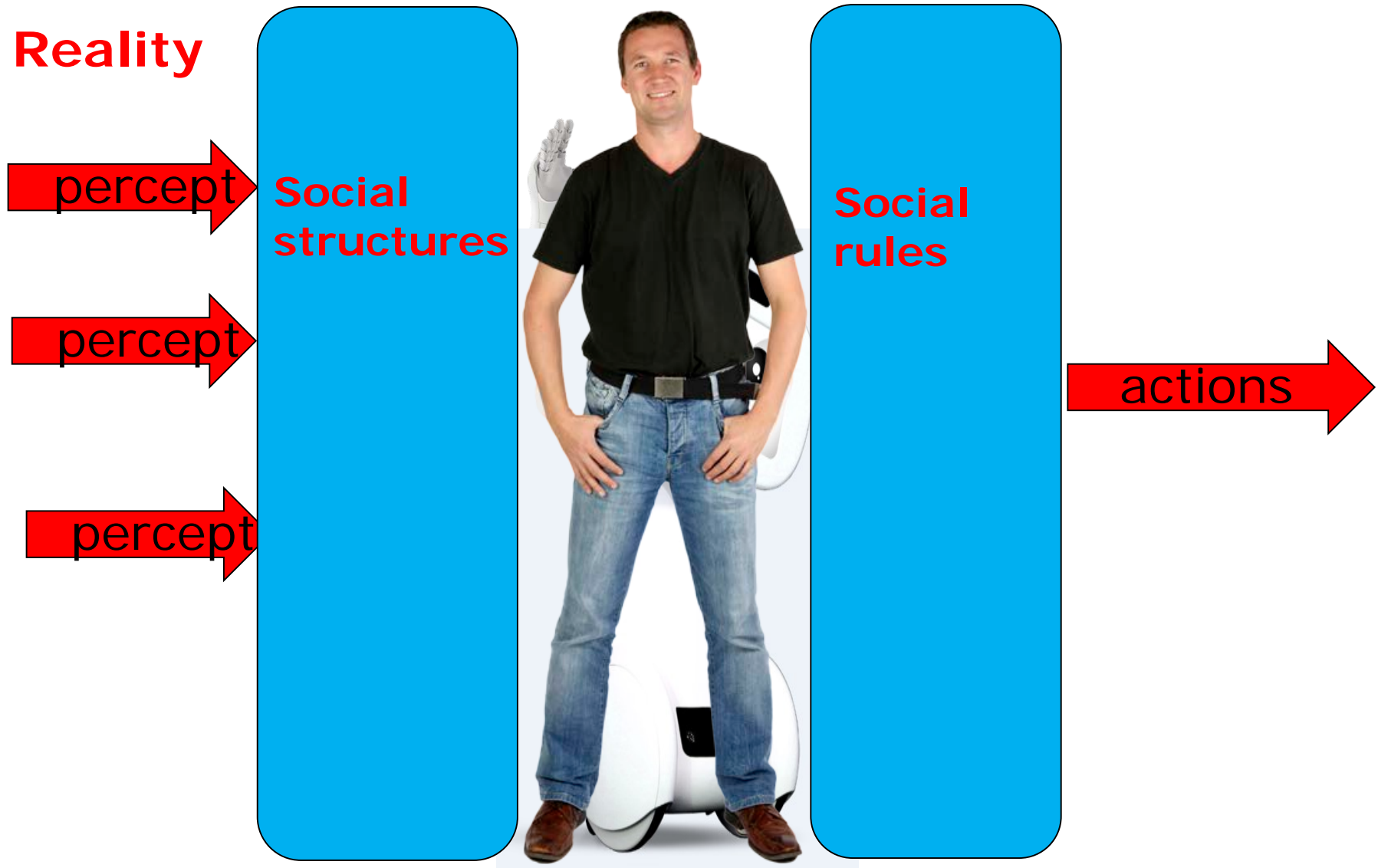
Cognitive

- When to violate a norm?
- Plan with norms
- Expect behavior based on norms

Open issues:

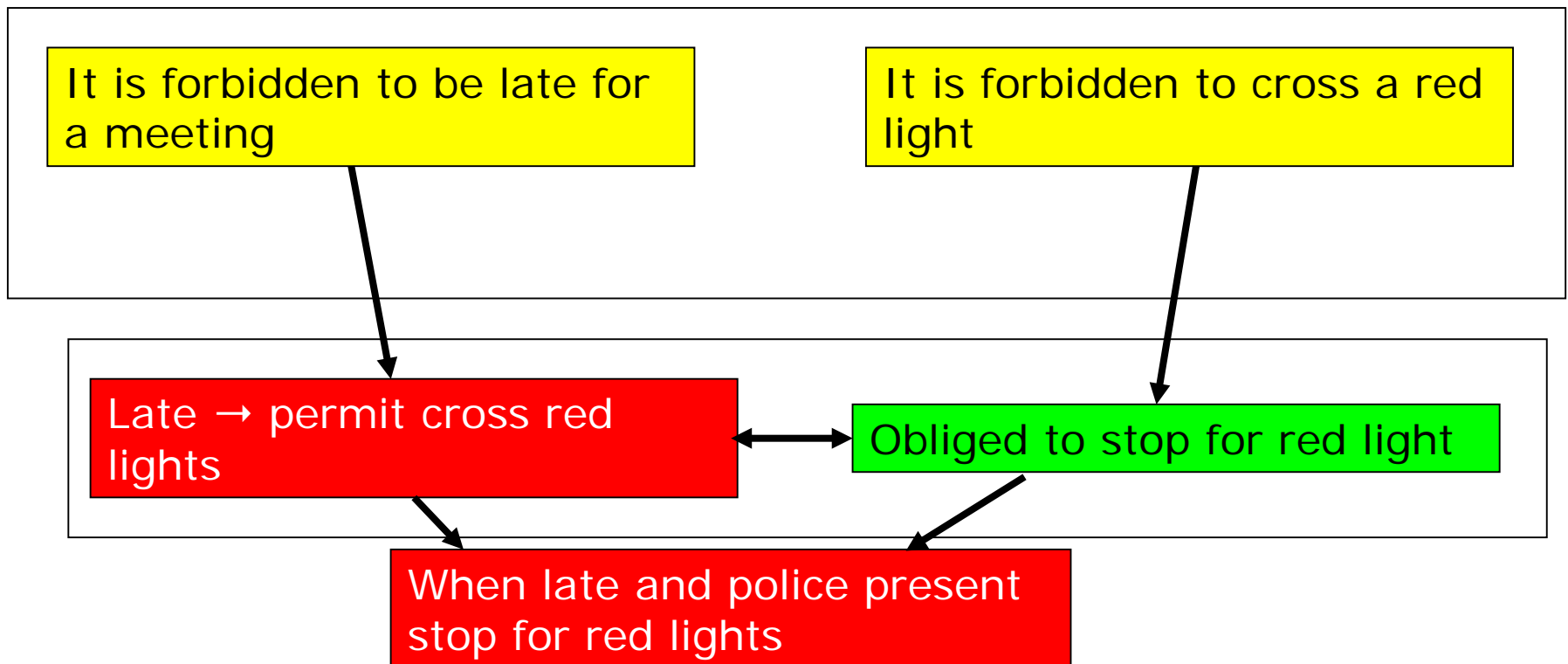
- 1. Modularity and compositionality of social and cognitive models**
- 2. Dynamicity of social reality**

Add social modules to the AI system?



Are the norms specified consistent and/or complete?

- Assumption: Check consistency in norms module
- Problem: connection with planning



Consequence for AI

Social and cognitive aspects of AI systems have to be developed in synchronization.

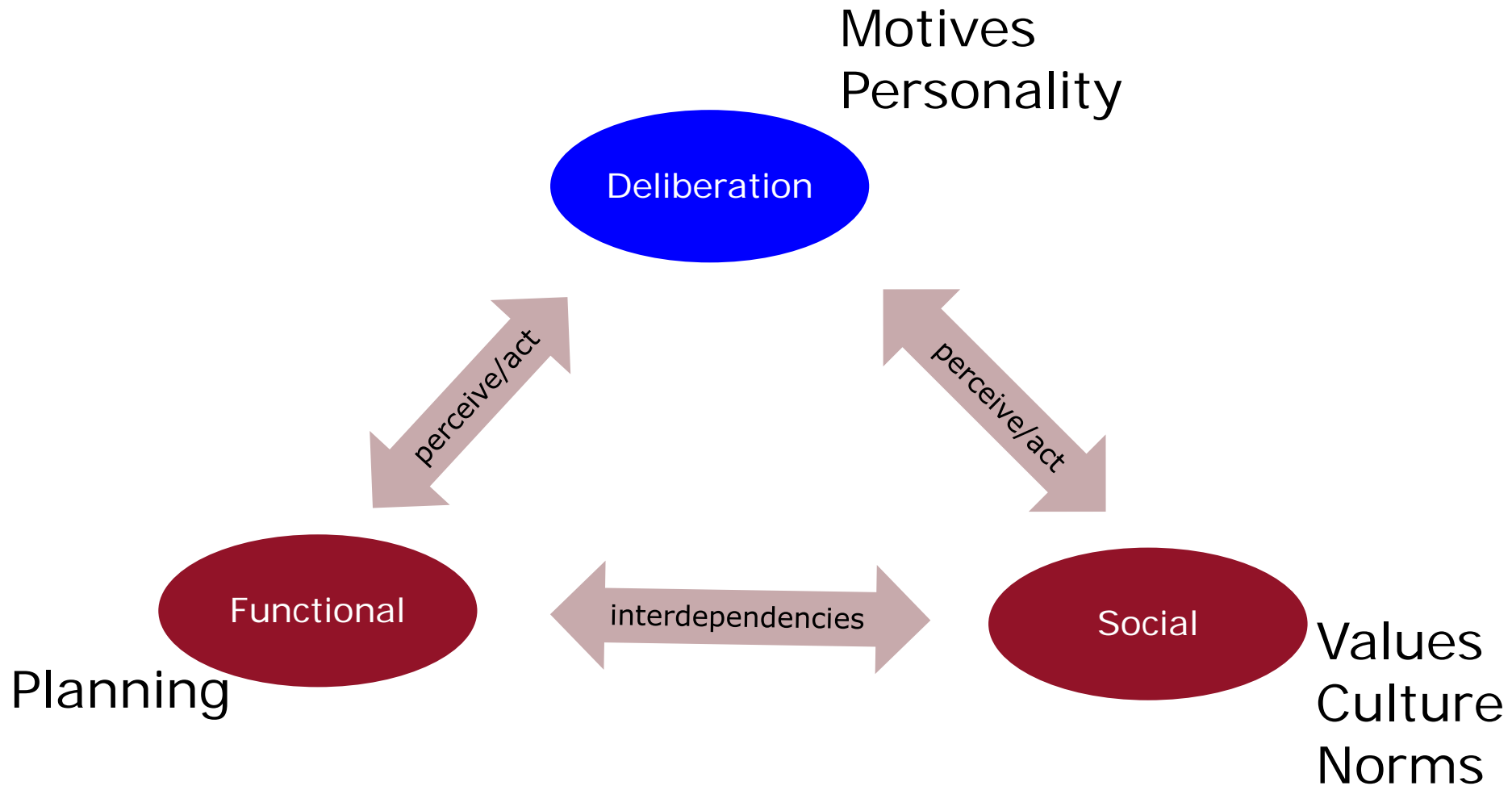
Dynamics:

Social structures motivate, emerge, adapt,...

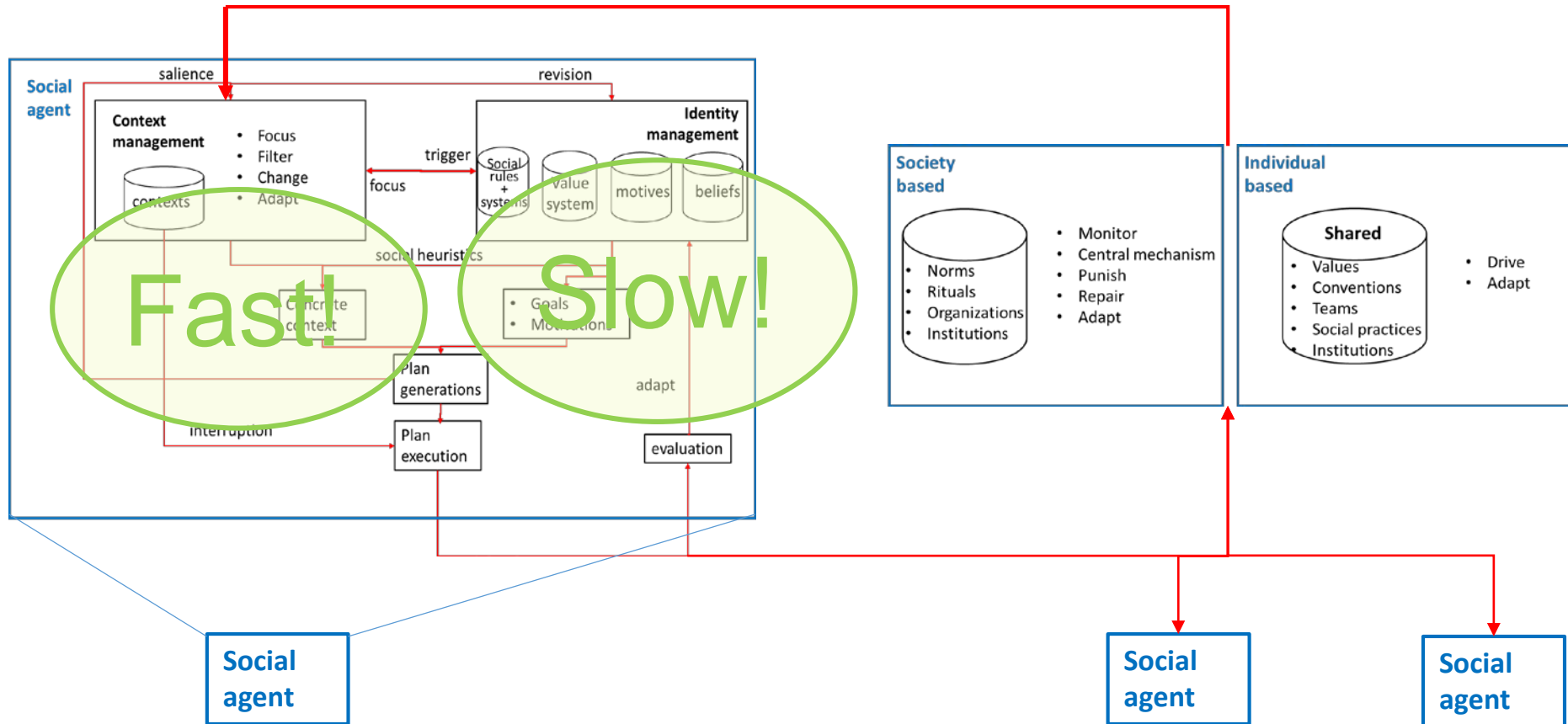
- Persons influence each other **through** social structures, **using** social structures and **because of** social structures



Social deliberation



Sketch of a social AI system



How to use theory on social intelligent behaviour?

1. During design of interactive systems
2. For designing socially intelligent systems
3. For designing social simulations
4. For designing MAS supported socio-technical systems

Designing applications in a social context

Social practices
Norms
Motives



Cardiomyopathy

Social practices
Norms and values
Social status



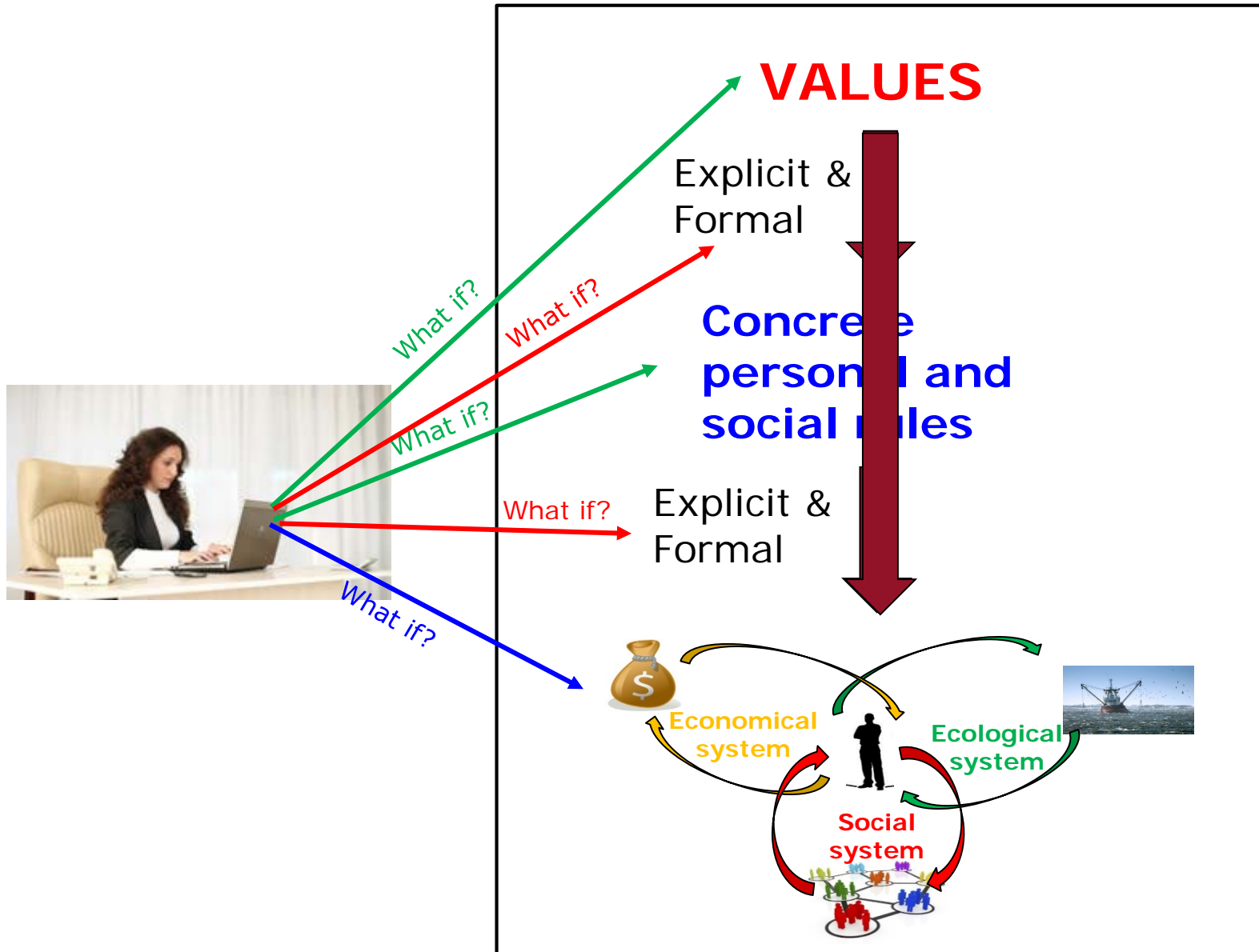
Child power

Design socially intelligent systems

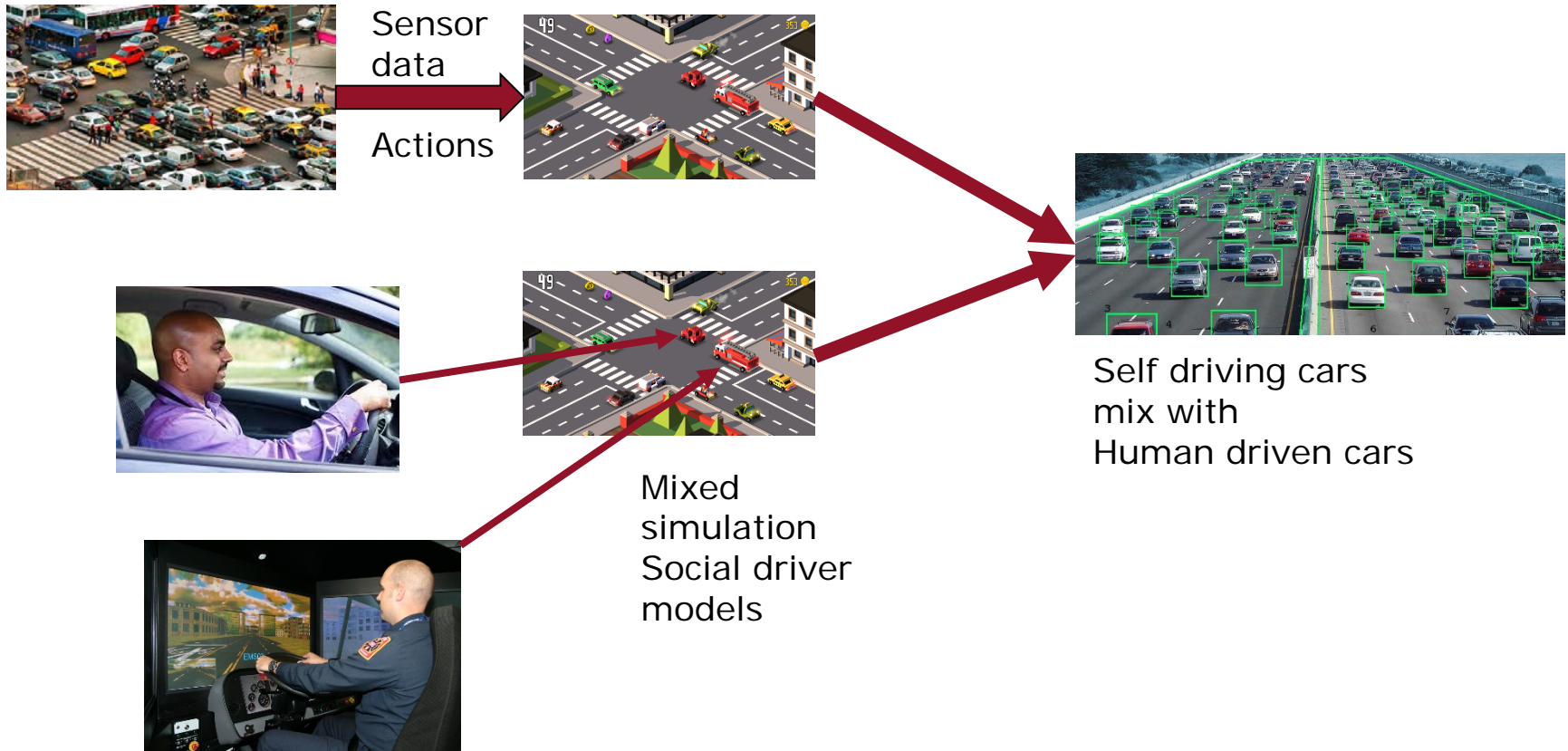
Social practices
Norms
Roles



Design of social simulations



Designing MAS supported socio-technical systems



Conclusions

1. **Start** of exciting new development
2. Possibly huge impact
3. Possibly too early and big failure
4. But I have made promises before:

PhD defence
18 December 1989

PROMISES

WASP lecture
15 January 2020

1. I promise that I will work in service for the welfare of humanity.
Therefore I will not work on projects for pure own profit or only monetary gain.
In particular I will not join projects that, according to my opinion, serve the arms race directly.
2. I promise not to start projects of which I, based on my knowledge in that area, can predict that they will not lead to the result as envisioned by the client or funding agency.
3. I will not use my title to pretend to have knowledge or skills I do not have. Thus, if I am asked *professionally* for a service for which I do not have the skills I will state this clearly.
4. At all times, I will assume full responsibility for the work that I have done.
5. I acknowledge the right of all persons for self determination; this means that I will not cooperate in any way on exploitation or suppression of fellow humans. This holds in particular in so far this is the consequence of the use (autonomous) software systems. I will therefore critically follow the developments and the boundaries of computer science and AI and inform the general public about these developments.

Frank Dignum



UMEÅ UNIVERSITY