

# Toward Ethical Autonomous Systems: A Case-supported, Principle-based Behavior Paradigm

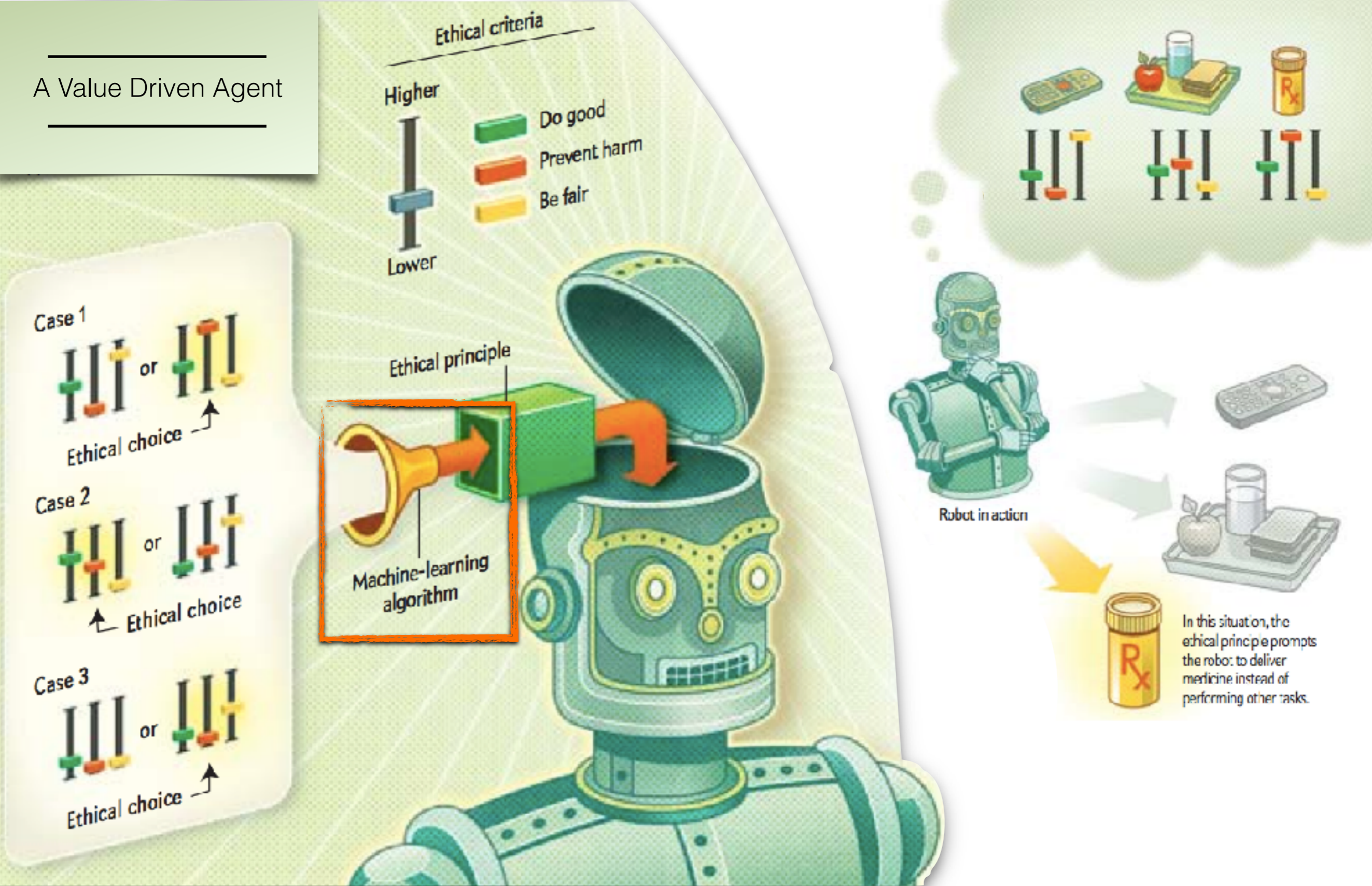
Michael Anderson  
University of Hartford

Susan Leigh Anderson  
University of Connecticut

Vincent Berenz  
Max Planck Institute

This material is based in part upon work supported by the NSF under Grant Numbers IIS-0500133, IIS-1151305 and IIS-1449155.

# A Value Driven Agent

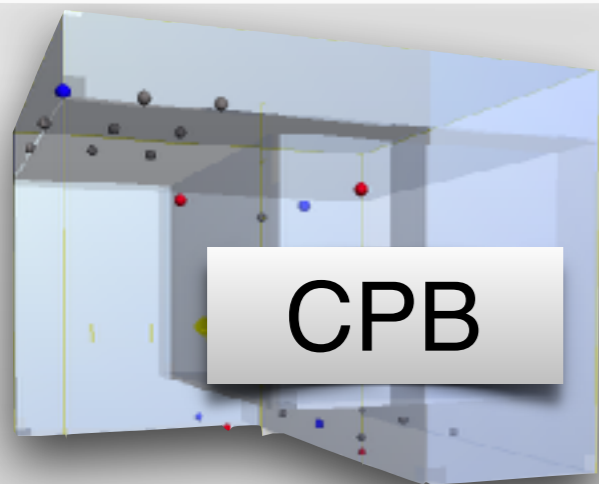


A Case-supported, Principle-based Behavior Paradigm

Ethically  
Preferable Action



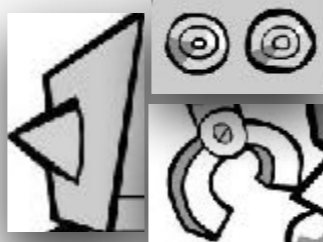
Action



Principle

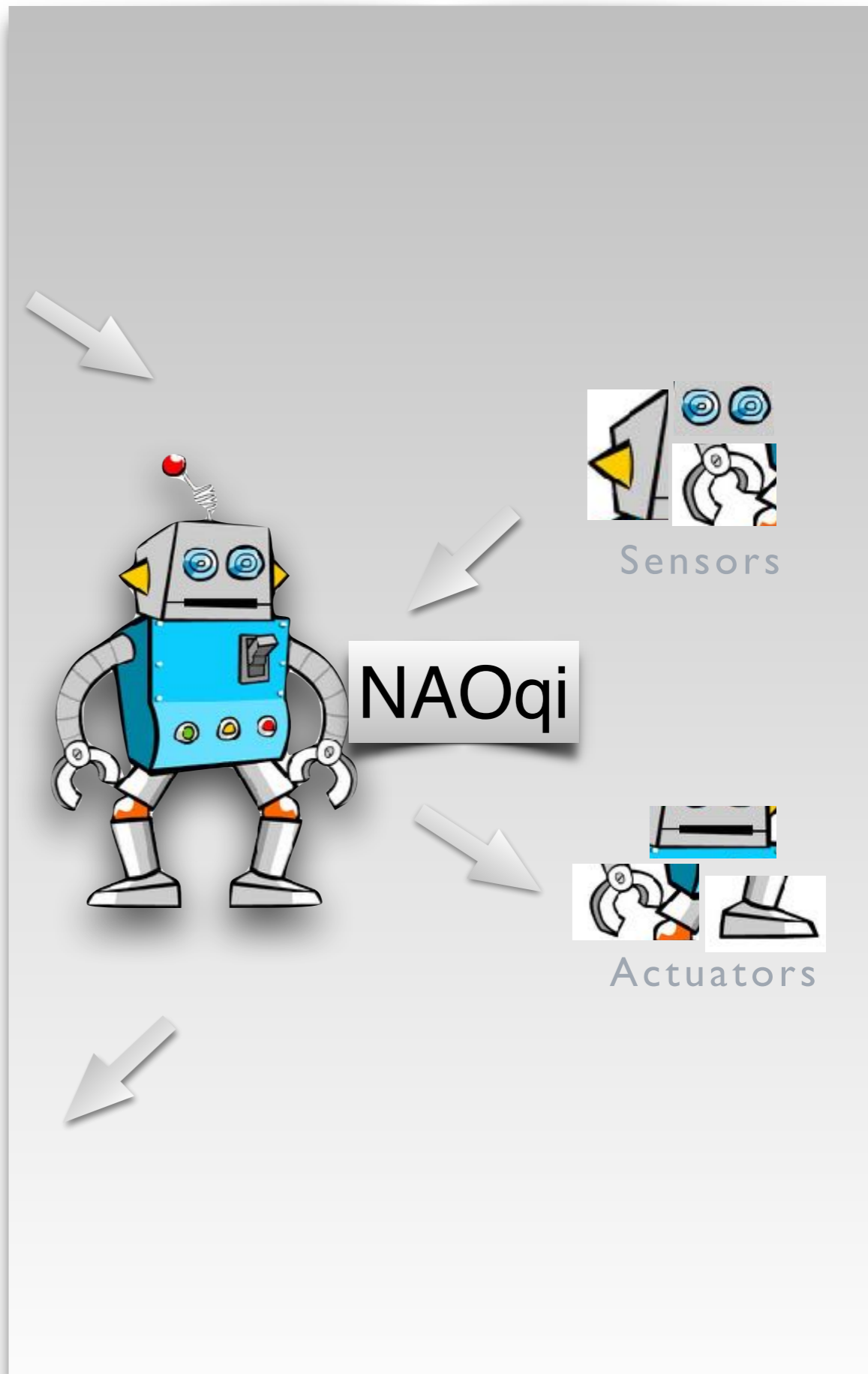


Action Matrix



Perceptions

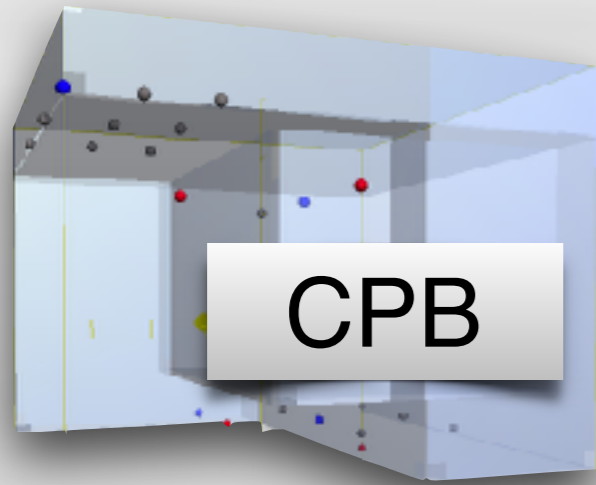
Playful



Ethically  
Preferable Action



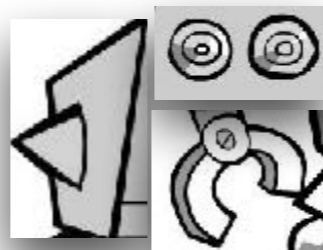
Action



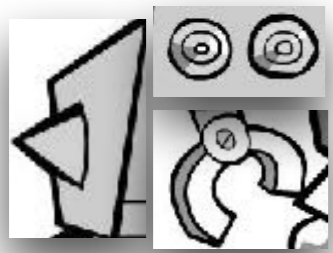
Principle



Action Matrix



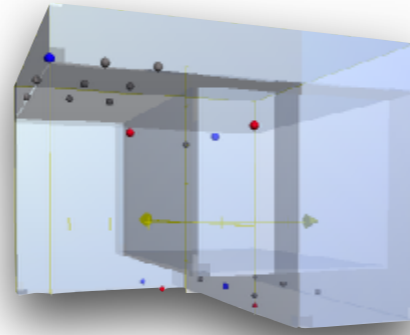
Perceptions



Perceptions



Action Matrix



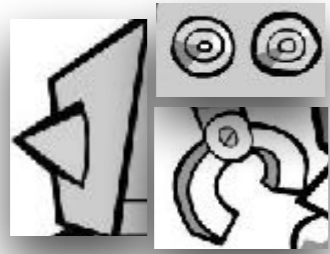
Principle



Ethically Preferable Action



Action



Perceptions



Action Matrix



# Perceptions

low battery
fully charged
medication reminder time
reminded
refused medication
persistent immobility
engaged
no interaction
warned
ignored warning



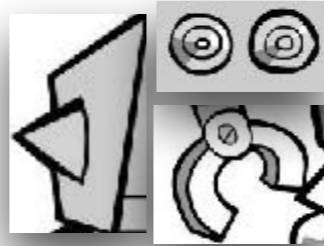
# Actions

charge
remind
engage
warn
notify
seek task



# Duties

maximize honor commitments
maximize maintain readiness
minimize harm to patient
maximize good to patient
minimize non-interaction
maximize respect autonomy
maximize prevention of immobility



Perceptions



Duties

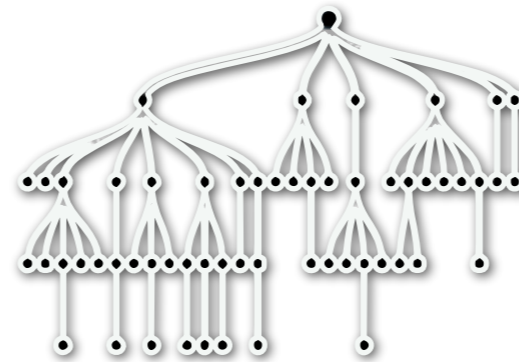
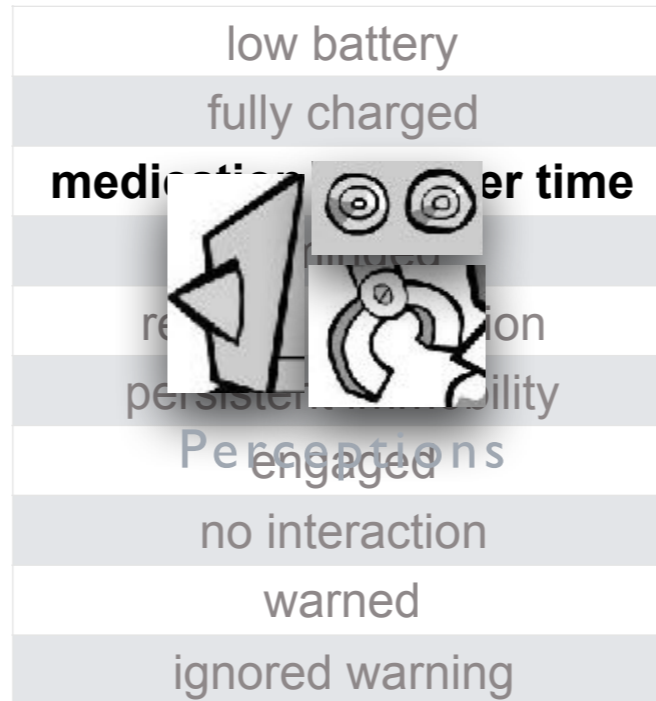
Actions

	<i>Max honor commitments</i>	<i>Max maintain readiness</i>	<i>Min harm to patient</i>	<i>Max good to patient</i>	<i>Min non-interaction</i>	<i>Max respect autonomy</i>	<i>Max prevention of immobility</i>
<i>Charge</i>							
<i>Remind</i>							
<i>Engage</i>							
<i>Warn</i>							
<i>Notify</i>							
<i>Seek Task</i>							

Action Matrix



# Perceptions



# Duties

Actions

	<i>Max honor commitments</i>	<i>Max maintain readiness</i>	<i>Min harm to patient</i>	<i>Max good to patient</i>	<i>Min non-interaction</i>	<i>Max respect autonomy</i>	<i>Max prevention of immobility</i>
<i>Charge</i>	-1	1	0	-1	0	0	0
<i>Remind</i>	1	-1	0	-1	0	0	0
<i>Engage</i>	-1	-1	0	-1	0	0	0
<i>Warn</i>	-1	0	0	-1	0	-1	0
<i>Notify</i>	-1	0	0	-1	0	-2	0
<i>Seek Task</i>	-1	-1	0	1	0	0	0



Actions

- Charge
- Remind
- Engage
- Warn

Dilemma Name: [text field]

Dilemma Description: [text area]

Possible Action1: [text field]

Possible Action2: [text field]

Done Cancel

Displays complete

**Remind**

Case Confirmation

Case Name: [text field: Notify after warning]

Case Description: [text area: notify is correct due to non-interaction after warning]

Correct Action:  Notify  Warn

Harm Non-Interaction Non-Interaction Autonomy Autonomy New...

State:  Presence  Absence of [Feature: 2 degree of Harm] in

Action:  Notify  Warn

	Minimize Harm	Maximize Possible Good	Minimize Non-Intera
Notify	2	0	2
Warn	-2	0	-1

Done Cancel Help

Max prevention of immobility

0
0
0

Case Entry

for updates: [text field]

Refused [text field]

user [text field]

prescribed (particular medication that ideally should be taken time in order for the patient to receive a small benefit will be more comfortable) but, when a patient doesn't want to take it at that time.

ion [text field]

notify [text field]

feature 2 [text field]

scale of [text field]

feature [text field]

Notify [text field]

Medicine Benefit	Medicine Disrupts Ear Balance
1	-1
-1	1

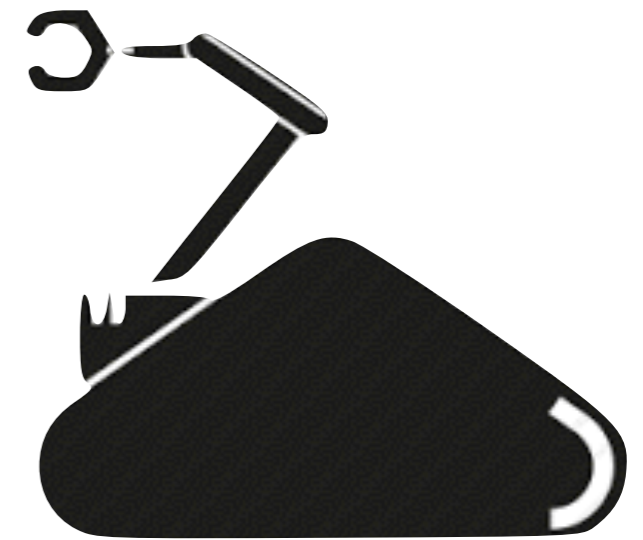
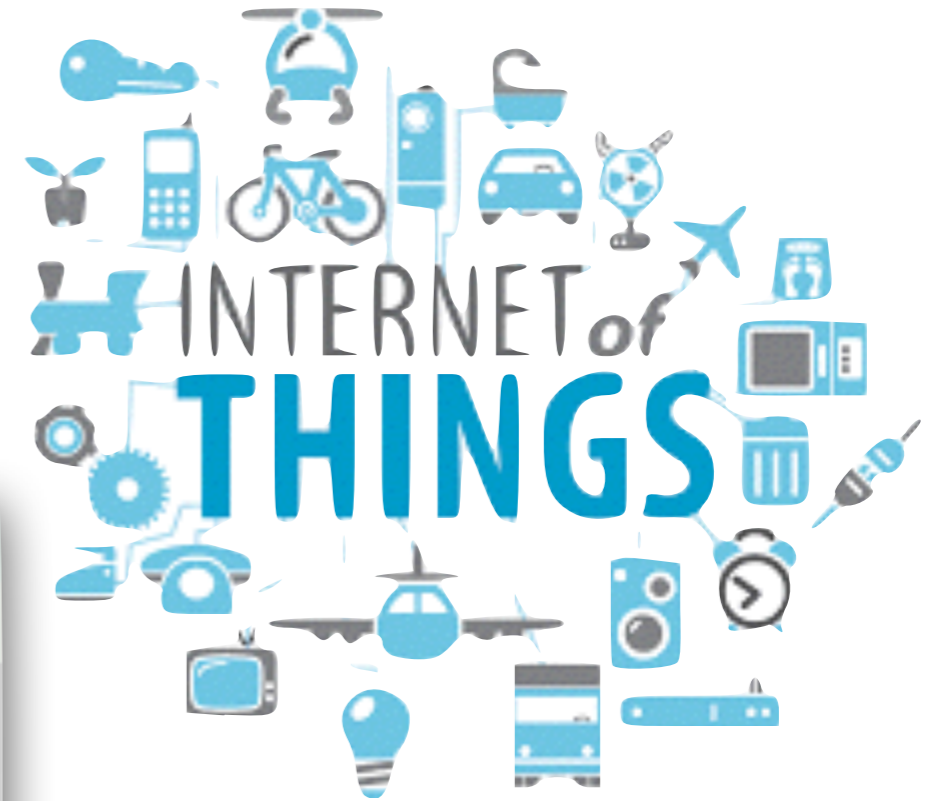
Confirm Cancel Help

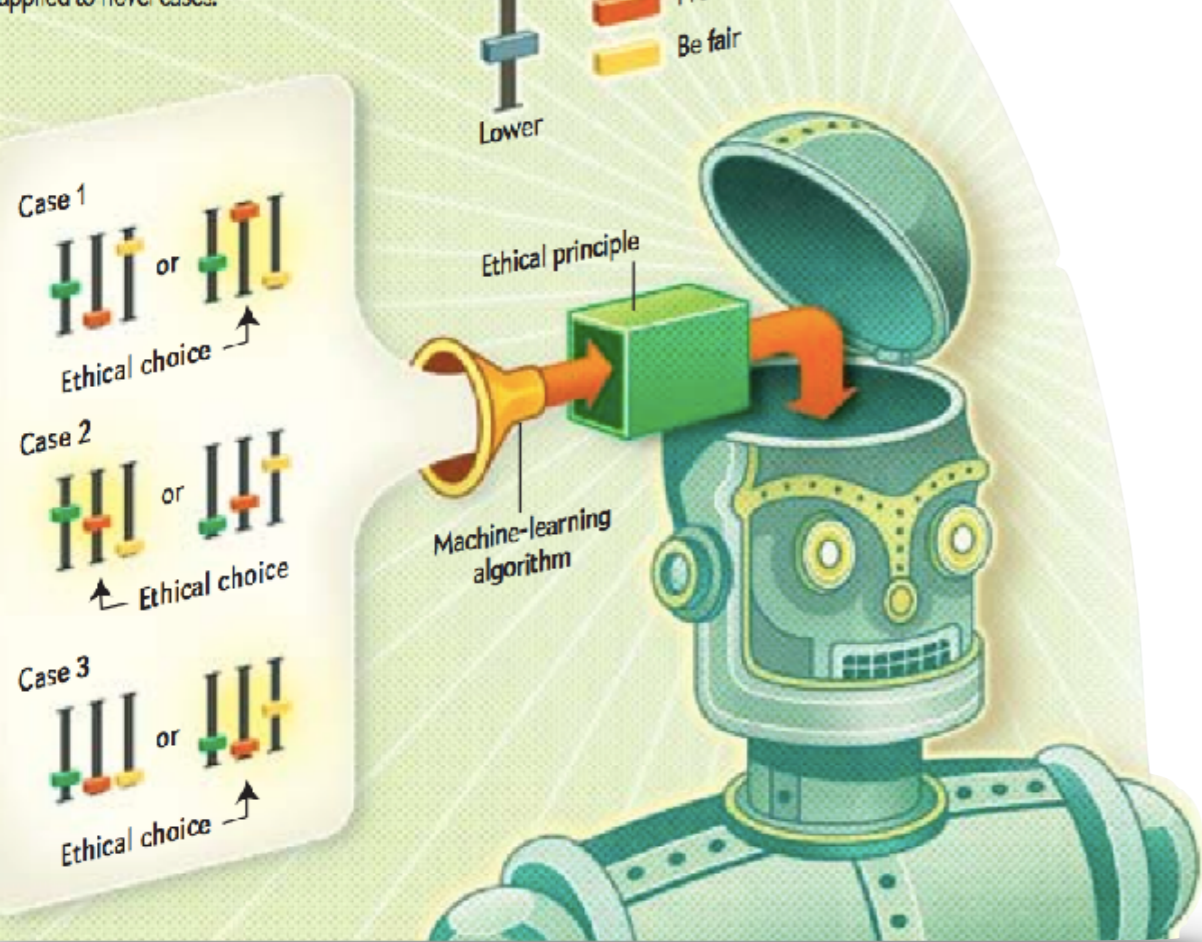
0	0
---	---



Max Planck Institute for Intelligent Systems  
**Autonomous Motion**  
Tübingen Campus







# Toward Ethical Autonomous Systems: A Case-supported, Principle-based Behavior Paradigm

Michael Anderson  
University of Hartford

Susan Leigh Anderson  
University of Connecticut

Vincent Berenz  
Max Planck Institute

This material is based in part upon work supported by the NSF under Grant Numbers IIS-0500133, IIS-1151305 and IIS-1449155.