

Ralf Schneider

intensiver Ideenaustausch mit Prof. Werner, Dr. Reinmuth, Institut für Philosophie

Vati, was hat denn das
Zebra noch gerufen?

Keine Ahnung. Irgendwas
von "Ethik", oder so...



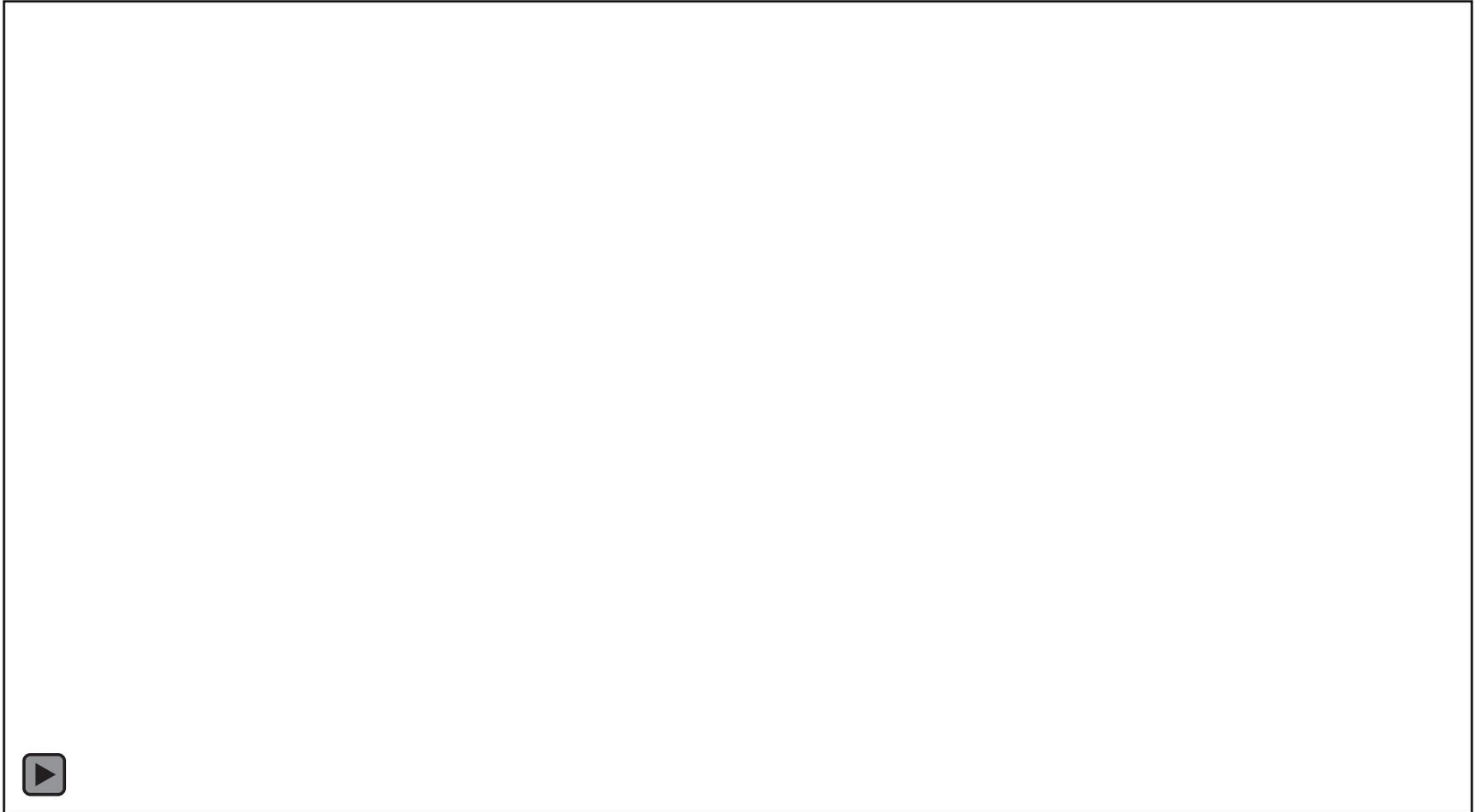


Ethik ist die Wissenschaft vom Handeln der Menschen

- Moralisches Dilemma: Zwickmühle
- mehrere Entscheidungsmöglichkeiten: jede falsch
- Wahl durch Moral bestimmt



Trolley Problem



Philippa Foot, 1967

<https://www.youtube.com/watch?v=MhOJp1DcabM>



3 Faktoren



1. Muss ich selber aktiv eingreifen, oder lasse ich nur gewähren?
2. Welche Konsequenzen hat das eigene Handeln, und zwar beabsichtigte wie auch unbeabsichtigte?
3. Welche gesellschaftlichen Vorstellungen gibt es, wozu ist man moralisch verpflichtet – und was ist tabu?



Hintergrund



Etwa 80 % würden Hebel umlegen (Forschungsgruppe Uni. Barcelona)

Vorstellung aktiv an der Tötung eines Menschen beteiligt zu sein, erzeugt Abwehr und starkes moralisches Unbehagen

Emotionale Distanz ändert Bewertung moralischer Dilemmas

Fremdsprache: Menschen risikobereiter und weniger emotional
→ 44% würden jetzt den dicken Mann von der Brücke schubsen

Anderer Kulturkreis: im ostasiatischen Raum fast nie
Entscheidung dazu, in solche Situationen einzugreifen



Moral, Ethik, Recht



Moral = gelebte Sitten und Traditionen

Ethik = kritischer Diskurs über Moral und Recht

Recht = staatlich sanktionierte Normen, Strafandrohungen:
Geldstrafe , Freiheitsentzug (Todesstrafe) (BGB, StGB)

Vorrang der Moral über Ethik und Recht (Fundamentalismus)

Vorrang des Rechts über Ethik und Moral (Legalismus)

Vorrang der Ethik über Moral und Recht (Ethischer Rigorismus)

→ **Abwägungsprozesse**



Utilitarismus



Utilitarismus (lat. *utilis*=nützlich):

Kriterium der Sittlichkeit ist die Optimierung des Glücks aller Betroffenen aufgrund

einzelner Handlungen (Handlungs-U.)

von Handlungsregeln (Regel-U.)

von Präferenzen (Präferenz-U.)

Utilitätsprinzip: Nutzen für das in sich Gute

Konsequenzprinzip: Konsequenzen einer Handlung

Sozialprinzip: Es geht nicht um das Glück von einzelnen, sondern um das Glück von allen

J. Bentham, Th. Hobbes, J.S.Mill



Deontologische Ethik



Deontologische Ethik (gr. *to deon*=das Erforderliche):

Handlung gilt als sittlich richtig, wenn sie Maximen folgt, die an sich gut sind (kategorischer Imperativ)

es geht um die moralische Handlung an sich, zunächst unabhängig von den Folgen oder Konsequenzen einer Handlung

christliche Ethik, Kant



Diskursethik



Diskursethik:

strittige soziale und politische Fragen sollten nach dem Prinzip einer gewaltfreien, rationalen und allgemein zustimmungsfähigen Lösung aufgrund eines auf Konsens hin geführten Diskurses gelöst werden

von Habermas (1983), Kant, Ch. S. Peirce



Erstes Fazit



Deontologische Ethik: Gesinnung, Demokratischer
Mehrheitsentscheid

Utilitarismus: Nutzen, Frage nach den Konsequenzen

Diskursethik: der kleinste gemeinsame Nenner bringt Konsens

....

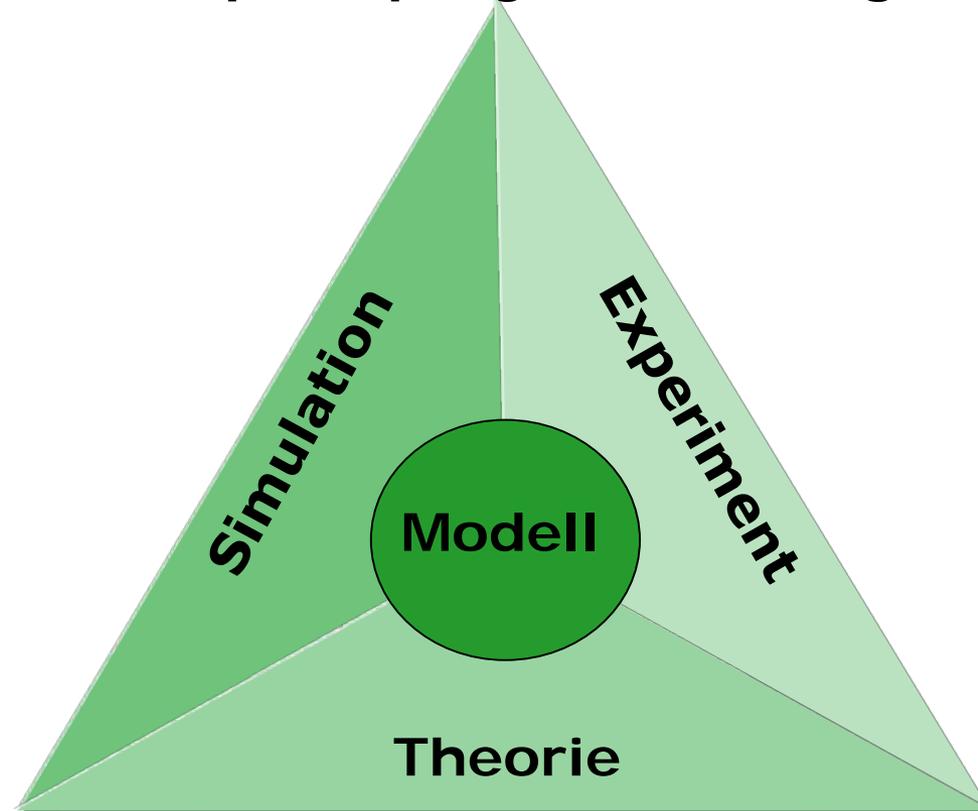
Umsetzung als Ethikmodul??



Computational Science



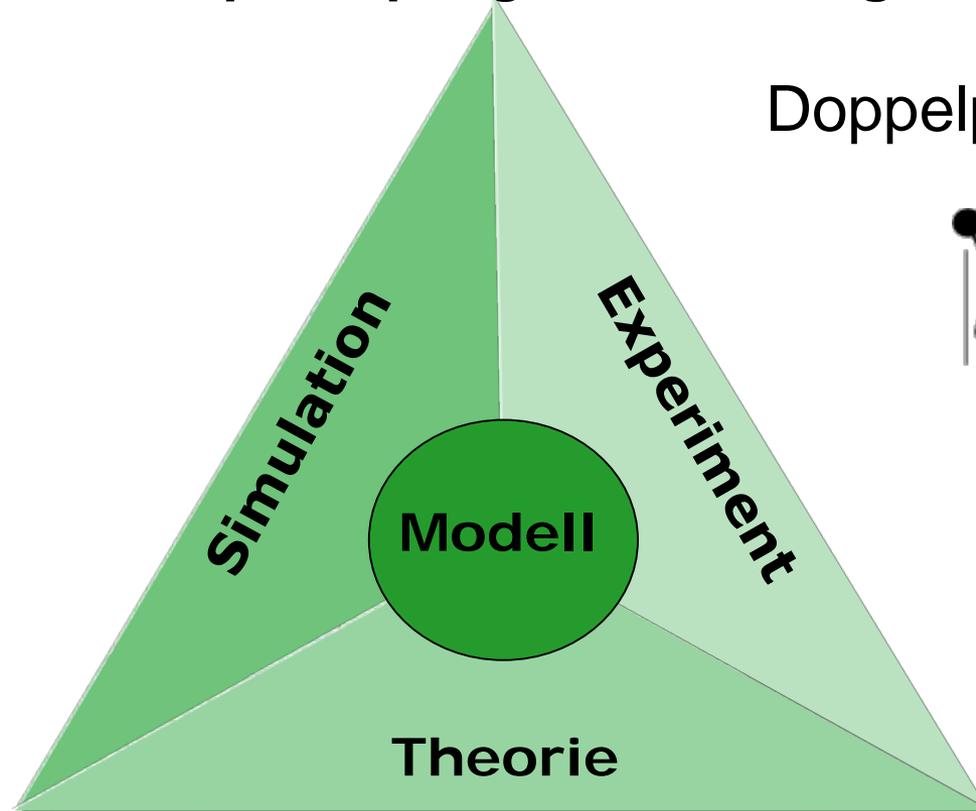
**Architektur
(Computerprogrammierung)**



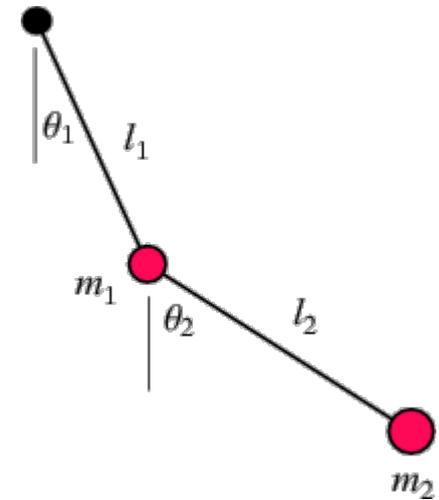
**Algorithmus
(Mathematisches Modell)**

**Anwendung
(Wissenschaftliches Problem)**

Architektur
(Computerprogrammierung)



Doppelpendel:



Algorithmus
(Mathematisches Modell)

Anwendung
(Wissenschaftliches Problem)

$$\dot{\theta}_1 = \frac{\partial H}{\partial p_{\theta_1}} = \frac{l_2 p_{\theta_1} - l_1 p_{\theta_2} \cos(\theta_1 - \theta_2)}{l_1^2 l_2 [m_1 + m_2 \sin^2(\theta_1 - \theta_2)]}$$

$$\dot{\theta}_2 = \frac{\partial H}{\partial p_{\theta_2}} = \frac{l_1 (m_1 + m_2) p_{\theta_2} - l_2 m_2 p_{\theta_1} \cos(\theta_1 - \theta_2)}{l_1 l_2^2 m_2 [m_1 + m_2 \sin^2(\theta_1 - \theta_2)]}$$

$$\dot{p}_{\theta_1} = -\frac{\partial H}{\partial \theta_1} = -(m_1 + m_2) g l_1 \sin \theta_1 - C_1 + C_2$$

$$\dot{p}_{\theta_2} = -\frac{\partial H}{\partial \theta_2} = -m_2 g l_2 \sin \theta_2 + C_1 - C_2,$$

where

$$C_1 \equiv \frac{p_{\theta_1} p_{\theta_2} \sin(\theta_1 - \theta_2)}{l_1 l_2 [m_1 + m_2 \sin^2(\theta_1 - \theta_2)]}$$

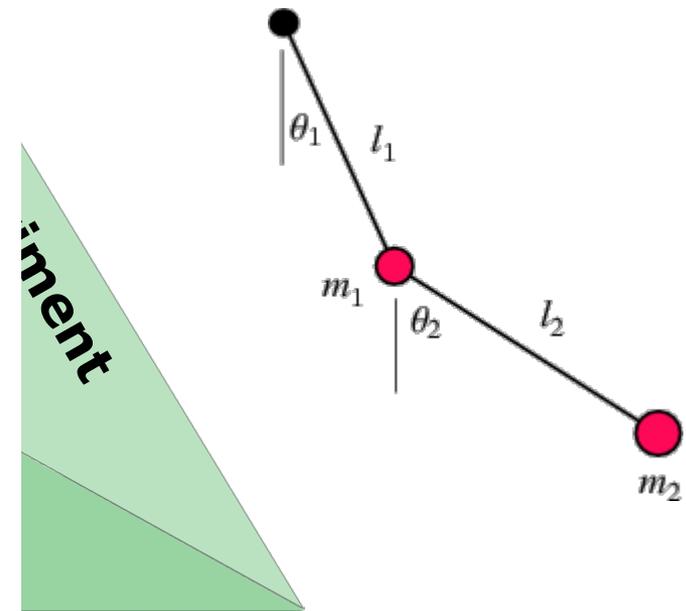
and

$$C_2 \equiv \frac{l_2^2 m_2 p_1^2 + l_1^2 (m_1 + m_2) p_2^2 - l_1 l_2 m_2 p_1 p_2 \cos(\theta_1 - \theta_2)}{2 l_1^2 l_2^2 [m_1 + m_2 \sin^2(\theta_1 - \theta_2)]^2} \sin[2(\theta_1 - \theta_2)].$$

Algorithmus
(Mathematisches Modell)

ierung)

Doppelpendel:



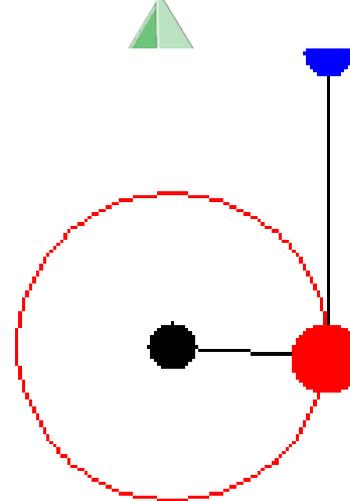
Anwendung
(Wissenschaftliches Problem)



Computational Science



Architektur
(Computerprogrammierung)



Algorithmus

(Mathematisches Modell)

Anwendung

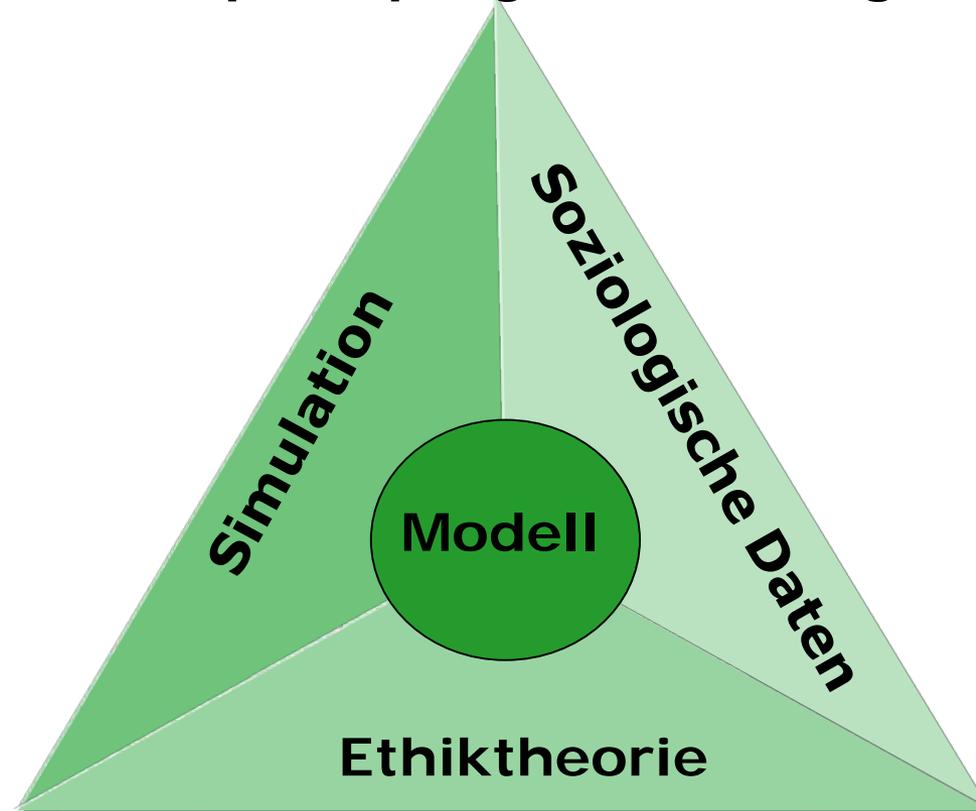
(Wissenschaftliches Problem)



Computational Ethics



Architektur
(Computerprogrammierung)



Algorithmus

(Mathematisches Modell)

Anwendung

(Wissenschaftliches Problem)



Untersuchung von sozialen Normen und Moral im Kontext von Agentensystemen für soziale Gruppen

erfordert algorithmische Erfassung von ethischen und moralischen Theorien im Sinne von „Moralmodulen“ (Entscheidungsmatrix)

Regeln müssen entsprechend einer vorgegebenen Moraltheorie konsistent aufgestellt und programmiert werden (am einfachsten für Pflichtenethik)



Entscheidungen nach einer Hierarchie von Werten und Prinzipien und unter Vorgabe von Regeln

Logikalgebra:

in Abhängigkeit von Eingangsvariablen Ausgangsvariablen;
auch über Wahrscheinlichkeiten möglich (Monte Carlo),
ebenfalls für statistische Bandbreite von Eingangsvariablen

Programmierung der Agenten und ihrer Interaktion: Mischung
aus deterministischen Vorgaben und statistischen Anteilen

Konsequenzen aus den jeweiligen Moraltheorien in
verschiedenen Testszenarien: auch zur Validierung der
Programmierung (Konsistenz mit Ethikmodellen)



Maschinenethik



Roboterethik, Maschinenethik: (teil-)autonome Systeme wie Agenten, Roboter, Drohnen, Computer im automatisierten Handel und selbstständig fahrende Autos

Autonomie von Maschinen: selbstständiges Denken und Handeln (als Subjekte) macht ihre Moral möglich und notwendig

notwendige Moral von Maschinen: macht Autonomie erst möglich, sonst kein autonomes Subjekt

Autonomie von Maschinen: Objekte von Moral, da sie Regeln folgen müssen

**Ethik im Sinne einer
Maschinenethik
erfassbar, falls
normative Ansätze
algorithmisch
verarbeitbar**

