

# Rationality and Cooperation

Julian Nida-Rümelin

Helsinki October 10th, 2007

# Rationality and Cooperation

I Consequentialism .....	3
II Deontology and Decision Theory...	7
III Structural Rationality.....	24
IV Cooperation.....	30
V Structural Rationality and Cooperation.....	43

# Consequentialism

The predominant view on practical rationality is consequentialist:  
Responsibility concerning action is responsibility for actual or probabilistic causal consequences of actions.

→ Objectivism vs subjectivism  
(Bayesianism) in probability theory

# Consequentialism

Decision theoretic utility theory seems to yield a neutral framework which allows to reconstruct all kinds of rational action as consequentialist.

# Consequentialism

The consequentialist paradigm emphasizes acting as some way of changing the world: If a person had not acted this way the world would look different. Therefore a rational person cannot act without exposing some subjective rank-order over possible worlds (propositions as sets of possible worlds).

# Consequentialism

The ideal type of consequentialist practical rationality results in a complete ordering of possible worlds and their lotteries.

# Deontology and Decision Theory

This section is based on my article: „Why Rational Deontological Action Optimizes Subjective Value“ in : *ProtoSociology* 21 (2005), 182 – 193.

# Deontology and Decision Theory

Does rational deontological action optimize subjective value? Can one be a deontologist and at the same time adhere to decision theory as an all-embracing theory of practical rationality? I think the answer to both of these questions is *yes*.

# Deontology and Decision Theory

There are two basic intuitions which frame the bigger part of practical philosophy and which seem to be incompatible. One intuition is *teleological* or more specifically *consequentialist* according to which rational action optimizes its consequences. The other intuition is *deontological* or *rule-oriented* according to which rational action is guided by certain rules.

# Deontology and Decision Theory

(1) Consequentialism as a theory of *rationality* is false

(2) Rational action maximizes subjective value

respectively

(1') Consequentialism as a theory of *morality* is false

(2') Moral action maximizes subjective value.

# Deontology and Decision Theory

- Thesis: The teleological intuition can be upheld not in the original *consequentialist*, but in a weaker *coherentist* form.
- Rational deontological action is compatible with standard decision theoretic axioms. *Rational deontological action optimizes subjective value.*

# Deontology and Decision Theory

Decision theory attributes two real-valued functions to the rational actor: the function of *subjective probability* and the function of *subjective value*.

# Deontology and Decision Theory

Decision theory does not begin with postulating probability and value functions. Rather decision theory begins with describing *basic properties* of preference relations.

# Deontology and Decision Theory

Four substantive postulates:

- I. Connectivity
- II. Transitivity
- III. Monotonicity
- IV. Continuity

# Deontology and Decision Theory

*THOU SHALT NOT STEAL.*

The conflict between deontology and consequentialism is apparent: Even if stealing in some cases might have overall good consequences, you are not allowed to steal.

# Deontology and Decision Theory

- Three possible states of the world  $x, y, z$ .
- Three alternative feasible actions  $u, v, w$
- The person prefers  $x$  to  $y$  and  $y$  to  $z$ .
- $u$  is an act of stealing, whereas  $v$  and  $w$  are no acts of stealing (and are not forbidden by other deontological rules).
- It is not possible to optimize consequences and at the same time conform with the deontological rule in question.

# Deontology and Decision Theory

The rational person's preferences regarding states of the world do not necessarily determine the person's preferences regarding actions.

# Deontology and Decision Theory

*The conceptual exclusion of deontological rationality is unacceptable for any adequate theory of practical reasons and it is unnecessary in order to preserve the conceptual frame of decision theory.*

# Deontology and Decision Theory

*The wedge between choice (of actions) and preference (between states of the world).*

- *Not all actions can be represented by their prospects in the sense of probability distributions over the set of their consequences, if “consequences” are understood as “states of the world” or “sets of states of the world”.*
- *It is not universally true that one can substitute actions for prospects within a rational person’s preferences. Deontological actions cannot be represented by their prospects in a rational persons’ preference relation.*

# Deontology and Decision Theory

- Ramsey-Coherence
- Transitivity of preferences is a requirement of practical coherence that can not be dismissed if we move from consequentialist reasons to deontological ones.
- The same holds for *reflexivity* and *completeness*.

# Deontology and Decision Theory

- Deontologically motivated preferences require *comprehensively described* actions.
- Prospects of actions in the usual consequentialist sense do not include enough information for the deontological agent to rank his actions.

# Deontology and Decision Theory

- A rational person has coherent preferences regarding probabilistic relations between actions.
- *Extended Monotonicity Axiom.*
- *Extended Continuity Axiom.*

# Deontology and Decision Theory

*A rational person following deontological reasons optimizes subjective value, because her preferences adequately i.e. comprehensively described are Ramsey-coherent.*

# Structural Rationality

- A fictitious biological example: a society of ants gradually gaining consciousness.
- Ants-society governed by a complex and perfect (or even optimal) system of rules that result in specific types of interactions and cooperations. The individual ants follow these behavioural rules non-intentionally (genetic outfit).

# Structural Rationality

Suppose a singular ant gains consciousness.

Suppose she continues to conform to the rules.

(1) She develops a certain disposition to follow the system of rules. Combining practical reason (in fostering an optimal system of rules) with constraining free choice.

(2) Let us assume the ant remains free. If the ant decides to act in conformity for the reason that she takes the structure of ant-behaviour established by the rules as optimal, this would be an example of what I will call *structural rationality*.

# Structural Rationality

The main characteristics of structural rationality:

- (1) 'Structurally rational' is a predicate to be applied to (individual) actions (token), i. e. to the primary objects of decision.
- (2) An action can be structurally rational only if it is the object of free choice, i.e. it may not be dispositionally or otherwise determined.
- (3) A positive valuation regarding the structure by the actor. 'Valuation' here is not meant necessarily in an axiological sense. The actor might foster or preserve a structure for deontological reasons, too.
- (4) Intentionality. The act is chosen with the motivating intention to conform with the structure.

# Structural Rationality

Two possible attitudes:

## (1) Solipsistic attitude

Cf. Immanuel Kant: *Grundlegung zur Metaphysik der Sitten*.  
Riga 1785

## (2) We-attitude

Cf. Raimo Tuomela: *A Theory of Social Action*. Dordrecht  
1984, Chap. 4

# Structural Rationality

If we ascribe we-intentions of the Tuomela-kind, every ant would have a motivation to conform with the established system of rules of behaviour, i. e. to act *structurally rational*.

# Structural Rationality

Structural Rationality has two aspects.

- One is the role of reasons for action. Reasons establish structures of individual forms of life and collective forms of interaction. Reasons constitute a form of life. Cf JNR: *Strukturelle Rationalität*. Stuttgart 2001
- The second aspect is game theoretic: structures of interaction are essential for rational action

# Cooperation

## 2x2-prisoners dilemma (PD)

If one of the prisoners ( $B$ ) confesses, then it is in the other prisoner's ( $A$ ) interest to confess as well. If  $B$  does not confess, then it is nevertheless in  $A$ 's interest to confess, because then  $A$  will be let off free as supergrass witness. The same reasoning is true for  $B$ , i.e., completely irrespective of what the other prisoner does, it is always in the interest of each to confess

# Cooperation

## The Prisoners Dilemma

A: DC CC DD CD

B: CD CC DD DC

	CC	DC
CD	DD	

# Cooperation

Four classical solutions

(1) The Irrationality Solution

(2) The Shaftesbury Solution

(3) The Hobbesian Solution

(4) The Kantian Solution

# Cooperation

## The Assurance Game

A: [C,C], [D,C], [D,D], [C,D]

B: [C,C], [C,D], [D,D], [D,C]

# Cooperation

## The Other Regarding Game (OR)

A: [C,C], [C,D], [D,C], [D,D]

B: [C,C], [D,C], [C,D], [D,D]

# Cooperation

The PD structure in terms of the *interest structure* of the interaction situation.

		B	
		C	D
A	C	3,3	1,4
	D	4,1	2,2

# Cooperation

Egoistic preferences

A:  $(4,1) > (3,3) > (2,2) > (1,4)$

B:  $(1,4) > (3,3) > (2,2) > (4,1)$

# Cooperation

Justice

A:  $(3,3) > (2,2) > (4,1) > (1,4)$

B:  $(3,3) > (2,2) > (1,4) > (4,1)$

# Cooperation

Justice combined with altruism

A:  $(3,3) > (2,2) > (1,4) > (4,1)$

B:  $(3,3) > (2,2) > (4,1) > (1,4)$

# Cooperation

Altruism superseding justice

A:  $(3,3) > (1,4) > (4,1) > (2,2)$

B:  $(3,3) > (4,1) > (1,4) > (2,2)$

# Cooperation

Pure altruism

A:  $(1,4) > (3,3) > (2,2) > (4,1)$

B:  $(4,1) > (3,3) > (2,2) > (1,4)$

# Cooperation

## Masochism

A:  $(1,4) > (2,2) > (3,3) > (4,1)$

B:  $(4,1) > (2,2) > (3,3) > (1,4)$

# Cooperation

## Sadism

A:  $(4,1) > (2,2) > (3,3) > (1,4)$

B:  $(1,4) > (2,2) > (3,3) > (4,1)$

Further (all symmetrical) variants are discussed in JNR: *Economic Rationality and Practical Reason*. Dordrecht: Kluwer 1997. Chap 8

# Structural Rationality and Cooperation

- Practical Reason requires that individual actions fit into a system of rules that the agent rationally can accept.
- This system of rules cannot be deduced from the principles of an ethical theory, but are constituted by a form of life and the reasoning that represent it (Wittgensteinian perspective).
- Cooperative reasons cannot be transformed into optimizing outcomes - however the valuation - they represent a paradigmatic form of structural rationality.

# Structural Rationality and Cooperation

$$v_i(a) = \text{fct} (O, \text{prob} (O), S^*)$$

a: action of the agent i,

$v_i(a)$  : value of the action a for agent i.  $i = 1, 2, \dots, n$

O: outcome of the game;  $o_i^j$  outcome for i, given the solution  $j = 1, 2, \dots, m$

$O = \langle O_1, O_2, \dots, O_n \rangle$ ;  $O_i = \langle o_i^1, o_i^2, \dots, o_i^m \rangle$

$S^* = \langle S, I \rangle$ ; S Game Format; I Intentional Structure

# Structural Rationality and Cooperation

$v_i(a) = \text{fct } (o_i)$  conventional game theory, one solution

$v_i(a) = \text{fct } (\langle o_1, o_2, \dots, o_{i-1}, o_i, o_{i+1}, \dots, o_n \rangle)$  ethical game theory, one solution.

$v_i(a) = \text{fct } (o_i^1, o_i^2, \dots, o_i^m)$  conventional game theory, games with m solutions. In analogy ethical game theory with m solutions.

$v_i(a) = \text{fct } (\langle o_i^1, \text{prob } (o_i^1) \rangle, \langle o_i^2, \text{prob } (o_i^2) \rangle, \dots, \langle o_i^m, \text{prob } (o_i^m) \rangle)$  bayesian game theory, m solutions. In analogy ethical bayesian game theory

$v_i(a) = \text{fct } (O, \text{prob } (O))$  bayesian game theory

$v_i(a) = \text{fct } (O, \text{prob } (O), S^*)$ , with  $S^* = \langle S, I \rangle$ . **structural rationality**

$v_i(a) = \text{fct } (O, \text{prob } (O), S^*, S^{**}, S^{***}, \dots)$  **extended structural rat.**

$v(a) = v_i(a) = \text{fct } (O, \text{prob } (O), S^*)$ , with  $S^* = \langle S, I \rangle$  **objectivist str rat**

# Structural Rationality and Cooperation

- $v(D) = \text{fct}(2)$  ; conventional game theory, one solution: DD
- $v(D) = \text{fct}(\langle 2, 2 \rangle)$ ; ethical game theory, one solution
- $v(D) = \text{fct}(\langle 2, 2 \rangle, \langle 4, 1 \rangle)$  ethical game theory, two solutions
- $v(D) = \text{fct}(\langle 2, \text{prob}(DD) \rangle, \langle 4, \text{prob}(DC) \rangle)$  ; Bayesian game theory
- $v(D) = \text{fct}(\langle \langle 2, 2 \rangle, \text{prob}(DD) \rangle, \langle \langle 4, 1 \rangle, \text{prob}(DC) \rangle)$  ; ethical B. gt.
- $v(D) = \text{fct}(\langle \langle 2, 2 \rangle, \text{prob}(DD) \rangle, \langle \langle 4, 1 \rangle, \text{prob}(DC) \rangle, S^*, S^{**}, S^{***}, \dots)$  (extended) structural rationality.

# Structural Rationality and Cooperation

- $v(C) = \text{fct}(1)$ ; conventional game theory, one solution: DD
- $v(C) = \text{fct}(\langle 1, 4 \rangle)$ ; ethical game theory, one solution etc.
- $v(C) = \text{fct}(\langle 1, 4 \rangle, \langle 3, 3 \rangle)$  ethical game theory, two solutions
- $v(C) = \text{fct}(\langle 1, \text{prob}(CD) \rangle, \langle 3, \text{prob}(CC) \rangle)$ ; Bayesian game theory
- $v(C) = \text{fct}(\langle \langle 1, 4 \rangle, \text{prob}(CD) \rangle, \langle \langle 3, 3 \rangle, \text{prob}(CC) \rangle)$ ; ethical Bayesian game theory
- $v(C) = \text{fct}(\langle \langle 1, 4 \rangle, \text{prob}(CD) \rangle, \langle \langle 3, 3 \rangle, \text{prob}(CC) \rangle, S^*, S^{**}, S^{***}, \dots)$  (extended) structural rationality.

# The End

$$v_i(a) = \text{fct} (O, \text{prob}(O), S^*, S^{**}, \dots)$$

**Cooperative action can be rational, depending on the reasons in favour of it.**

**Cooperative action can be rational only in terms of structural rationality.**

**Structural rationality is different both from consequentialist optimization and from principles guided deontology.**

**Structural rationality is Ramsey-coherent and can be embedded in any coherent account of deontological ethics (extended structural rationality).**

# The End

Thank you for listening

[www.Julian.Nida-Ruemelin.de](http://www.Julian.Nida-Ruemelin.de)

Universität München