

Julian Nida-Rümelin, ZiF March 5-7, 1998. ZiF4.doc

Rationality: Coherence and Structure¹

I Introduction

The ideal rational person has coherent preferences, i.e. preferences which fulfil the decision theoretic axioms of rational choice. A person whose preferences are e.g. intransitive can not be perfectly rational. The analogy seems to be true for every single axiom which constitutes the theory of expected utility maximization. The theory of expected utility maximization is a weak theory of rationality. It is unable to discriminate between good and bad reasons for acting. It only qualifies revealed preferences as coherent or incoherent.

Revealed preferences are the result of practical reasoning. In the last resort rationality is constituted by actions which are guided by good reasons. Therefore we do not know whether a person is rational if we know that the person maximizes expected utility. If a person does not maximize expected utility we know that she has incoherent preferences and is not rational therefore.

Let us call this view the compatibility assumption: Whatever the concrete reasons which guide a person's actions, action-guiding rational preferences fulfill the standard axioms of rational choice. Rational action must be compatible with a description for which the person maximizes expected utility.

The intuitive underpinning of the compatibility assumption is not consequentialist but strictly coherentist. It is not based on the idea that rational action means to optimize consequences, because there are many types of good reasons for action which can not be reconstructed within a

¹ I am grateful that I had the opportunity to discuss my arguments with John Broome, Ed McClennen, Michael Slote, Wolfgang Spohn and others during the conference. I thank Martin Rechenauer and Thomas Schmidt for reading this paper and proposing valuable amendments.

consequentialist theory of rationality². Certainly this is only a prima facie argument against consequentialism. It does not prove that consequentialism is wrong. Since prima facie good reasons tell against consequentialism, consequentialism could be taken to be true only as a result of some convincing theory. In this paper I do not exclude that such a convincing theory could exist. What I take to be granted is that there are prima facie arguments against consequentialism such that consequentialism can not be the starting point of the argument.

Two examples might suffice for rendering the assumption plausible that many of our good reasons for action - moral and extra-moral ones³ - are non-consequential.

First Example.

That Tom has promised at time *t* to come can be a good reason for Tom to come at some later time *t'*. It is not necessary to spell the conditions out under which a given promise in fact is a good reason to keep it (transforming "can" into "is" within the above sentence) in order to see that these conditions can not be confined to outcome-considerations alone. If Tom keeps his promises only in case he expects the outcomes to be optimal, he has not really understood what it means to give a promise. If Suzan knows that he keeps his promises only in case he expects the outcomes to be optimal she will not trust him. In this case promise-giving would lose its power to coordinate intentions and actions adequately. If we take it to be that the intention to keep the promise is part of the act of (genuine) promise giving - as John Austin assumed in his speech-act analysis⁴ - then a consequentialist intention would even be incompatible with promising. This does not exclude that certain or probable

² cf. J.N-R. Why Consequentialism Fails, in Contemporary Action Theory. ed.... and in more detail id. Kritik des Konsequentialismus, München: Oldenbourg 1993, 2nd. (paperback-) ed. 1995, part III and IV.

³ In fact, I think that this dichotomy is one of the inventions of modern ethical theory which should be given up (cf. my "Subjective and Objective Reasons", 1998), but for the moment we do not need to introduce this argument here.

⁴ cf. John Langshaw Austin, How to Do Things with Words, ...!!!!

consequences of keeping the promise vs. not keeping it are relevant for determining whether Tom has a good reason to come. It does exclude though that the outcomes of keeping the promise vs. not keeping it are exclusively relevant for determining whether Tom has a good reason to come.⁵

Second Example

Tom might have the intention to cooperate in a specific situation with Suzie, expecting that Suzie has the same intention⁶. Let us assume that cooperation can be defined by using a matrix of outcomes which shows the structure of a Prisoner's Dilemma. Tom might intend to cooperate although he is aware of the outcome structure of the game he plays. It is not necessary that Tom attributes some kind of intrinsic value to the act of cooperation itself, he might deliberately choose a dominated strategy in cooperating. In this case Tom would not maximize the expected value of outcomes. Nevertheless Tom could have coherent preferences. The intention to cooperate should be compatible with having coherent preferences. According to the definition above a coherent preference relation fulfills the decision-theoretic axioms. It should therefore be conceptually possible that Tom (in cooperating) maximizes expected utility although he chooses (in cooperating) the dominated action regarding outcomes. This might seem strange at first sight but it leads to the core of an adequate analysis of rational action which drives a wedge between consequentialism and coherentism.

The observation that there are good reasons for action which are not consequentialist, i.e. which recommend actions which do not maximize the expected values of their outcomes, can lead to quite different reactions:

⁵ Note that we do not have to subscribe to some kind of institutional analysis like the one presented in Searle (Speech Acts, ch. 8 ??...!!!). It is not necessary to assume that the fact that X has uttered l which under some conditions means that X has promised to f alone is sufficient for X having a good reason to f. The naturalist fallacy debate is of no relevance here.

⁶ One might suppose that with this description we opted implicitly for an outcome-structure of the Assurance Game type. This is not the case, as I tried to show in „Metapreferences or Practical Reason...“.

(1) One might assume that those prima facie reasons which can not be integrated are no genuine good reasons (this account could be called „consequentialist“). The problem with this account is that some of our most central types of reasons would have to be excommunicated.

(2) One might confine the range of application of rational choice and exclude moral and other types of good reasons for action (this could be called the „narrow rationality account“). Rational choice theory would then not be an all-embracing theory of practical rationality any more.

(3) One might give up some of the axioms constitutive for standard rational choice theory as it is done e.g. by McClennen in his theory of resolute choice (this account could be called „revisionist“)

(4) One might redesign the conceptual framework, i.e. re-interpret the basic concepts of rational choice such that non-consequentialist reasons for action can be integrated (I call it the account of „structural rationality“ for reasons which might become clearer later on).

The observation that there are good reasons for action which are not consequentialist, i.e. which recommend actions which do not maximize the expected values of their outcomes, can lead to quite different reactions:

In this paper I delineate some arguments which speak in favor of the forth reaction.

II. Weak Coherentism

We call a preference relation „weakly coherent“ or „Ramsey-coherent“ if it accords with the standard decision-theoretic axioms. „Weakly“ because what is rational depends on the whole frame of conative and epistemic attitudes, but this frame is not considered in testing whether a preference relation accords with the decision-theoretic axioms. An action is

rational in the full sense iff it coheres adequately with this frame.

Preferences which do not accord with the standard decision-theoretic axioms are not coherent. But Preferences which do accord with the standard decision-theoretic axioms can still be incoherent with the broader frame of conative and epistemic attitudes. Preferences which accord with the standard decision-theoretic axioms are weakly coherent. The standard decision-theoretic axioms (collectively) are necessary and sufficient criteria for weak coherence, but they are only necessary criteria for coherence in the broader sense.

We can think of many different types of conative attitudes desires, wishes, hopes, intentions, preferences and of many different theories how to systematize this plurality, but it seems that, independently from the type of the respective conative attitudes and their systematization, there is an initial plausibility of (weak) coherentism or Ramsey-coherentism, because (and insofar as) conative attitudes result in concrete actions. There might be irresolvable conflicts between different desires, hopes and wishes, there might even be genuine moral dilemmas, but in the end we have to integrate our conative attitudes such that we are able to act. If, for example, our action guiding (overall-) preferences were incomplete then in some situations we could not act or - to put it differently: our actions would lose their basis in our conative attitudes. The violation of each singular axiom would have the same devastating result for the agency of individuals, who try to decide rationally. Preferences - however they may be constituted - have to be Ramsey-coherent if they are thought to guide action.

Irrationality of action guiding preferences has two main sources:

- (1) how they are based on conative (and epistemic) attitudes and
- (2) on which conative (and epistemic) attitudes they are based.

But even if the action guiding preferences can not be criticized for being inadequately based on conative attitudes or for being determined by inadequate conative attitudes (because the determining conative attitudes are inadequate), they can still be criticized if they are (weakly) incoherent.⁷

Since Ramsey-coherent preferences can be represented by a real valued utility function (determined up to positive linear transformation) such kind of quantitative representability reveals that a minimal requirement of ideal rationality is fulfilled. This explains the use of the term "compatibility-assumption": The plurality of our reasons for actions has to be made coherent such that the action guiding overall preferences conform with the axioms of rational choice. If they conform we do not know whether these preferences are rational but if they do not conform we know that at least one of these preferences is irrational. The compatibility is between reasons for actions on the one hand and coherence of action guiding preferences on the other. If our reasons for actions are such that they are incompatible with Ramsey-coherent action guiding preferences they have to be modified, whatever their concrete content may be⁸. It can not e.g. be rational to have action guiding (overall-) preferences which are not transitive. No type of reason for action can be held responsible for violating reflexivity, completeness, transitivity, continuity, reduction, substitution or monotonicity.⁹

Some familiar arguments in favour of the rational choice axioms like the money pump argument or dutch book argument

⁷ There is a critical point, here: Is this statement also true for the ideal theory of practical rationality? Or should we postulate that a complete theory of good reasons for action is such that weak incoherence of action-guiding preferences is incompatible with being adequately based on conative attitudes and for being determined by adequate conative attitudes?

⁸ Those who argue that there are genuine moral dilemmas think that our moral reasons for action can not be made coherent in this sense, cf. Gowans, Moral Dilemmas.!!!!

⁹ Contrary to the first three mentioned axioms the last four mentioned axioms make use of probability considerations and this might undermine their initial plausibility. I think they are equally plausible, but we can not go into it here. I argue that even within a deontological framework probability considerations can be integrated in. §.51 of my Kritik des Konsequentialismus, München 1993, second ed. 1995

might inhibit a proper understanding of my argument. I postulate an initial plausibility of the axioms of rational choice as minimal coherence requirements of action-guiding preferences, whereas the traditional defenders of the axioms of rational choice use consequentialist arguments. But since many of our good reasons for actions are not consequentialist we should not rely on consequentialist, but on coherentist intuitions. Our coherentist intuitions are in favour of the compatibility assumption. There is a prima facie justification of the rational choice axioms which does not depend on consequentialist metaphysics.

Whereas standard arguments in favour of the axioms of rational choice are based on consequentialist intuitions there is a kind of critic of the axioms of rational choice which also is consequentialist in spirit. If an agent, who maximizes expected utility, is in the end worse off than an agent who doesn't consequentialism seems to tell against the axioms of rational choice. To avoid confusions it is helpful to discriminate terminologically between „general consequentialism“ as the view that consequences and their value determine exclusively what is rational and the more specific view that a concrete singular act is rational if and only if it's consequences are optimal - let us call this latter view „agency consequentialism“.

At first sight it seems that this discrimination does not make much sense. Agency consequentialism is obviously a specialization of general consequentialism. It seems that a general consequentialist must be agency consequentialist, too. The reversal though does not hold, naturally. Agency consequentialism seems to be implied by general consequentialism.¹⁰

The discrimination between agency consequentialism and general consequentialism makes sense only because there are self-defeating arguments against general consequentialism. Tom, who

¹⁰ If one adopts the plausible stance that only concrete actions of natural individuals can bear the predicate „rational“, the consequentialist account had to be confined to agency consequentialism anyway. Every different usage would on this account be at it's best acceptable as merely metaphorical.

judges the course of his life up to the present using the time integral of some value function which - say - attributes to his mental state at every point of time a real number representing it's subjective quality, realizes that he would have done much better if he had refrained from choosing every singular action such that it optimized it's causal (including probabilistic) consequences. Consequentialism applied to whole courses of life is incompatible with agency consequentialism. General consequentialism is self-defeating.

All forms of consequentialism assume that there is no intrinsic value of rules or types of actions and that in the last resort it is utility or preference fulfillment which is decisive whether an action is rational or not. In this weaker sense one can be consequentialist without being in the stricter sense consequentialist, i.e. agency-consequentialist. A person, acting rational in the sense of agency-consequentialism can be overall irrational regarding overall individual preference fulfillment. This difference is used, e.g. by Edward McClennen in his theory of resolute choice which uses general consequentialist arguments against the specific form of agency-consequentialism. I agree with McClennen that practical rationality is not the result of point-wise maximization, in my words: Agency consequentialism is no convincing general theory of practical rationality. But (other than McClennen) I am convinced that this this does not force us to give up the theoretical core of decision theory, if only we rely on a strictly coherentist interpretation (or foundation) of it's axioms.

The observation that there rational actions which do not maximize the expected values of their outcomes, leads McClennen to give up axioms constitutive for standard rational choice theory. Whereas I plea for redesigning the conceptual framework, i.e. re-interpret the basic concepts of rational choice such that non-consequentialist reasons for action can be integrated (compatibility assumption). The theory of resolute choice favors the third reaction of the above mentioned whereas I think we should strive at the fourth.

My justification of the compatibility assumption consists of two elements. The first element is to stress its initial plausibility and the second is to show that arguments which tell against the compatibility assumption are not convincing. In a recent illuminating article A. Sen¹¹ has developed arguments which if they were stringent would undermine the initial plausibility of the compatibility assumption. In order to refute them as arguments against the compatibility assumption I delineate an alternative interpretation of the rational choice framework. I hope that in doing this it shall become transparent how arguments of this type in general are to be answered.¹²

II Comprehensive Description

Let us sketch the idea of a "comprehensive" description of decisions using one of Sen's examples: you arrive at a garden party, and can readily identify the most comfortable chair. You would be delighted if an imperious host were to assign you to that chair. However, if the matter is left to your own choice, you may refuse to rush to it. You select a "less preferred" chair. The reason for this behavior is that you want to be polite and that you take it as a politeness-constituting-rule not to take the only most comfortable chair in such situations. It is not the case that you merely want to give the impression to be polite, neither do you fear any bad consequences if you were not polite - you just want to be polite. To want to be polite can not be excluded as a conative attitude which may become reason-giving in adequate circumstances. To strive at being polite can not be charged with being an irrational attitude, I suppose. This conative

¹¹ A. Sen, Maximization and the Act of Choice, Econometrica 65 (1997) 745-779

¹² In fact I think that there is not only an underdetermination of theories given the empirical data, there is also an underdetermination of application. This underdetermination can take the form of different possible interpretations of basic terms. Those who argue against the compatibility assumption stay to a more conventional interpretation as it is mostly used in economic and other social science applications of rational choice theory. Lastly it is theoretical fruitfulness of conceptual frameworks which decides.

attitude is determined by accepting (and finally acting in accord with) those rules which constitute politeness.

In this example there seems to be a chooser dependence involved which we can not get rid of in passing from self-oriented (egotist) preferences to an all-comprising ones which regard (i.e. consequentialist) the whole situation including the distribution of chairs to all the other guests since we additionally assumed that you prefer to have the most comfortable chair if you get it as a result of another persons choice (the imperious host).

To accomodate preferences of this kind it is necessary to take the act of choice itself into consideration. In this example it is obviously relevant who chooses even if the outcomes remain unchanged. A comprehensive description of the situation should therefore include the information who chooses. Such a description can not be integrated into consequentialism, because from a consequentialist point of view it is the subjective value of consequences alone (i.e. the state of affairs as it results from the respective choice) which determines rational choice. If polite behaviour can be rational and if polite behavior implies chooser dependence, then it can not be described by reference to consequences (in the technical described) alone. The rationality of politeness requires a non-consequentialist conceptual frame: We have to know more about the options than merely their consequences. If actions are described by an exclusive reference to lotteries which have genuine consequences as outcomes, it is impossible to integrate chooser dependence into the rational choice framework. And if it is impossible to integrate chooser dependence, it is impossible to characterize this kind of polite behavior as rational. Since rational agents ought to be polite like any other agents, the theory of rational choice should be modified such that conceptually they can be polite (this is a peculiar case of OUGHT implies CAN).

Since there is no justification for the assumption that to have chooser dependent preferences implies to have incoherent preferences, this modification should not run against the core of rational choice theory. An adequate theory of rational

choice should allow for including preferences of this kind (i.e. chooser dependent preferences). We should not accept a theory which implies that chooser dependent preferences necessarily lead to a preference relation which is Ramsey-incoherent. Chooser dependent preferences have to be described such that the corresponding subjective preference relation of the rational person is coherent. A strictly coherentist interpretation of the decision-theoretic conceptual frame should be compatible with all kinds of motivation. It is not the content (the motive on which a preference is founded), but the form in which it relates to other preferences which is decisive, whether it can be integrated into a coherent preference relation. To have chooser dependent preferences, which is one kind of rule sensitive preferences, does not necessarily mean to have incoherent preferences.

In order to clarify the formal implications let X be the set of alternatives where each alternative is characterized by the respective seat you are seated on. Then it seems that the principle of revealed preference is violated in our example: You have a preference for the state of affairs in which you sit on the most comfortable chair, but you choose to sit on another chair.

The principle of revealed preference:

$$\forall x, y \in X: \langle x, y \rangle \in g(i) \leftrightarrow [X^\circ \subseteq X \wedge x \in C(X^\circ) \wedge y \in X^\circ]$$

g : preference structure

$g(i)$: (binary) preference ordering of i .

$i \in K$: the individual person out of the collective K

X : optional set, set of alternatives

C : choice function (representing the agent's actual choices)

Let us take a new situation into account in which another equally comfortable chair is added. In this situation you could choose one of the two most comfortable chairs knowing that there is still another most comfortable chair which can be offered. With this choice two quite basic conditions of

rationality - the principle of contraction coherence a and the principle of conjunction coherence g - seem to be violated.

Principle of contraction coherence a:

$$\forall x \in X: \forall S, T \in \text{Pot}(X):$$

$$[S \subseteq T \wedge x \in S \wedge x \in C(T) \rightarrow x \in C(S)]$$

Principle of conjunction coherence b:

$$\forall x \in X \forall S, T \in \text{Pot}(X):$$

$$[x \in C(S) \wedge x \in C(T) \rightarrow x \in C(S \cup T)]$$

Whereas the expansion coherence principle (g) is not violated in this example

$$\forall x, y \in X: \forall S, T \in \text{Pot}(X):$$

$$[S \subseteq T \wedge x, y \in C(S) \rightarrow [x \in C(T) \rightarrow y \in C(T)]]$$

We said that being polite should not be necessarily irrational. If being polite results in accepting the rule that one should not choose the most comfortable chair, following this rule should be compatible with being rational. But also it would not be justified to postulate that preferences which are the expression of politeness, i.e. which you have because you wish to conform with this rule (not choosing the only most comfortable chair) should result in a respective modification of how to rank overall states of affairs in order to save the conceptual framework of rational choice. It is the acceptance of this respective rule of politeness and not some modification of how to value state of affairs (i.e. consequences of one's actions) which motivates these preferences. It is obvious that the supposed rationality of these preferences is a result of some inadequate description. We should neither accept that preferences motivated by some rule of politeness have to be irrational nor that it can be rational to have preferences which violate the principles of expansion or contraction coherence. The way out is to redescribe these preferences motivated by this rule of politeness such that they can be rational and coherent.

This redescription can be done in the following manner: a comprehensive description of the situation considers the difference between

(1) a choice between alternatives where one of these alternatives results in violating the rule of politeness and the other doesn't and

(2) a choice between alternatives which do not include violation of this principle of politeness.

In some cases expanding or contracting the set of alternatives can result in a change from the first type of choice to the second. In this case it would be inadequate to apply the principles of expansion and contraction coherence because the elements of the respective sets change if the set of alternatives expands or contracts. It would be inadequate because it makes a difference whether

(1) to decide between x and y in case choosing x violates a rule of politeness and choosing y doesn't or

(2) to decide between x and y in case neither choosing x nor choosing y violates a rule of politeness.

Expanding the set of alternatives may change the alternatives between to choose if they are recognized in a sufficiently comprehensive manner. A conceptual framework of rational choice which is committed to agency-consequentialism does not allow for a sufficiently comprehensive description. It precludes an adequate description of many choice-situations. In some cases it renders preferences irrational which aren't irrational. To be polite is not necessarily irrational, but in some cases being polite is not compatible with agency-consequentialism. If the alternatives are comprehensively described being polite does not violate the rationality principles a or g.

III Ramsey Coherence

Let us turn back to the compatibility assumption. Is there any reason to suppose that polite preferences are not Ramsey-coherent? In fact properties of choice functions and properties of corresponding preference relations are closely related.¹³ For example: If the choice function is well-defined contraction coherence is sufficient for rendering the corresponding preference relation acyclical. On the other hand if the preference relation is acyclical the corresponding choice function is well-defined. If the conditions a, b and g are fulfilled for some well-defined choice function the corresponding preference relation is an ordering. Therefore, Ramsey-coherence would be secured if we had no reason to assume that polite preferences violate the rationality conditions a, b and g. To be sure, this argument does not prove that polite preferences of this kind in general fulfill a, b and g but since we rejected the thesis that polite preferences are incompatible with basic principles of coherence there seems to be no good reason anymore to doubt the compatibility assumption (given its initial plausibility).

At this stage we should include Broome's Argument and draw the line which separate his and Vallentyne's view from mine.¹⁴ Broome, too, rejects arguments which tell that some principle of Ramsey-coherence is violated. But Broome (and Vallentyne) understand their rejections as a general justification of a teleological theory of rational choice. It is not quite clear whether this kind of teleology covers what I called agency-consequentialism. If it did it would be a case of NON-SEQUITUR. On the other hand, if Ramsey-coherence would be an argument in favor of agency-consequentialism (the „teleological theory“), my argument above would be proven to be wrong. I have to show that it is not.

Let us pick out Broome's somewhat famous example. In order to "develop the general argument" of his book Broome rejects an

¹³ cf. A. Sen, Collective Choice and Social Welfare, London et al. 1970, ch.. 1 and 1*; id.. Social Choice Theory : A Re-Examination, Econometrica 45 (1977, 53-89); and L. Kern / J. Nida-Rümelin, Logik kollektiver Entscheidungen, München 1994, ch. 1.2-1.4.

apparent counter example of the transitivity of preference relations. The counter example is the following: Maurice prefers visiting Rome (R) to mountaineering in the Alps (M) and he prefers staying at home (H) to visiting Rome. However, he does not prefer staying at home to mountaineering; if he had a choice between this two alternatives, he would take the mountaineering trip. So Maurice's preferences are intransitive.¹⁵ Maurice's claim to rationality is that mountaineering frightens him, so he prefers visiting Rome. Sightseeing bores him, so he prefers staying at home. But to stay at home when he could have gone mountaineering would, he believes, be cowardly. This is the reason why he chooses mountaineering instead of the alternative to stay at home. Whereas if he chose to visit Rome instead of mountaineering this could not be taken as an expression of cowardice.

It seems that this is an example that options change with their alternatives. Staying at home is a different option depending on whether the alternative is visiting Rome or mountaineering . Let us call this phenomenon „choice-dependence of alternatives“. A most general way to deal with choice-dependence of alternatives would be to index every alternative with the given choice just as a matter of care. A closer look might then reveal (as a second step in the analysis) that many of these indexes can be dismissed. But in the marginal case of a general indexing of alternatives the comparability of alternatives would collapse completely. If it is the choice itself which makes the difference this tragedy would not save Ramsey-coherence (i.e. the compatibility assumption) but on the contrary would undermine its first constitutive principle and this is the connectivity axiom. We should not save transitivity by giving up another equally essential principle of preference coherence.

Indexing choice-dependent alternatives:

$H > R \wedge R > M \wedge M > H$ is an intransitive preference relation.

¹⁴ cf. P. Vallentyne, Gimmicky Representations of Moral Theories, Metaphilosophy 3&4 July/October 1988, 253-263

¹⁵ cf. J. Broome, Weighing Goods. Equality and Certainty , and Time, Oxford 1991, ch. 5.4

If H in the first choice situation is different from H in the third we can index H like "H_R" and "H_M" where H_R is "staying at home if the alternative is to visit Rome" and H_M is "staying at home if the alternative is to mountaineer in the Alps".

This allows for a new transitive preference relation

$H_R > R \wedge R > M \wedge M > H_M$. We passed from the 3-set of alternatives $\{H, R, M\}$ to the 4-set $\{H_R, R, M, H_M\}$

One interpretation of indexed alternatives would result in a radical contextualisation: by definition there would be no preference between H_M and H_R possible, because e.g. the index M makes sense only if the preference is between staying at home and mountaineering.

Broome's example is open to a consequentialist re-interpretation: The state of affairs "Maurice staying at home with the alternative to visit Rome" could be different from the state of affairs "Maurice staying at home with the alternative to mountaineer in the Alps". Maurice could for example feel better in the first in comparison to the second and in this way he could have the preference for H₁ compared to H₂. But in fact it is not at all sure that such a consequentialist solution exists. In case of the consequentialist solution the choice situation were only indirectly relevant for determining the respective alternatives. If mental-state-optimization was the metaphysical background of our theory of rational choice we should have to assume from the start that a choice-dependent preference relation can be rational only if there is a consequentialist (and more specifically mental-state-) re-interpretation which would render the apparent choice-dependance as merely an indirect effect of mental-state-optimization. But we have already seen that it is not reasonable to take this stance.

IV Weighing Reasons

Radical contextualisation undermines practical rationality. There must be something in common which gives reason to have

this preference now and that preference then. We refer to such common traits in telling reasons for action. If choice-sensitivity is rational there must be some reason to be choice-dependent. One reason could indeed be that the agent wants to optimize her mental state. This reason not only renders choice-dependence indirect (or not genuine) but allows also for a wide comparability of alternatives: If the agent expects to feel better if the first of two alternatives is realized she will opt for the first, independently of how the choice situation is. If the choice situation has an influence on how the agent feels this effect has to be included into a full description.

If a reason is responsible which is directly related to the choice situation as in the example above (choosing chairs) it offers another track of comparability beyond the context of the choice given. To act on the basis of reasons secures comparability. Weighing reasons in general secures universal comparability. The coherence requirement of completeness is *prima facie* plausible not only within the traditional consequentialist perspective of weighing goods (i.e. values of outcomes) but also in the broader perspective of weighing reasons.

The account of rationality which is to be defended here can be understood as a proposal of weighing reasons - instead of weighing goods. Reasons are the fundamental elements out of which a theory of practical rationality has to be constructed. But in order to choose coherently we have to weigh them and this includes probability considerations. Weighing reasons differs from traditional deontological accounts insofar as it postulates preference coherence (i.e. quantitative representability of rational epistemic and conative attitudes by real-valued belief- and value-functions) and it differs from consequentialist (or teleological) accounts in so far as it takes reasons into account which can not be reduced to optimizing state of affairs or outcomes.

Most of our *prima-facie* good reasons for action entail genuine choice-dependence. The outcomes alone, i.e. the resulting states of affairs do not determine whether the choice conforms

with the respective reason. Therefore the comparability of alternatives can only be saved if they are comprehensively described i.e. described such that their traits which are relevant for judging whether the choice conforms the reason are included. If alternatives of this kind are described adequately the reason in question is part of the description. If we compare alternatives and the reason in question is part of their description we implicitly have to weigh reasons.

V Structural Rationality

I conclude with a programmatic outlook on an account which I call „structural rationality“. It integrates non-consequential reasons for action and does not touch the core of rational choice theory.¹⁶ It starts from the assumption that point-wise outcome-optimization is (a) incompatible with many of our good reasons for action, (b) would desintegrate the structure of interaction on which social cooperation is based and (c) would desintegrate the person. Act-utilitarians argue that this assumption would not be true for perfectly rational utilitarian agents: The ideal utilitarian agents - they think - could adopt agency-consequentialism without destroying the structure of interaction on which social cooperation is based and without desintegrating their own personality. For act-utilitarianism the conflict with seemingly good reasons for action is merely an indication that our moral intuitions are no reliable. For the purpose of my argument it is not necessary to prove the truth of this claim.¹⁷ For even if the ideal moral person in the sense act-utilitarianism defines her could live on the basis of point-wise maximization, it would still be true that the rest of humankind could not. The rest of humankind differs in ranking states of affairs and weighing reasons, in personal projects and interests. A general theory

¹⁶ I used that term already in my habilitation-thesis which was handed in december 1988 (chap. 4, p. 417-479). A shortened version of this thesis was completed shortly after but published only four years later in 1993 because the book-series "Ethica" which de Gruyter planned and in which it should be the first publication was withdrawn. This is the reason why in *Kritik des Konsequentialismus* there is no reference to publications which appeared later than 1988 - e.g. Edward McClennen on resolute choice or John Broome on weighing goods.

¹⁷ I am convinced that act-utilitarianism is not right in this , see *Kritik des Konsequentialismus* §§ 26-29, 36

of practical rationality should not presuppose that the overall value functions representing revealed preferences are interpersonally identical.

A general theory of practical rationality should not be too restrictive. It should not prescribe which personal projects we try to realize in our lives, with whom we want to cooperate, which rules we want to follow, which reasons are the most relevant for our form of life. Some element of Humeanism is healthy for a general theory of rationality: what we are aiming at, how our conative attitudes are constituted is largely up to the acting person. If a person wants to be polite and therefore accepts a certain rule of politeness we should not call this person irrational only because the respective rule is not outcome oriented. We saw above how it is possible to integrate such a conative attitude into a general framework of rational choice. Radical Humeans would stop here and refuse to go further, because they think that there is no way to develop rational criteria for ends.

The theory of structural rationality has its starting point where radical Humeans stop. The idea is that - in analogy to descriptive beliefs which are neither isolated elements of an unstructured set - our actions reveal a whole system of subjectively accepted reasons which are closely interconnected and integrated. In order to conform with such a structured system of reasons we have to be careful in choosing our actions such that they conform with it as a whole. Point-wise outcome-optimization would result in a form of life which would not accord with the conative attitudes of most persons. It is perfectly rational to refrain from point-wise optimization because you do not wish to live the life which would result.

It is helpful to introduce the idea of structural rationality in a strictly consequentialist framework (i.e. the only relevant values would be related to outcomes or state of affairs which are caused by actions) emphasizing the tension between point-wise-optimization of consequences and overall optimization of consequences. The modification then would be to give up agency-consequentialism in order to save overall-

consequentialism. But this first step of the analysis opens the door for a more comprehensive understanding of practical rationality which leaves consequentialism behind.

Structural rationality covers the relation of singular actions and overall preferred (action-) structures in general. If you prefer a form of society in which politeness plays a role and if a specific rule is constitutive of your understanding of politeness this has an influence on which act is rational for you. Even if your personal rule-conformity or rule-defection has no influence on the respective social states and processes as a whole it is not necessarily irrational to act in conformity with the preferred structure of social action. The conative attitude referring to a social structure coheres with the respective conative attitude in the singular case.

The individuation of acts is modulo intention and the individuation of intentions is modulo reasons. In many cases an agent's intention can not be realized by a singular action. A whole sequence of actions can be necessary to realize the intention. This sequence as a whole can again be described as an action modulo this embracing intention. Depending on the agent's intentions some actions are done in doing other, more fine-grained actions. The minimal units are constituted by the intentional state of the agent in the moment the action is done. There is no agent-relative or even universal repertoire of atomic actions. Which actions are atomic depends completely on the respective choice situation and the intentional state of the agent in this situation. The more experienced the agent is the bigger are her atomic action units. The atomic actions are chosen within the broader framework of the agent's embracing intentions. The intention to do x_1 is motivated by the intention to do y because you have the intention to do y and know that doing $x_1, x_2, x_3, \dots, x_N$ would result in doing y . The relation between fine-grained and embracing intentions should be understood without falling back into a consequentialist interpretation. The relation is not one of causal determination of outcomes. Structural traits of rationality are constitutive parts of the intentional sets which guide our actions. An adequate theory of practical

rationality is not possible without taking them into consideration. As shown in this paper this can not be done within the conventional consequentialist interpretation of rational choice. But if this biased interpretation is given up structural traits are compatible with Ramsey-coherence, the core element of rational choice theory.