To appear in *Topoi*, special issue guest-edited by Simone Duca.

Odd choices

On the rationality of some alleged anomalies of decision and inference

Hans Rott

13 September 2010
with author's corrections 11 November 2010

Abstract. This paper presents a number of apparent anomalies in rational choice scenarios, and their translation into the logic of everyday reasoning. Three classes of examples that have been discussed in the context of probabilistic choice since the 1960s (by Debreu, Tversky and others) are analyzed in a non-probabilistic setting. It is shown how they can at the same time be regarded as logical problems that concern the drawing of defeasible inferences from a given information base. I argue that initial appearances notwithstanding, these cases should not be classed as instances of irrationality in choice or reasoning. One way of explaining away their apparent oddity is to view certain aspects of these examples as making particular options salient. The decision problems in point can then be solved by 'picking' these options, although they could not have been 'chosen' in a principled way, due to ties or incomparabilities with alternative options.
Keywords. rational choice; irrationality; inference; non-monotonic logic; context effects

1. Rational choice

The classical theory of rational choice is a paradigm for rational action. It is the preferences of a person that determine her actions. A rational person decides for the option that maximizes her preferences. The meaning of the preferences is not fixed. Usually they are not to be interpreted as just representing the egoistic interests of the decision-maker, but they may well take into account moral maxims and social goods. In the following, σ will denote a *choice function* (or *selection function*) that determines, for any set S of available options—for any *menu* or *issue S*—the subset of optimal, i.e. rationally choosable options. For simplicity, we assume throughout this paper that S may be any subset of a finite domain of potential options. For a person characterized by σ, each element of the set $\sigma(S)$ would be an optimal solution of the choice problem represented by S. It is important to note that a choice function σ is defined for many different menus. It is a matter of contingent fact (and irrelevant for the following discussion) which menu S the person is *in fact* facing. The function σ represents a set of *potential* choices or, to put it differently, the choice *dispositions* of the person. For a rational person, precisely the elements that are maximal with respect to $<$ are choosable, that is, $\sigma(S) = \{x \text{ in } S\text{: there is no } y \text{ in } S \text{ such that } x<y\}$.[1] We need to emphasize that the preference relation is not supposed to depend on the choice situation (which is represented by the menu S). In his Nobel prize lecture, Daniel McFadden (2001, p. 356) coined the slogan *Desirability precedes availability* for this idea. The choice dispositions of a person are called *rationalizable* if there

[1] I will take the asymmetric relation $<$ as primitive in this paper (i.e., $<$ is not taken to be derived from a primitive reflexive relation $\leq$).

is a fixed preference relation that allows us to explain the (potential) choices of the person as maximizing her preferences. *Rationalizability* is thus identified with *relationalizability*.

Intuitively one should think of the choice situation as comparatively transient and contingent: Which options S are open for a person in a given situation is (or, at least¸ frequently appears to be) a matter of chance. It is, however, a matter of a person's relatively persistent, stable character what she considers choosable given certain circumstances.[2] In this picture, the choosing and acting of a person are *determined* by her character and the given choice situation. A rational being *must* act in such a way that she maximizes (or: as if she maximized) her preferences. The theory of rational choice stipulates that she possesses at any point of time a decision-guiding preference relation across many different potential choice situations. Thus she is *programmed*, as it were, by her own preferences, here choices invariably *realize* or *reveal* her preferences. In this perspective, the nature of a human being becomes comparable to that of a chess computer that always chooses, and has to choose, the move that is being judged best by its internal evaluation algorithm.[3] One may be tempted to think that this conception contradicts the idea of human freedom. But I do not think it is adequate to speak of a restriction of freedom here. What needs to be said in this regard was expressed more than 300 years ago by the following words of John Locke:

> *To be determined by our own judgment, is no restraint to Liberty*. This is so far from being a restraint or diminution of *Freedom*, that it is the very improvement and benefit of it; ... 'tis as much *a perfection, that desire or the power of Preferring should be determined by Good*, as that the power of Acting should be determined by the *Will*, and the certainer such determination is, the greater is the perfection. Nay were we determined by any thing but the last result of our own Minds, judging of the good or evil of any action, we were not free, the very end of our Freedom being, that we might attain the good we chuse. And therefore every Man is put under a necessity by his constitution as an intelligent Being, to be determined in *willing* by his own Thought and Judgment, what is best for him to do: else he would be under the determination of some other than himself, which is want of Liberty.[4]

What guarantees freedom is that the judgments, thoughts or preferences at work are *the person's own* judgments, thoughts or preferences. Once the personal preferences are given, it does not make sense to complain that the person is the slave of these preferences. If there is anything to criticize, then it is the preferences themselves, but the process of forming personal preferences does not belong to the domain of rational choice theory.

It is initially surprising that the mere fact that a person's choice dispositions derive from an underlying preference relation determines certain structural properties of the choice dispositions. For instance, if an option x is optimal in S, that is, if x is contained in $\sigma(S)$, and the restriction of the menu S to a submenu S' does not exclude x, then x must be in $\sigma(S')$, too. If an option y is optimal in a menu S and likewise optimal in another menu S', then y must be optimal in the union set $S \cup S'$, too. The two statements are valid just because of the existence of a rationalizing preference relation, completely irrespective of the form and content of this

[2] I will neglect in the present paper the fact that in doxastic applications of rational choice theory, the "character" of a person encoded by a choice function or by rationalizing preferences is not so persistent and stable after all. It will typically change in response to the experiences the person makes, or more precisely, to the pieces of information ϕ_1, ϕ_2, ϕ_3, … that happen to come in. On the plurality of methods to change qualitative doxastic preferences, see Rott (2009).
[3] The likening to chess computers is kind of flattering, as many machines have become stronger than any human chess player now. Kant (1996, p. 218) famously talked about the "freedom of a turnspit [*Freiheit eines Bratenwenders*], which, when once it is wound up, also accomplishes its movements of itself".
[4] Locke (1975, section II.xxi.48, p. 264).

relation. How deep the connections are between the properties of choice functions over finite domains and the properties of rationalizing preferences, is summarized by the following well-known theorem.[5]

Theorem. (a) σ is rationalizable if and only if σ is rationalizable by the "revealed" preference relation $<$ defined by

$y < x$ iff x is in, but y is not in $\sigma(\{x,y\})$ — *base preferences*[6]

(b) σ is rationalizable if and only if it satisfies (I) and (II):

(I) If $S \subseteq S'$ then $S \cap \sigma(S') \subseteq \sigma(S)$ — *Sen's property α*

(II) $\sigma(S) \cap \sigma(S') \subseteq \sigma(S \cup S')$ — *Sen's property γ*

(c) σ is transitively rationalizable if and only if it satisfies (I), (II), and (III):

(III) If $S \subseteq S'$ and $\sigma(S') \subseteq S$ then $\sigma(S) \subseteq \sigma(S')$ — *Aizerman's property*

(d) σ is modularly[7] rationalizable if and only if is satisfies (I) and (IV):

(IV) If $S \subseteq S'$ and $\sigma(S') \cap S \neq \emptyset$ then $\sigma(S) \subseteq \sigma(S')$ — *Sen's property β+*

That a choice function is *transitively* or *modularly rationalizable* means that it is rationalizable by a transitive or modular preference relation. Correspondingly, (IV) is essentially stronger than (III).[8] The modularity of the relation $<$ together with its asymmetry implies its transitivity. Modularity intuitively means that all elements are comparable with respect to their desirability. In finite domains, this equivalent with the feasibility of assigning numerical values to the options which determine the preferences between them, in the sense that $x<y$ iff $val(x)<val(y)$.

This characterizes the basic ingredients of the classical theory of rational choice. The theory is philosophically plausible and formally of captivating elegance. But unfortunately it appears to have problems that were recognized early on in its development. A well-known example is due to Luce and Raiffa (1957). Compare two alternative scenarios of a visit to a restaurant. In the first scenario, only salmon (A) and steak (B) can be found on the restaurant's menu, and the customer decides to order salmon. Even though this customer normally prefers steak, he refrains from choosing it because the small menu makes him afraid that the cook might spoil the steak. In the second (alternative, not temporally successive) scenario there is a menu that offers, besides salmon and steak, also snails (C) and fried frog's legs (D). In this situation the customer decides to take the steak. Even though he likes neither snails nor frog's legs, the fact that they are listed on the menu indicates that this is a good restaurant that can be trusted to know how do to a steak (see fig. 1).

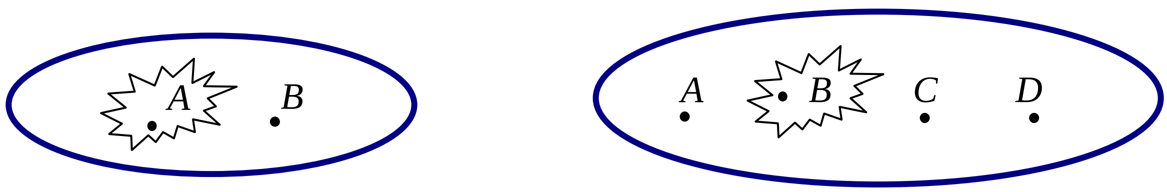


Fig. 1: The example of Luce and Raiffa

[5] Cf. Sen (1971) and Moulin (1985).

[6] Here it is of course still assumed that all two-element sets $\{x,y\}$ are in the domain of σ.

[7] A relation $<$ is *modular* if and only if $x<y$ implies that for every z, either $x<z$ or $z<y$.

[8] (IV) is stronger than (III) if the choice set $\sigma(S')$ is always required to be non-empty.

Intuitively, nothing in this scenario suggests that the customer is irrational. But his choice dispositions violate Sen's property α and Aizerman's property: Option *B* is chosen in the second scenario but not in the first scenario, even though it is on offer there as well. Option *A* is chosen in the first scenario, but not in the second scenario, even though all optimal solutions of the latter scenario are available in the first scenario. The reason for this odd choice behaviour can be seen in the fact that the menu in this example does not only list the available options, but has some *informational value* of its own: In the respective scenarios, the menus convey information (that need not be true) about whether the restaurant is mediocre or rather ambitious. And this seems to lead to a violation of the conditions of the classical theory of rational choice. But it is exactly the information conveyed by the menu that casts doubt on the adequacy of the representation of the cases in question. In the first scenario, the offer is steak-in-a-mediocre-restaurant, while in the second scenario, it is steak-in-an-ambitious-restaurant, and similarly for the salmon. In a sense, then, the menus are not comparable, even if some items appearing on the menu are literally identical. Seen in this light, there is no variation of any preferences in the two scenarios, just the objects being compared are different (steak-in-a-mediocre-restaurant is different from steak-in-an-ambitious-restaurant). The problem here seems to be one of under-representation.

2. Choosing and inferring

A related problem is discussed in Rott (2004) where the focus does not lie on the choosing of goods or actions but on the rational formation of beliefs. The example given there concerns the filling of a vacant position in metaphysics in a philosophy department. There is only a single job, and there are four finalists on the shortlist: Amanda Andrews (*A*) is an excellent metaphysician, Bernice Becker (*B*) a very good metaphysician with a substantial competence in logic, Carlos Cortez (*C*) a brilliant logician who also did some work in metaphysics and David Donaldson (*D*) who is generally perceived to be the candidate to win the competition. All this information is publicly available. But now we get confidential and surprising information by the dean of the faculty, a person who is known to be a very knowledgeable, serious and sincere person. Consider three alternative scenarios (again, not scenarios succeeding one another). In scenario 1 the dean tells us that the position will be given to either *A* or *B*. Given this information, we believe that the position will go to *A*, because the job was advertised for metaphysics and thus *A* appears to be better qualified than *B*. In scenario 2, the dean says that *C* will get the job, which makes us infer that *C* will get it. In scenario 3, the dean informs us that either *A* or *B* or *C* will get the job. This sets off a somewhat complex train of thought. We conclude from the fact that *C* is considered to be a serious candidate that logic plays a certain role in the committee's decision. Still we do not believe that *C* is the most suitable candidate. But now we may well reach the conclusion that *A*'s advantage with respect to metaphysics is not strong enough to compensate *B*'s expertise in logic (see Fig. 2).[9]

[9] Rott (2004) gives a numerical reconstruction that makes the example plausible for a comparatively wide variety of weights that may be given to logic (in relation to metaphysics). Critical discussions of this example have been provided by Hill (2008), Stalnaker (2009) and Arló-Costa and Pedersen (2010).

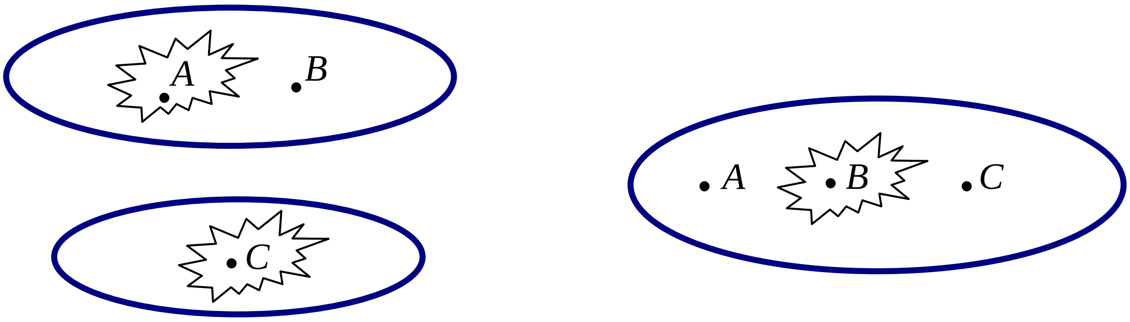


Fig. 2: The example of the appointment procedure

One can see by comparing scenario 1 with scenario 3 that Sen's property α and Aizerman's property are violated, similarly to the example of Luce and Raiffa. The diagnosis looks similar, too, at first sight at least. The fact that the logician Cortez is still in the race attests to the relevance of logic for the appointment, just as snails and frog's legs—though not first choices themselves—attest to the quality of the restaurant. But there is an important difference.[10] While Luce and Raiffa's example concerned a real choice to be made ("which meal should we order?"), we have nothing to decide in the story about the selection procedure. We are just being informed about some facts that contradict our expectations; like everyone else, we had been presuming that Donaldson would get the job. This is not a practical, but a theoretical or cognitive problem in which preferences don't express values or utilities, but degrees of plausibility. That we "prefer Andrews over Becker" in scenario 1 means, for instance, that we think it is *more plausible* or *more possible in our opinion* that Andrews gets the job than that Becker gets the job. The preference relation involved represents doxastic, not evaluative or volitive attitudes.

We can now ask whether it makes sense to apply the theory of rational choice to such kinds of preferences. This question was answered in the positive by various authors in the 1990s.[11] The theory of choice may in fact be applied very fruitfully, and it generates, depending on the strength of the constraints on rational choices, an assortment of logics that have been known and independently motivated in the literature on everyday reasoning. In what sense can we speak of a "logic" here? We have been talking about meals and applicants to make a choice from, not of propositions or sentences. But it is easy to convert the format of the examples. Let us abbreviate the sentences "*A* will get the job", "*B* will get the job" etc by "*a*", "*b*" etc. We will now depict the situation in such a way that the alternative pieces of information received by the dean support varying conclusions, and we call the drawing of such a conclusion an *inference*. The scenarios can be summarized as follows. In the initial situation in which we have not yet talked to the dean, we take it for granted that *D* will be appointed:

$$d \in Inf(\varnothing) \qquad \text{and} \qquad \neg a, \neg b, \neg c \in Inf(\varnothing)$$

Inf is a function that assigns, for every sentence ϕ of the language, the set $Inf(\phi)$ of consequences or sentences inferrable from ϕ. $Inf(\varnothing)$ denotes the set of sentences inferrable

[10] There are other important differences. The options involved are identical, since we are talking about the same persons across the different scenarios. So there must be some variation concerning the relevant choice functions or preferences.

[11] See Lindström (1991), Rott (1993, 1994, 2001), Schlechta (1997), Freund (1998), Lehmann (2001), and more recently Bonanno (2009) and Arló-Costa and Pedersen (2010).

from no information at all, or from a tautological piece of information.[12] In the alternative scenarios, we can make varying inferences on the basis of the new information received:

scenario 1	$a \in Inf(a \vee b)$	and	$\neg b, \neg c, \neg d \in Inf(a \vee b)$
scenario 2	$c \in Inf(c)$	and	$\neg a, \neg b, \neg d \in Inf(c)$
scenario 3	$b \in Inf(a \vee b \vee c)$	and	$\neg a, \neg c, \neg d \in Inf(a \vee b \vee c)$

It is natural to regard *Inf* as a logical operation. The striking point, however, is that a strengthening of the premise (in the sense of classical logic) does not lead to an augmentation of the conclusions. The premise $a \vee b$ of the first scenario is stronger, that is, more contentful, than the premise $a \vee b \vee c$ of the third scenario, but the first scenario does not vindicate b as a conclusion, which is exactly what we infer in the third scenario. In fact, in the first scenario we even infer the negation $\neg b$ of b. We can now fix a *first central statement concerning everyday inferences*:

> Everyday inferences are nonmonotonic, that is, if a sentence ϕ follows from another sentence ψ according to the canons of classical logic, this does not guarantee that the consequences of ϕ are included in the consequences of ψ.[13]

The news value of this statement is of course limited, since it has been known from discussions about induction and inductive logic for a long time. But its full impact was only recognized in the 1980s when nonmonotonic logics (also known as "logics for defeasible reasoning" or "default logics") developed into a research field of their own right.[14] One might be afraid that with the loss of the monotonicity property, all other logical patterns of inference would go away as well. But this is not true. This is the *second central statement concerning everyday inferences*:

> The loss of the monotonicity condition does not mean that anything goes in everyday reasoning; on the contrary, it has been argued that many substantial logical laws remain valid.

It is one of the fundamental insights of the research on nonmonotonic logic that commonsense inferences exhibit considerable regularities even in the absence of monotonicity. The following logical properties do not only serve as examples for this claim, but will also be of special interest for the later parts of this paper. They have been regarded by many researchers as particularly plausible and highly desirably conditions for everyday inferences.

(Cumulative Monotony)	If $\psi \in Inf(\phi)$, then $Inf(\phi) \subseteq Inf(\phi \wedge \psi)$
(Or)	$Inf(\phi) \cap Inf(\psi) \subseteq Inf(\phi \vee \psi)$
(Conditionalization)	$Inf(\phi \wedge \psi) \subseteq Cn(\, Inf(\phi) \cup \{\psi\}\,)$

[12] *Inf* can also take finite sets of sentences, if they are conjoined by the conjunction $\wedge$. An extension to infinite sets of sentences is non-trivial.

[13] As long as we stick to finite sets of premises and identify them with their conjunctions, this implies that a subset X of a premise set Y may have consequences that Y itself does not have. Concerning the use of variables in this paper, a, b, c etc stand for particular sentences occurring in our examples, while ϕ and ψ are schematic for arbitrary sentences.

14 On induction see Levi (2005) and Spohn (2005). On nonmonotonic logics see Ginsberg (ed., 1987) and Makinson (2005).

In the formulation of these conditions we continue to use *Inf* for the notion of everyday inference and distinguish from it the monotonic, essentially classical background logic *Cn*.[15] In this notation, the condition of Nonmonotonicity reads thus: $\phi \in Cn(\psi)$ does not imply $Inf(\phi) \subseteq Inf(\psi)$.

If we consider our example of the appointment procedure as an exercise in everyday reasoning, we see that it violates all three conditions just mentioned. First, we find, against Or, that $Inf(a \vee b) \cap Inf(c)$ fails to be a subset of $Inf(a \vee b \vee c)$; second, against Conditionalization, $Inf(a \vee b)$ is no subset of $Cn(Inf(a \vee b \vee c) \cup \{a \vee b \vee \neg c\})$. The sentences $a \vee c$ and $\neg b$ may serve as witnesses for these claims. Third, $a \vee b$ is contained in $Inf(a \vee b \vee c)$, but $Inf(a \vee b \vee c)$ is no subset of $Inf(a \vee b)$, contradicting Cumulative Monotony. This is witnessed, for instance, by the sentences b and $\neg a$.[16]

It seems justified to say that these logical properties are violated *just because* the underlying principles of rational choice are violated. The breach of Or and Conditionalization is due to the fact that Sen's condition α is violated: B is a best option in $\{A,B,C\}$ and of course it is an option in $\{A,B\}$, but B not a *best* option in $\{A,B\}$. The breach of Cumulative Monotony can be attributed to a corresponding violation of Aizerman's property: All best options of $\{A,B,C\}$ lie within $\{A,B\}$, and A is a best option of $\{A,B\}$, yet A at the same time fails to be a best option of $\{A,B,C\}$. Seemingly odd choices are the reason for seemingly odd inferences.

How can the idea that logical conditions derive from conditions of rational choice be substantiated? One can think of choices as applying to two different levels. On the semantic level, the plausible inferences used in everyday reasoning consist in the choice of the *best possible worlds*. If a person has the information ϕ as her premise,[17] then she looks for the most plausible, that is, most belief-conforming, possible worlds at which ϕ is true. On the syntactic level, the choices concern the *worst beliefs*: If a person has the information ϕ as her premise, then she looks for the least plausible, that is, least well-founded sentences that had helped to derive the initial belief or expectation that $\neg\phi$. In both readings, plausibility is assessed by the standards of the person's doxastic choice function. A sentence ψ can be inferred from a premise ϕ on the semantic level if and only if ψ is true at all most plausible ϕ-worlds; or, on the syntactic level, if $\phi \to \psi$ does not belong to the set of most implausible *Cn*-consequences of $\neg\phi$.[18] The logical conditions Or and Conditionalization are equivalent, given some other, more basic conditions. Applying the theory of rational choice, one can show that Or and Conditionalization correspond exactly to Sen's property α, while Cumulative Monotony corresponds to Aizerman's property. There is a list of other correspondences that suggest that logical conditions for nonmonotonic inferences can be interpreted as constraints

[15] Compare the first chapter of Rott (2001).

[16] Note that $(a \vee b \vee c) \wedge (a \vee b)$ is *Cn*-equivalent with $a \vee b$, and that it should therefore allow the same inferences as $a \vee b$.

[17] Here and in the following, this formulation is supposed to mean that the premise ϕ contains *all the information* available to the person. This point is crucial to the understanding of nonmonotonic reasoning.

[18] Unfortunately, this latter condition is rather hard to understand. It can be replaced by the condition that $\phi \to \psi$ is not among the most implausible elements of $\{\phi \to \psi, \phi \to \neg\psi\}$. For the motivation of both conditions see Rott (2001), pp. 172-181. There are bridge principles showing that the concept of a "best possible world" has the same theoretical power as the concept of a "weakest belief". The following bridge principles are suitable to bear this out: A best ϕ-world is one at which ϕ and all except the least plausible *Cn*-consequences of $\neg\phi$ are true. Conversely, a sentence χ in a set of sentences M is least plausible in M if and only if among the best M-falsifying worlds (i.e., the best possible worlds at which at least one element of M is false) there is a world at which χ is false (cf. Rott 2001, pp. 208-213). Here I am not presuming that a choice function needs to be rationalizable by a preference relation.

for rational choice. We don't need to present a complete list here, but the following pairings will turn out to be relevant. The choice-theoretic condition (II) and its strengthening

(II^+) If x is in $\sigma(S)$ and y is in $\sigma(S')$, then either x or y is in $\sigma(S \cup S')$

correspond, in this order, to the logical conditions

(Very Weak Disjunctive Rationality) $Inf(\phi \vee \psi) \subseteq Cn\,(Inf(\phi) \cup Inf(\psi))$

and

(Disjunctive Rationality $Inf(\phi \vee \psi) \subseteq Inf(\phi) \cup Inf(\psi)$

Almost all the choice-theoretic constraints have exactly the same effect when applied to the semantic level (plausible worlds) as when applied to the syntactic level (implausible sentences). Sen's condition γ is an (almost singular) exception. Very Weak Disjunctive Rationality is the effect of (II) applied to the semantic level, while placing (II) on the syntactic level yields a weak form, and not only the very weak form of Disjunctive Rationality.[19]

I advocate the following thesis: *All problems of rational choice can be represented, without much difficulty, as problems concerning the drawing of everyday inferences.* In the following I will not elaborate this thesis abstractly. Instead I will carry on in the next section with the example of the metaphysics appointment, and vary it in three different ways. The practical problem of choosing people for a job from different sets of candidates will be transferred into a problem of drawing inferences on different bases of information just as effortlessly as in the last example. This, I hope, will suffice as evidence for the validity of my thesis.

At the end of this section we should comment on the "anomaly" exhibited by the above example. Are we facing a case of plain irrationality? No, the story is just too plausible, and the reasoning is too easy to comprehend. But the explanation of the Luce-Raiffa example does not seem to be suitable here. It may have been legitimate there to object that the content of the menus smuggled in extra information about the restaurant. In the example of the appointment procedure, however, the case is different in that here we are dealing with a problem concerning precisely the formation and transformation of beliefs. So it seems odd to complain that information is "smuggled in" by the premises. Any kind of consideration that we have informally described should be representable explicitly in the belief change model, given the specific background of the beliefs and assumptions of the recipient of the information. The solution must lie elsewhere. The point is, as already sketched in Rott (2004), that the hearer in the third scenario does not receive the information that $a \vee b \vee c$, but the information "The dean of the faculty tells me that $a \vee b \vee c$". And *this* is what triggers the comparatively complex train of thought that leads to the conclusion that b. Had the information that $a \vee b \vee c$ been obtained, for instance, by the observation that all the other candidates have left the place with a long face, then we would not have been able to conclude that logic matters and that B might end up being the best candidate for this appointment. But if we annotate the new information explicitly with the source of information, then the example loses any appearance of irrationality. From "The dean of the faculty tells me that $a \vee b$ " it simply does not follow that "The dean of the faculty tells me that $a \vee b \vee c$". So this example has something in common

[19] The condition of Weak Disjunctive Rationality is unfortunately not very intuitive. It says that $Inf(\phi \vee \psi) \subseteq Cn(Inf(\phi) \cup \{\psi\}) \cup Cn(Inf(\psi) \cup \{\phi\})$.

with the Luce-Raiffa case after all: The initial representation of the example may be too simplistic.

3. New cases

In the last section we have described a close correspondence between requirements for rational choices and requirements for nonmonotonic everyday inferences. But we also looked at two strange examples that gave us reason to doubt even the most natural and paradigmatic of all constraints for rational choices, namely Sen's property α. In response to Luce and Raiffa's example, we sketched the defence that the options were poorly represented by the first modelling, and that, for instance, the option "steak" should rather have been designated as "steak-in-a-good-restaurant". The appointment procedure example was presented as a logic problem right from the start, but we saw that interpreted as a problem of choice, it showed some similarity with the Luce and Raiffa case.

We will now consider three new types of cases. They have been discussed for a long time in the context of so-called *probabilistic choice*, which examines the frequencies of certain choices of the persons in a reference group.[20] We will discuss only very simplified, purely qualitative versions of these cases. In contrast to the examples presented so far, Sen's central property α is not called into question. But other important conditions like Sen's property γ and Aizerman's property are violated. Our question will be whether these are cases of irrational choice or not. According to the classical theory of rational choice we should call them irrational, but the examples might raise the converse suspicion that the classical theory itself is deficient (for reasons that may still be unclear).

We will focus (i) on the effect of *similar options*,[21] (ii) on the effect of *compromise options*,[22] as well as (iii) on the effect of *dominating options*.[23] In all three types of examples, we have that both A and B can be rationally chosen in the menus consisting only of the two options (see Fig. 3).

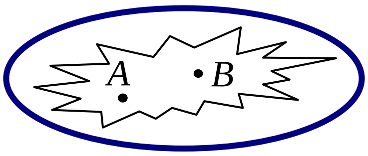


Fig. 3: The new cases—menu with two elements

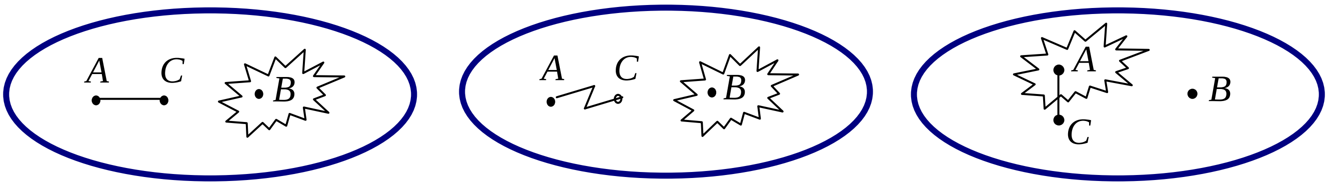
[20] We can leave it open here whether this research is meant to imply that every individual has a specifiable propensity to choose a certain option in a given single situation.

[21] Debreu (1960). This three-page review of Luce (1959) was seminal for the theme of the present paper. The bone of contention, *Luce's Choice Axiom*, is worth reproducing once again, in a variant that is sometimes called the *Constant Ratio Rule*: If A and B are elements of a menu S and $P_S(A)$ denotes the probability that the element A is chosen from the menu S, then $P_S(A)/P_S(B) = P_{\{A,B\}}(A)/P_{\{A,B\}}(B)$. Luce's Axiom formulates a sort of "independence of irrelevant alternatives" (cf. Luce 2008). The qualitative analogue is this: If $S \subseteq S'$ and $\sigma(S') \cap S \neq \emptyset$, then $\sigma(S) = \sigma(S') \cap S$. This condition is known as *Arrow's Axiom*, and it is equivalent with the conjunction of (I) and (IV) (cf. Rott 2001, p. 154). – Debreu uses an example of a menu with two similar recordings of the eighth symphony of Beethoven and one recording of a string quartet of Debussy. Important papers for the history of the problem of similar options were Chipman (1960), Tversky (1972) and McFadden (1974). In the literature the problem is widely discussed by means of red and blue busses as vehicles that are not different in any significant respect (the *red bus/blue bus problem*).

[22] Simonson (1989).

[23] Huber, Payne and Puto (1982), Simonson and Tversky (1992), Shafir, Simonson and Tversky (1993), and Shafir and Tversky (1995). The effect of asymmetrically dominated options is also called *attraction effect* and *decoy effect* in the literature.

Each of the three cases presents its characteristic problematic when the menu is enlarged by a third option *C* (see Fig. 4).



(i) similarity *(ii) compromise* *(iii) dominance*

Fig. 4: The new cases—menus with three elements

The choices within the three-element menus are made on the basis of certain relations between *A* and *C*, while *B* is incomparable with both *A* and *C* and occupies as it were an isolated position. As in the previous example, the choices within the three-element competitions are contrasted with the choices in two-element competitions. Except for case (iii) in which *A* is preferred to *C*, both options are eligible in any two-element menu, either because they are equally good (*A* and *C* in case (i)) or because they are incomparable (all other pairwise competitions). Considering the fact that *A* (in case (iii)) or, respectively, *B* (in cases (i) and (ii)) are clearly preferred in the three-element menus, this is forbidden by the classical theory of rational choice. The single prevailing option in the larger competition should still prevail alone if some other (unsuccessful) candidate is cancelled from the menu. It seems odd that this idea is not heeded here.

Let us for the sake of illustration continue the above example of the appointment procedure, but let us make a few changes to the field of applicants. Suppose that Cortez and Donaldson are not in the race, but that Amanda's sister Caroline has applied as well.[24] Let *C* be the short name for Caroline. For the similarity case, we can imagine that Amanda and Caroline are twins who are hard to tell apart, even regarding their competences in metaphysics and logic. In the compromise case, however, the sisters are just opposites. Whereas Amanada is an excellent metaphysician who does not know much about logic, this is just the other way round for Caroline.[25] Amanda and Caroline are well (in fact, extremely well) comparable in case (i). But they can be regarded as incomparable in case (ii), at least so long as there is no standard that puts competences in metaphysics into relation to competences in logic: Amanda is very good in metaphysics and has no expertise in logic, Caroline is just the reverse. In case (iii) we have a good comparability again, but this time asymmetrically: Amanda is somewhat better suited for the job than Caroline in all relevant respects. The following constellation is common to the three cases: Amanda and Bernice are incomparable, but as soon as Caroline enters the picture, her presence emphasizes certain relevant features of Amanda that motivate a decision between Amanda and Bernice after all. Even though Caroline herself is not capable of encroaching upon the decision, she acts as a catalyst to the decision-making process.

In the *similarity* case (i), there are three options, two of which (*A* and *C*) having very similar properties and thus being evaluated as equally good. *B* is different in kind from *A* and *C*,

[24] Being relatives does not exclude being competitors.

[25] We pretend for the purposes of this example that good metaphysics is possible without logic, and good logic without metaphysics. It is crucial for the examples adduced in this paper that there are different factors or criteria that are, and are thought of as, independent of each other.

incomparable with these in its properties. As a result, a two-element competition between *A* and *B* has no definite winner, but in a three-element competition many people choose option *B*. The reason for this is evidently that the test persons have no clues whatsoever to distinguish between *A* and *C*, and so *B* is salient as the only option with a recognizable "unique selling proposition".

The *compromise* case (ii) is similar in structure to, but quite different in substance from, case (i). Here *A* and *C* are not similar in their property profile, but just very different and mark, as it were, two extreme options within a certain decision spectrum. As already mentioned, a typical case is one in which there are two relevant properties ("dimensions") and *A* satisfies the first, but not the second very well, while *C* conversely satisfies the second, but not the first very well. *B* lies in both respects "between" the extremes *A* and *C*. In their own ways, all candidates are suitable for choice to the same (or better: to an incomparable) degree. But *B* possesses a unique selling point due to its being a compromise of the two extremes, and this is presumably why it is frequently chosen.[26]

Matters are different in the *dominance* case (iii). *A* and *C* are very easily comparable here, and the result of the comparison is that *A* is clearly preferable to *C*. Like in case (i), *B* is hard to compare with both *A* and *C*. But since the dominance of *A* over *C* provides a "reason" to prefer *A*, option *A* is chosen more frequently in empirical situations that offer a three-element choice between *A*, *B* and *C*. If there is no contrast with *C*, however, *A* is not distinguished against *B* and there is no significant propensity to choose *A*.

Which coherence constraints are violated in these cases? First, it is Sen's property γ, that we also called (II) above. In the cases of similarity and compromise, *A* is optimal in the pairwise comparison with *C* as well as in the pairwise comparison with *B*, but in the three-element competition it is only *B* that gets chosen. Option *A*'s eligibility is "lost" when the union of the smaller menus is taken. In the dominance case, the same happens to *B*. Second, Aizerman's property (III) gets violated. Even though all optimal elements in the three-element menu (*B* in (i) and (ii), *A* in (iii)) are contained in the set $\{A,B\}$, it is not true that all optimal elements of the restricted set—remember that $\sigma(\{A,B\})=\{A,B\}$—remain optimal within the three-element menu.

The violation of the choice properties (II) and (III) has the following effects on the logical level. We focus on the cases (i) of similarity and (ii) of compromise; the dominance case (iii) is analogous, with *a* and *b* changing roles. First, the condition of Very Weak Disjunctive Rationality does not hold. While we can defeasibly infer from the disjunction $a\vee b\vee c$ that $\neg a$, we cannot draw any substantial conclusions from either $a\vee b$ or $a\vee c$, so $Cn(Inf(a\vee b)\cup Inf(a\vee c))$ does not permit us to infer $\neg a$.[27] Second, the condition of Cumulative Monotony gets violated. Even though $a\vee b$ may be inferred from $a\vee b\vee c$, not all consequences of $a\vee b\vee c$ are at the same time consequences of $a\vee b$.[28] Sentence *b* can be inferred from the weaker $a\vee b\vee c$, but not from the stronger $a\vee b$.

[26] The original appointment example of Rott (1994) may also be described as a compromise case. Here, however, the effect of the presence of the third candidate Cortez in the race even reverses the choice between *A* and *B* – at least in the opinion of the person informed by the dean.
[27] The situation is actually worse. Still supposing that only a single person can be employed, the union of $Inf(a\vee b)$ and $Inf(a\vee c)$ even has *a* among its *Cn*-consequences.
[28] Notice that $(a\vee b\vee c)\wedge(a\vee b)$ is *Cn*-equivalent with $a\vee b$ and should therefore license the same inferences as the latter.

The ease and effortlessness with which the problems of choice can be mirrored into the realm of everyday reasoning opens up a new research program. The task is to study systematically whether the intense discussions and varied proposals how to interpret and resolve problems of choice can be of help in the evaluation of the virtues and vices of existing systems of everyday reasoning.

4. Discussion

The types of cases considered are plausible constellations, and it has been shown that they have empirical reality in many situations across diverse populations. Empirical studies established probabilistic connections telling us how frequently persons in the relevant scenarios decide for which of the options. The transfer from probabilistic choice to the normative area of individually rational choice is not trivial, but I suggest that it be regarded as feasible. A few points need to be observed, though, that separate empirical studies from the classical theory of rational choice. First, in an empirical situation one can make one and only one decision, while the normative theory of choice can designate several options as rational and thus admissible for choice. Second, in an empirical situation only a single set of options, i.e., a single menu can be offered to be chosen from (a second offer could only be made in another, subsequent situation), while the theory of rational choice characterizes precisely the structural connections between different potential menus. In empirical situations one tries to solve the problem by performing repeated experiments, assuming that individual preferences are stable across a series of choice situations, but this can only be an approximation of the real facts of the matter. Third, the particular contents of the menu will always tend to convey information about the empirical situation (as in the example of Luce and Raiffa), a phenomenon that the theory of rational choice attempts to filter out. Finally, it is difficult to decide on the basis of an empirical finding whether an option is rational (admissible for choice). Clearly, it is not sufficient if a few test persons out of a large sample choose a particular option. These persons could well be irrational outliers. But what would be the threshold percentage that allows us to say that the approval within a population is so great that one cannot help but concede a certain rationality to choosing this option? Or is the gap between descriptive findings and normative recommendation so deep that it cannot be bridged at all? Could it happen that a choice principle fails to be rational even if all the members of a population obey it? We screen off all these problems and remain at the level of thought experiments that appeal directly to the reader's rationality intuitions.

It has perhaps become clear how important the concept of *incomparability* is. Incomparability has been rarely made a topic of research,[29] but it is crucial to our interpretation of these cases. In all of them, we have been confronted with options that can be evaluated by more than one criterion or, as I prefer to say, in more than one dimension.[30] As far as I can see, it is sufficient to take into account two dimensions. According to a single criterion, or along a single dimension, all options are perfectly *comparable*. Either A is better than B, or B is better than A, or A and B are equally good. Matters are entirely different as soon as two criteria are to be

[29] A notable exception is Ruth Chang (1997) who would probably subsume our type of incomparability under the heading "diversity of values" or "bidirectionality" – and deny that such phenomena give rise to genuine incomparability.
[30] The talk about *dimensions* is supposed to suggest that the various aspects of evaluation are independent of each other. The recognition of multiple criteria or dimensions opens up completely new questions. The very wide field of *social* choice theory gets immediately relevant if the preferences or evaluations according to various criteria involved in a decision are treated like the preferences or evaluations of different members in a social community. The impossibility theorems of Arrow and others have to be kept in mind. Cf. Arrow, Sen and Suzumura (2002) and Figueira, Greco and Ehrgott (2005).

observed. If both criteria speak in favour of the same option or if one of the criteria sees the options *A* and *B* as equally good, then there is no problem, because it is clear which option is to be preferred overall. The interesting case, and the one that is relevant in all our examples, is the one in which the two criteria pull in different directions for at least two options: *A* does better according to criterion 1 (metaphysics) and *B* does better according to criterion 2 (logic). So we have a case of incomparability, at least unless further considerations come into play that could decide the case. A ranking of criteria (in the purely qualitative modelling) or a weighting between them (in a quantitative modelling) might resolve the conflict, but otherwise there cannot be a meaningful comparison. Based solely on the choice behaviour, one cannot decide whether two options are incomparable or tied. In either case both options may be rationally chosen.[31]

The three classes of phenomena discussed can be subsumed under the common heading of *context effects*.[32] The third option *C* introduces a context to the choice between *A* and *B* (and it is this choice we are actually interested in). This context has an influence on the relative appreciation of *A* and *B*. As we have seen, this influence depends heavily on the relative "location" of the new option. Fig. 5 gives a qualitative sketch of the areas for *C* that introduce a choice bias in favour of either *A* or *B*. Notice that sometimes only pragmatic considerations may decide whether a particular case of slight dominance is effective as one of similarity or rather one of dominance—with opposite tendencies, as we have seen.

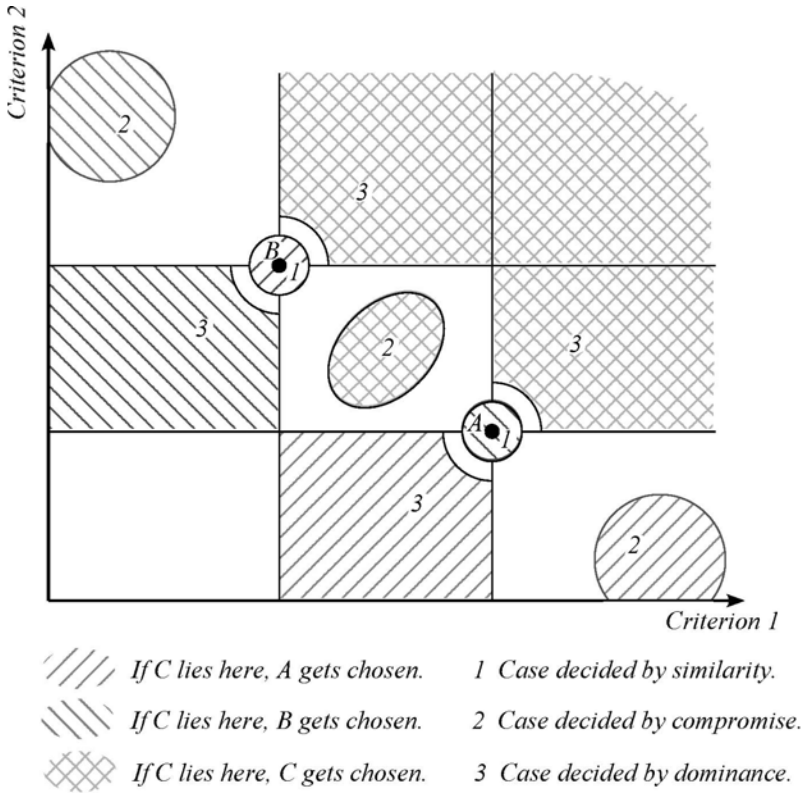


Fig. 5: Contexts for the effects of similarity, compromise and dominance

[31] Giacomo Bonanno (personal communication) has suggested that context effects may exist even in cases when *A* and *B* are tied rather than incomparable. Suppose for instance that *A* and *B* are applicants who are exactly alike in all *relevant* respects, but that *A* is female and *B* is male. If a third candidate *C* enters the competition, who is inferior to both *A* and *B*, his or her sex (a property that everyone agrees should be *irrelevant* for the appointment) may still tip the scales in favour of *A* and *B*. This is a very interesting hypothesis which, if true, I would not know how to explain. I do not know whether it is empirically valid.
[32] Cf. Simonson and Tversky (1992) and Tversky and Simonson (1993).

Odd choices are not necessarily symptoms of inconsistency of fickleness. They have to be taken seriously and should not rashly be discredited as irrational. Possible answers to the examples of Luce and Raiffa (1957) and Rott (2004) were sketched in the first two sections above. It seems to me that the problems presented by the new cases can be resolved in a way that explains away their seeming irrationality. Our above findings are in fact logically compatible with the assumption of rational choices on the basis of fixed, context-independent values of the options. Neither the empirical studies nor our intuitions in thought experiments have to be interpreted in such a way that options that are equally eligible in a pairwise comparison suddenly get evaluated differently in a three-element competition. An alternative interpretation is possible that distinguishes carefully between genuine choosing and mere picking.[33] *Choosing* refers to the rational determination of an option that is guided by the person's preferences, values, desires etc. *Picking*, on the other hand, is the deciding for an option without any such guidance. That does not necessarily mean that picking is always done completely at random. Decisions that are not guided by preferences, values, desires etc. may still be influenced or caused by other factors, factors that are not relevant for any assessments of the rationality of the choice. In our three types of examples—similarity, compromise and dominance—, the distinguishing property of the chosen option may be regarded as a property that guarantees *some* sort of salience within the three-element menu. But neither those properties themselves ("being dissimilar to other options that are similar to each other", "being 'between' other, antipodal options", "dominating some other option") nor the salience that an option may enjoy due to its having such properties, are value-increasing or value-decreasing. The distinctive features just introduce some differences in prominence between options amongst which no rational decision, i.e., no decision justifiable in terms of preferences, values, desires etc., can be made. In this way they overcome incomparabilities, unlock an incipient choice blockade, and prevent imminent torpidity. So these properties are obviously relevant for the process of decision-making. Because such an unlocking is procedurally rational, one may perhaps raise the question whether these features are modulating the formation of our preferences between the options after all. But as they do not attach to the options themselves, but are rather *context-dependent* or *relational* properties, it is unclear whether they can be value-conferring in any sense.

Much more elaborate explanations of the choice "anomalies" we discussed have been given in the psychological literature.[34] Psychologists want to explain and predict human behaviour, it is not their business to judge it in terms of rationality. Economists are closer to the philosophers' concern, since they aim at saying which decisions ought to be taken in certain situations. The most popular view here still appears to be that deviations from classical choice models are irrational.[35] But it seems to me that the problematic has some genuinely philosophical, metaphysical content. The question is: Does an option (e.g., a bundle of goods, the result of an action, or a possible world) possess some value in and of itself? Or could it be that any value that an option possesses is dependent on the context of the available alternatives? (Here, as always, "value" should not be understood just in a selfishly individual sense, but may also include social or otherwise "objective" aspects.) To measure the goodness of an option by assigning a numerical value to it, for instance a certain utility or amount of money, presupposes an absolute notion of value, and exactly this absoluteness is called into

[33] See Ullmann-Margalit and Morgenbesser (1977).
[34] See the excellent survey of Rieskamp, Busemeyer and Mellers (2006). It was only after the present paper was essentially finished that I became aware that a few psychological papers of the last 10 years juxtapose examples illustrating the similarity, compromise and dominance effects in a way that is similar to the presentation here.
[35] See, for instance, the first chapter, "The Truth about Relativity", of Ariely (2008). I thank Jim Delgrande for pointing me to this book.

question by the examples that we have discussed. We have seriously to face the possibility that the absolute value of an option is just a piece of fiction that at best represents an abstraction of the truly fundamental values belonging to the option in different contexts of available alternatives. Unfortunately, this is too big a question to be even addressed in the present contribution.

References

Ariely, Dan: 2008, *Predictably Irrational. The Hidden Forces that Shape Our Decisions*, New York: Harper-Collins.

Arló-Costa, Horacio, and Arthur P. Pedersen: 2010, 'Social Norms, Rational Choice and Belief Change', in Erik J. Olsson and Sebastian Enqvist (eds.), *Belief Revision Meets Philosophy of Science*, Dordrecht: Springer, pp. 163-212.

Arrow, Kenneth, Amartya K. Sen and Kotaro Suzumura (eds.): 2002, *Handbook of Social Choice and Welfare*, Band 1, Amsterdam: Elsevier.

Bonanno, Giacomo: 2009, 'Rational choice and AGM belief revision', *Artificial Intelligence* 173, 1194-1203.

Chang, Ruth: 1997, 'Introduction', in Ruth Chang (ed.), *Incommensurability, Incomparability and Practical Reason*, Cambridge: Harvard University Press, pp. 1-34.

Chipman, John: 1960, 'The Foundations of Utility, *Econometrica* 28, 193-224.

Debreu, Gerard: 1960, 'Review of R.D. Luce’s Individual Choice Behavior', *American Economic Review* 50, 186-188.

Figueira, José, Salvatore Greco and Matthias Ehrgott (eds.): 2005, *Multiple Criteria Decision Analysis: State of the Art Surveys*, New York: Springer.

Freund, Michael: 1998, 'Preferential orders and plausibility measures. Journal of Logic and Computation 8, 147-158.

Ginsberg, Matthew L. (ed.): 1987, *Readings in Nonmonotonic Reasoning*, Los Altos CA: Morgan Kaufmann.

Hill, Brian: 2008, 'Towards a "Sophisticated" Model of Belief Dynamics. Part II: Belief Revision', *Studia Logica* 89, 291-323.

Huber, Joel, John W. Payne and Christopher Puto: 1982, 'Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis', *Journal of Consumer Research* 9, 90-98.

Kant, Immanuel: 1996, *Critique of Practical Reason*, reprinted in: *Practical Philosophy*, *The Cambridge Edition of the Works of Immanuel Kant*, ed. by Mary J. Gregor, Cambridge: Cambridge University Press 133-271 (orig. 1788).

Lehmann, Daniel: 2001, 'Nonmonotonic Logics and Semantics', *Journal of Logic and Computation* 11, 229-256.

Levi, Isaac: 2005, 'Inductive Inference as Ampliative and Non-monotonic Reasoning', in: *Proceedings of the 10th Conference on Theoretical Aspects Of Rationality And Knowledge*, ed. Ron van der Meyden, National University of Singapore, ACM Digital Library 2005, pp. 177-192.

Lindström, Sten: 1991, 'A Semantic Approach to Nonmonotonic Reasoning: Inference Operations and Choice', manuscript, Department of Philosophy, University of Uppsala. (Slightly updated version in *Uppsala Prints and Preprints in Philosophy*, 1994, no 10.)

Locke, John: 1975, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch, Oxford: Oxford University Press (orig. 1690).

Luce, R. Duncan: 1959, *Individual Choice Behavior: A Theoretical Analysis*. New York: John Wiley and Sons.

Luce, R. Duncan: 2008, 'Luce's choice axiom', *Scholarpedia* 3(12):8077, http://www.scholarpedia.org/article/Luce's_choice_axiom

Luce, R. Duncan, and Howard Raiffa: 1957, *Games and Decisions*, New York: Wiley.

Makinson, David: 2005, *Bridges from Classical to Nonmonotonic Logic*, London: King's College Publications.
McFadden, Daniel: 1974, 'Conditional Logit Analysis of Qualitative Choice Behavior', in Paul Zarembka (ed.), *Frontiers in Econometrics*, New York: Academic Press, pp. 105-142.
McFadden, Daniel: 2001, 'Economic Choices,' *American Economic Review* 91, pp. 351-378. (Also in Torsten Persson (ed.), *Nobel Lectures in Economic Sciences 1996-2000*, Singapore: World Scientific 2003, pp. 330-365)
Moulin, Hervé: 1985, 'Choice Functions Over a Finite Set: A Summary', Social Choice and Welfare 2, 147-160.
Rieskamp, Jörg, Jerome R. Busemeyer and Barbara A. Mellers: 2006, 'Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice', *Journal of Economic Literature* 44, 631-661.
Rott, Hans: 1993, 'Belief Contraction in the Context of the General Theory of Rational Choice', *Journal of Symbolic Logic* 58, 1426-1450.
Rott, Hans: 1994, 'Coherent Choice and Epistemic Entrenchment (Preliminary Report)', in Bernhard Nebel and Leonie Dreschler-Fischer (eds.), *KI-94: Advances in Artificial Intelligence*, Berlin: Springer, pp. 284-295.
Rott, Hans: 2001, *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*, Oxford Logic Guides, Bd. 42, Oxford University Press, Oxford.
Rott, Hans: 2004, 'A Counterexample to Six Fundamental Principles of Belief Formation', *Synthese* 139, 225-240.
Rott, Hans: 2009, 'Shifting Priorities: Simple Representations for Twenty-seven Iterated Theory Change Operators', in *Towards Mathematical Philosophy*, eds. David Makinson, Jacek Malinowski and Heinrich Wansing, Dordrecht: Springer, pp. 269-296.
Schlechta, Karl: 1997, *Nonmonotonic logics: Basic Concepts, Results, and Techniques*, Berlin: Springer.
Sen, Amartya K.: 1971, 'Choice Functions and Revealed Preference', *Review of Economic Studies* 38, 307-317, reprinted in A.K.S., *Choice, Welfare and Measurement*, Blackwell, Oxford 1982, 41-53.
Shafir, Eldar, Itamar Simonson and Amos Tversky: 1993, 'Reason-Based Choice', *Cognition* 49, 11-36.
Shafir, Eldar, and Amos Tversky: 1995, 'Decision Making', in Daniel N. Osherson and Eldar E. Smith (eds.), *Invitation to Cognitive Science: Thinking*, Cambridge, Mass., MIT Press, pp. 77-109.
Simonson, Itamar: 1989, 'Choice Based on Reasons: The Case of Attraction and Compromise Effects', *Journal of Consumer Research* (16), 158-174.
Simonson, Itamar, and Amos Tversky: 1992, 'Choice in Context: Tradeoff Contrast and Extremeness Aversion', *Journal of Marketing Research* 29, 281-295.
Spohn, Wolfgang: 2005, 'Induktion', in Wolfgang Spohn, Peter Schroeder-Heister and Erik J. Olsson (eds.), *Logik in der Philosophie*, Heidelberg: Synchron Wissenschaftsverlag der Autoren, pp. 137-159.
Stalnaker, Robert: 2009, 'Iterated Belief Revision', *Erkenntnis* 70, 189-209.
Tversky, Amos: 1972, 'Elimination by Aspects: A Theory of Choice', *Psychological Review* 79, 281-299.
Tversky, Amos, and Itamar Simonson: 1993, 'Context-Dependent Preferences: The Relative Advantage Model', *Management Science* 39, 1179-1189.
Ullmann-Margalit, Edna, and Sidney Morgenbesser: 1977, 'Picking and Choosing', *Social Research* 44, 757-785.

Department of Philosophy
University of Regensburg
93040 Regensburg
Germany

Email: hans.rott@psk.uni-r.de