HANNU NURMI[22]

Institution Design and the Preference Ranking Assumption

Introduction

The principle of rationality is often invoked in non-scientific contexts as a means of understanding behavior. The concept of rational behavior is also the main ingredient of modern decision theory. No decision theorist would maintain that all human behavior is rational, but many would probably argue that rational behavior provides a useful benchmark for evaluating, explaining and predicting various forms of behavior. In particular, if observed behavior is found to agree with the dictates of rationality, no further explanation is typically needed for it. It is behaviors that exhibit deviations from rationality that require explanation. But what is then rational behavior?

The most precise definition - due to Savage (1954) - is based on a simple choice situation involving two alternatives, say, A and B. Suppose that the individual making the choice has a strict preference over these two so that he/she (hereinafter he) strictly prefers A to B. Choice behavior is then called rational if it always, that is, with probability 1, results in A being chosen (see also Harsanyi 1977). Of course, we may encounter situations where the individual is physically prevented to choose A e.g. by making him believe that A is not really available or that by taking some new aspects of the situation into account, he does not prefer A to B or something similar. These kinds of considerations are, however, irrelevant since by suggesting that A is not available, the situation is no longer one involving a choice. Similarly, if the individual is led to believe that he is actually preferring B to A, the "original" preference no longer holds. So, we may argue that the definition holds at least as far as preference-based rationality is concerned. In this setting it is quite straight-forward and trivial to argue that rational behavior aims at utility maximization since by assigning a larger utility value to A than to B, we guarantee that preferences coincide with utility maximization.

Things get more complicated when the alternative set is expanded. The standard way to proceed is to impose conditions on preference relations that guarantee that acting in accordance with preferences amounts to maximizing utilities. In fact, the theory of choice under certainty, risk and uncertainty focuses precisely on those conditions. In what follows we shall restrict ourselves to the most elementary decision setting involving a finite set of more than two alternatives, viz. the choice under certainty. The standard representation theorem (see e.g. Harsanyi 1977: 31) states that if the individual is endowed with a continuous, complete and transitive weak preference relation over

[22] Constructive comments of Tapio Raunio on an earlier version are gratefully acknowledged.

the alternatives, then his choice behavior -- if it conforms with his preferences -- can be represented as utility maximizing. In the following sections we shall consider each one of these properties of preference relations in turn and discuss their plausibility. Our aim is to show that under relatively general circumstances each one of them can be questioned. We shall thereafter endeavour to show that individual preference tournaments could provide a useful starting point for modeling reason-based behavior and a more plausible benchmark than the traditional preference-based rationality.

Rationality vs. optimization

Is rationality always associated with optimization? This question has been discussed by many authors, especially after Simon introduced (1983) the idea of bounded rationality and satisficing. Simon argues that this is not the case. Mongin (2000) examines the plausibility of the opposite implication and asks whether optimization implies rationality. The question is well worth posing even if one accepts Simon's view which leaves this implication basically open. After careful analysis, Mongin concludes that the answer is negative, optimization does not imply rationality.

For our purposes it is not necessary to re-examine Mongin's and Simon's arguments in detail, but it is important to point out the difference of our focus vis-à-vis Mongin's and Simon's. We basically accept the thin rationality view, i.e. rational behavior amounts to acting upon one's preferences in two-alternative settings. In multi-alternative ones the conditions imposed on preferences (completeness, transitivity, continuity) guarantee the existence of a utility function and acting upon preferences, therefore, means maximizing one's utility function. We find this plausible. That is, if a person's preferences satisfy the conditions mentioned, then it makes sense to call his behavior rational to the extent he acts upon his preferences and maximizes his utility function. However, the conditions needed for the utility function to exist may not hold. It will be argued in the following, that reasonable (rational) choices can still be made and are actually being made. What is being rejected here is the position according to which complete and transitive preference relations are necessary and sufficient for rational behavior. They may be sufficient, but not necessary.

In the following some new arguments are presented in support of the claim that it is quite plausible that individual preferences are occasionally intransitive, incomplete and discontinuous. Our approach that stems from multiple criterion decision theory is new, although some of the conclusions - especially those that pertain to intransitivity - are well known from previous literature.

Can one question transitivity?

It is common to assume that preferences are revealed by choices. This is, in fact, stated in the definition of preference. In the world of empirical observations it may, however, happen that a person may, for one reason or another, occasionally choose B even though his preference is for A over B. It would, then, be more plausible to translate the preference of A over B into a probability statement according to which the probability of A being chosen by the person is larger than the probability that B is chosen. Starting from this somewhat milder probability definition of preference, we shall now consider the transitivity property. May (1954) suggests that the appropriate definition of preference-based choice is one that - in addition to choice probability - includes the alternative set considered as well as the description of the experimental setting. In this framework the preference for A over B is expressed as the following probability statement:

$$p(A|A, B, E) > p(B|A, B, E)$$

Here E denotes the experimental setup.

Suppose now that A is preferred to B and B is preferred to C. I.e.

$$p(A|A, B, E) > p(B|A, B, E) \quad (1)$$

$$p(B|B, C, E) > p(C|B, C, E) \quad (2)$$

Now, transitivity would require that inequality (1) and inequality (2) imply that

$$p(A|A, C, E) > p(C| A, C, E) \quad (3)$$

It is, however, difficult to associate this implication with rationality, since the alternative sets considered are different in each equation: in inequality (1) it is {A, B}, in inequality (2) it is {B, C} and in inequality (3) it is {A, C}. What May (1954: 2) argues is "that transitivity does not follow from this empirical [probabilistic] interpretation of preference, but must be established, if at all, by empirical observation." This point on which we completely agree leaves, however, open the possibility that transitivity would be normatively compelling (even if empirically contestable).

Our position is stronger here: while we agree that there are circumstances where transitivity seems normatively plausible[23], there are others where it is not. Hence, defining rationality so that transitivity of preferences is a necessary part of it, is not acceptable in our view.

The reason is rather straight-forward. The grounds for preferring A over B might well be different from those used in ranking B ahead of C. Hence, it is purely contingent whether these or other grounds are used in preferring C to A or vice versa. Alternatively, the decision maker may use several criteria of "performance" of alternatives. Each of these may result in a complete and transitive relation over alternatives, but when forming the overall preference relation on the basis of these rankings, the decision maker may well end up with a intransitive relation. Consider a fictitious example.

[23] E.g. in preferences over monetary payoffs.

Three universities A, B and C are being compared along three criteria: (i) research output (scholarly publications), (ii) teaching output (degrees), (iii) external impact (expert assignments, media visibility, R& D projects, etc.).

publications	teaching	external impact
A	B	C
B	C	A
C	A	B

Assuming that each criterion is of roughly equal importance, it is natural to form the overall preference relation between the universities on the basis of majority rule: which one of any two universities is ranked higher than the other on at least two criteria is preferred to the latter. In the present example, this leads to a cycle: A > B > C > A >... (Here ">" is to be read "is preferred to"). Hence, intransitive individual preference relations can be made intelligible by multiple criterion setting and majority principle (cf. May 1954; Fishburn 1970; Bar-Hillel & Margalit 1988).

Given that intransitivity of preferences may plausibly emerge in decision making under certainty, it is to be expected that it may also occur under the risk modality, i.e. in choosing among lotteries. The attractiveness of a lottery depends on two components: the amount of payoff and the probability of winning it. Let there be a sequence of two-outcome lotteries (A_i, p_i; 0, 1- p_i), i = 1, ..., k where a positive sum A_i is won with probability p_i and nothing is won with the complementary probability 1-p_i. Suppose that these lotteries form a sequence so that moving from top to bottom of the sequence, the payoffs diminish in each step, i.e. $A_j > A_{j+1}$, for each j = 1,..., k-1. Assume moreover that $p_{j+1} > p_j$, for each j = 1, ..., k-1, i.e. the winning probabilities increase when moving down the sequence. Then it is quite plausible to prefer the first lottery to the second one (because of the payoff difference), the second to the third etc., but at some point, say at stage m, to prefer the m'th lottery to the first one (because of the probability difference). In fact, Tversky (1969) found that this kind of behavior is common in choice experiments under risk.

Or completeness?

As we pointed out above, the completeness of weak preference relation entails that for any pair (A, B) of alternatives either A is preferred to B or *vice versa* or both. Stated in another way, completeness means that it cannot be the case that A is not preferred to B and B is not preferred to A. In the following we show that there is nothing unnatural or irrational in situations where there are grounds for saying that neither A is preferred to B nor B is preferred to A.

Perhaps the simplest, albeit heretofore unnoticed, way to show this is via a phenomenon known as Ostrogorski's paradox. It refers to the ambiguity in determining the popular preference among two alternatives (Daudt & Rae 1978). In the following we recast this paradox in an individual decision-making setting.

The nominating individual is to make a choice between two alternatives A and B, e.g. applicants to the chair of political science in a university. Three kinds of merits are deemed of primary importance for this office, viz. research accomplishments, teaching skills and ability to attract external funding to the university. The nominating individual has sought advice from three other individuals: one representing the peers (i.e. other political science professors), one representing the students of political science and one representing the university administration. The following table indicates the preferred applicant of each representative on each area of merit. Thus, e.g. applicant A has a more preferable research record according to the peers than applicant B. Similarly, the representative of the administration deems B preferable in each merit area.[24]

merit area	research	teaching	funding potential	row choice
advisor 1	A	B	A	A
advisor 2	A	A	B	A
advisor 3	B	B	B	B
column choice	A	B	B	?

Suppose now that the nominating individual forms his preference in a neutral and anonymous manner, i.e. each merit area and each advisor is considered equally important. It would then appear natural that whichever applicant is deemed more suitable by more advisors than its competitor, is preferable in the respective merit area. Similarly, whichever candidate is more suitable than his competitor in more merit areas is regarded as preferable by each advisor.

Under these assumptions the nominating individual faces a quandary: if the aggregation of valuations is first done over columns - i.e. each advisor's overall preference is determined first - and then over rows - i.e. picking the applicant regarded more appropriate by the majority of advisors - the outcome is that B cannot be preferred to A. If the aggregations are performed in the opposite order - first over rows and then over columns - the outcome is that A cannot be preferred to B. Hence, the preference relation over {A, B} is not complete.

It should be observed that there is nothing arbitrary or irrational in the above example. The use of expert information (advisors) or other evaluation criteria in assessing applicants would seem quite obvious way to proceed. Also, the duties to be performed by the successful applicant often have several aspects (merit areas) to them. Similarly, the use of majority principle in determining the "winners" of aggregation is quite reasonable, certainly not counterintuitive.

This chapter is by means the first to question the plausibility of the completeness condition. In the beginning of 1960's Aumann (1962) pursued the possibility of constructing a theory of utility without this condition. His focus is, however, in expected utility theory under risk where the decision maker is faced with risky prospects or

[24] The composition of the advisory body may raise some eyebrows. If so, then instead of these particular categories of advisors, one may simply think of a body that consists of three peers.

lotteries, i.e. probability distributions over certain outcomes of the type described in the last paragraph of the preceding section. Our interest is mainly in the more elementary setting, viz. decision under certainty. Despite this difference in focus, it is interesting to note that Aumann sees the completeness condition as the most problematic among the conditions underlying those guaranteeing that preferences can be represented by utility functions. In fact, some doubt concerning the empirical and normative plausibility was expressed by the pioneers of decision theory, von Neumann and Morgenstern (2004: 28), some seventy years ago.

What about continuity?

Continuity condition states that both the inferior and superior sets for any given alternative are closed (Harsanyi 1977: 31). To elaborate this a little, consider a set X of alternatives and an element x in it. Let now x_1, x_2, ..., a sequence of alternatives converging to x_0, have the property that for each x_i in the sequence, $x >_j x_i$. In other words, individual j prefers x to each element of the sequence. Then, continuity requires that $x >_j x_0$ as well. Similarly, if the sequence has the property that $x_i >_j x$, then $x_0 >_j x$ as well. Intuitively stated, continuity requires that small changes in the alternatives are accompanied with small changes in their desirability.

Let us now see how continuity assumption translates into multiple-criterion settings. We shall take advantage of Baigent's (1987) fundamental result in social choice theory. This result has subsequently been augmented, modified and generalized by Eckert and Lane (2002), Baigent and Eckert (2004), as well as by Baigent and Klamler (2004). We shall, however, first and foremost make use of the early version (Baigent 1987). It states the following.

Theorem (Baigent) *Anonymity and respect for unanimity of a social choice function cannot be reconciled with proximity preservation.*

Proximity preservation is a property defined for social choice functions. It amounts to the requirement that choices made in profiles more close to each other ought to be closer to each other than those made in profiles less close to each other. Profiles -- it will be recalled -- are n-tuples of preference rankings over the set of alternatives (n being the number of individuals). What this requirements intuitively means is that if we make a small modification in the preference rankings, the outcome of the social choice function should change less than if we make a larger modification. Anonymity, in turn, requires that relabeling of the individuals does not change the choice outcomes.

In multi-criterion setting anonymity means that permuting the criteria does not change the outcome of evaluation. Respect for unanimity is satisfied whenever the choice function agrees with a preference ranking held by all individuals, i.e. if $x >_j y$ for all individuals j, then this will also be the social ranking between x and y. In multi-criterion environment this amounts to the requirement that if all criteria suggest the same ranking of alternatives, then this ranking should also be the outcome.

To illustrate the incompatibility exhibited by Baigent's theorem, let us turn again to the fictitious example of nominating the chair of political science. Suppose that there are two applicants A and B. Moreover, only two criteria are being used by the nominating authority: research merits (R) and teaching record (T).[25] To simplify things further, assume that only strict preferences are possible, i.e. each criterion produces a strict ranking of the applicants. Four different configurations of rankings (S_1, ... , S_4) are now possible:

S_1		S_2		S_3		S_4	
R	T	R	T	R	T	R	T
A	A	B	B	B	A	A	B
B	B	A	A	A	B	B	A

Let us denote the rankings in various configurations by P_{mi} where m is the number of the configuration and i the criterion. We consider two types of metrics: one that is defined on pairs of rankings and one defined on configurations. The former is denoted by d_r and the latter by d_P. They are related as follows:

$$d_P (P_m, P_j) = \sum d_r (P_{mi}, P_{ji}).$$

In other words, the distance between two configurations is the sum of distances between the pairs of rankings of the first, second, etc. criterion.

Take now two configurations, S_1 and S_3, from the above list and express their distance using metric d_P as follows:

$$d_P (S_1 ,S_3) = d_r (P_{11}, P_{31}) + d_r (P_{12}, P_{32}).$$

Since, $P_{12} = P_{32} = A > B$, and hence the latter summand equals zero, this reduces to:

$$d_P (S_1 ,S_3) = d_r (P_{11} , P_{31}) = d_r ((A > B), (B > A)).$$

Taking now the distance between S_3 and S_4, we get:

$$d_P (S_3 , S_4) = d_r (P_{31} , P_{41}) + d_r (P_{32}, P_{42}).$$

Both summands are equal since by definition:

$$d_r((B > A), (A > B)) = d_r ((A > B), (B > A)).$$

[25] The argument is a slight modification of Baigent's (1987, 163) illustration.

Thus,

$$d_P(S_3, S_4) = 2\, d_r((A > B), (B > A)).$$

In terms of d_P, then, S_3 is closer to S_1 than to S_4. Intuitively this makes sense.

We now turn to procedures used in aggregating the information on criterion-wise rankings into an overall evaluation or choice. Let us denote the aggregation procedure by g. We make two intuitively plausible restrictions on choice procedures, viz. that they are anonymous and respect unanimity. In our example, anonymity requires that whatever is the choice in S_3 is also the choice in S_4 since these two profiles can be reduced to each other by relabeling the criteria. Unanimity, in turn, requires that $g(S_1) = A$, while $g(S_2) = B$. Therefore, either $g(S_3) \neq g(S_1)$ or $g(S_3) \neq g(S_2)$. Assume the former. It then follows that $dr(g(S_3), g(S_1)) > 0$. Recalling the implication of anonymity, we now have: $d_r(g(S_3), g(S_1)) > 0 = d_r(g(S_3), g(S_4))$. In other words, even though S_3 is closer to S_1 than to S_4, the choice made in S_3 is closer to - indeed identical with – that made in S_4. This argument rests on the assumption that $g(S_3) \neq g(S_1)$. Similar argument can, however, easily be made for the alternative assumption, viz. that $g(S_3) \neq g(S_2)$.

The example shows that small mistakes or errors in criterion measurements are not necessarily accompanied with small changes in evaluation outcomes. Indeed, if the true criterion rankings are those of S_3, then a mistaken report on criterion R leads to profile S_1, while mistakes on both criteria lead to S_4. Yet, the outcome ensuing from S_1 is further away from the outcome resulting from S_3 than the outcome that would have resulted had more - indeed both - criteria been erroneously measured whereupon S_4 would have emerged. This shows that measurement mistakes do make a difference.

It should be emphasized that the violation of proximity preservation occurs in a wide variety of aggregation systems, viz. those that satisfy anonymity and unanimity. This result is not dependent on any particular metric with respect to which the distances between profiles and outcomes are measured. Expressed in another way the result states that in nearly all reasonable aggregation systems it is possible that a small number of measurement errors has greater impact on evaluation outcomes than a large number of errors.

Eckert and Lane (2002) strengthen Baigent's theorem by showing that any choice rule satisfying anonymity and proximity preservation is imposed. A rule is imposed if it results in a constant outcome regardless of the opinions of the individuals. Hence, by adopting an anonymous and proximity preserving choice rule, one cannot guarantee even the slightest degree of responsiveness of the rule.

The two theorems (Baigent's as well as Eckert & Lane's) - when interpreted in the multiple-criterion choice context - do not challenge completeness or transitivity of individual preferences, but call into question the continuity of preferences, i.e. their representation by smooth utility functions.

Acting upon reasons

The upshot of the preceding is that all assumptions underlying the utility maximization theory can be questioned, not only from the descriptive accuracy but also from the normative point of view. The deviations from the assumptions described above are not unreasonable or irrational. In fact, it can be argued that they are just the opposite, viz. based on reasons for having opinions (cf. Dietrich & List 2013). Incompleteness of preference relations as exhibited by Ostrogorski's paradox is a result of a systematic comparison of alternatives using a set of criteria and a set of aspects or dimensions or purposes ("functions") that the alternatives are associated with. There is a reason for the incompleteness: simple majority rule gives different results when row-column aggregation or column-row aggregation is resorted to.

The simple majority rule is not the sole culprit: the paradox can occur with super-majority rules as well. The point is that one can build a plausible argument for the incompleteness under some circumstances. The same goes for intransitivity. The argument is, however, somewhat different in invoking reasons for having a given binary preference: the reason for preferring A to B may differ from the one for putting B ahead of C and this, in turn, may differ from the basis for preferring A to C or vice versa. As May (1954) pointed out, the basic sets from which choices are made are different in each of these three cases.

The eventual failure on continuity rests on yet another consideration. By Baigent's theorem any rule that is anonymous (does not discriminate for or against individuals) and respects unanimity (in agreeing with the ranking if that happens to be identical for all individuals) can lead to discontinuities. One could, hence, argue that any reasonable rule is prone to discontinuous utility representations.

To reiterate: the grounds for deviating from the assumptions of utility maximization are normative, not just descriptive. In other words, it makes perfect sense to have preferences that deviate from the assumptions. The question now arises: are there alternatives to these assumptions that could be used in analyzing individual choice behavior? In what follows we shall argue that there are and, moreover, these alternatives provide adequate foundations for institutional design.

Deriving rankings from tournaments

The most natural way of handling intransitive preference relations is to start from complete relations and look for methods to aggregate them. This approach has a long history. An early precursor is Ramon Lull, a 13th century Catalan monk and mystic (Hägele & Pukelsheim 2001; Szpiro 2010). Of the more recent pioneers, the most important is undoubtedly Ernst Zermelo (1929). The starting point is the concept of tournament, *i.e.* a complete and asymmetric relation. With a finite (and small) number k of alternatives this can conveniently be represented as a k-by- k matrix where the element a_{ij} on the i'th row and j'th column equals 1 whenever i'th alternative is preferred to the j'th one. Otherwise, the element equals 0.

Given an individual preference tournament we might be interested in forming a ranking that would preserve the essential features of the tournament, while at the same time augmenting it so that a complete and transitive relation emerges. The latter might be necessary *e.g.* for aggregating individual preference information to end up with a social ranking or choice. By the fundamental result of Edward Szpilrajn (1930) every partial order - that is a asymmetric and transitive relation - has a linear extension. In other words, if the individual gives a preference relation that is asymmetric (strict preferences only) and transitive, but not complete (not all pairs of alternatives are comparable), then preference rankings can be constructed that preserve those aspects provided by the individual.

The problem is that the resulting rankings are rarely unique. In fact, if x and y are two non-comparable alternatives in the relation given by the individual, there are rankings in which $x > y$ and rankings in which $y > x$ (Dushnik & Miller 1941). Thus, there seems to be no general way of extending a partial order into a unique linear one. However, tournaments impose less structure into individual preferences than partial orders. After all, they are complete and asymmetric, not necessarily transitive. Over past decades many ways of translating tournaments into rankings have been suggested. The usual way - called scoring method by Rubinstein (1980) - is the straight-forward summing of row entries in the tournament matrix whereby one ends up with a score s_i for each alternative i. The ranking over the alternatives is then determined by the order of scores. The resulting ranking is, of course, weak since several alternatives may receive the same score.[26] The scoring method may, however, lead to an outcome ranking where a higher rank is given to an alternatives that is deemed inferior to, or defeated by, one or several of the lower ranked ones. Several methods to avoid this problem has been suggested. Thus, for example, Goddard's (1983) proposal is to choose those rankings that minimize the number of times a binary preference between any two alternatives is upset (i.e. reversed) in the outcome ranking.[27] Upon closer inspection this proposal turns out to be similar to Kemeny's (1959) rule.

Viewed as a social choice function this rule has a host of desirable properties (see e.g. Nurmi 2012: 257). It is, however, intended for finding the "closest" social ranking for any given set of individual rankings over several alternatives. A function that - given a set of individual preference tournaments - looks for the collective one that is closest to the individual tournaments in a specific sense is - regrettably nowadays largely forgotten - Slater's (1961) rule. It seems identical with the rule that Goddard advocates. It works, as was already stated, on the basis of individual tournaments, *i.e.* complete and asymmetric relations. It then generates all k! complete and transitive relations (strict rankings) that can be obtained from the k alternatives and converts them into tournament matrices. Each of these generated matrices is then a candidate

[26] In social choice literature the scoring method described here is known as Copeland's rule and usually dated to 1950's. Arguably it was, however, introduced already in late 13'th century by Ramon Lull (Hägele and Pukelsheim 2001).

[27] Goddard is not the first to suggest this method. For earlier discussions, see Kendall (1955) and Brunk (1960).

for the collective preference tournament (i.e. the winning tournament). The winning tournament has the distinction that it is closest to the individual tournaments in the sense that it requires the minimum number of changes from 0 to 1 or vice versa in individual opinions to be unanimously adopted.

The principle of Slater's rule can, of course, be used in individual decision making as well. To wit, given an individual preference tournament one generates the tournaments corresponding to all k! preference rankings involving the same number of alternatives. One then determines whether the individual tournament coincides with one of them. If it does, then this gives us the ranking we are looking for. Otherwise one determines which of the generated tournaments is closest to the individual's. The closest one indicates the ranking. It may happen that there are several equally close tournaments and thus there may be several "solutions".

Zermelo's (1929) approach to tournaments is based on observations of chess playing contests which often take the form of a tournament.[28] Each player plays against every other player several times. The outcome of each game is either a victory of one player or a tie. We assume that the games are independent binomial trials so that the probability of player i beating player j is p_{ij}. Zermelo then introduces the concept of Spielstärke, playing strength, denoted by V_i, that determines the winning probability as follows:

$$p_{ij} = V_i / (V_i + V_j).$$

The order of the V_i values is the ranking of the players in terms of playing strength. Apparently player i is ranked no lower than player j if and only if $p_{ij} \geq 1/2$, *i.e.* players with greater strength defeat contestants with smaller strength more often than not. Now, given the matrix A of results, i.e. a k–by-k matrix of 0's and 1's denoting losses and victories of the alternatives represented by the rows, Zermelo defines maximum likelihood estimates, denoted by v_i, for the playing strengths of players. Consider any k vector of strengths v. One can associate with it the probability that the observed matrix A is the result of the tournament when the strengths are distributed according to v. The probability is the following:

$$p(v) = \prod_{i,j} (v_i / (v_i + v_j))$$

and this is what is to be maximized. Conditions under which a unique maximizing vector of strengths can be found are discussed by Zermelo and found to be rather general. A particularly noteworthy property of the Zermelo rankings is that they always coincide with the rankings in terms of scores defined above. So, were one interested in rankings only, the easy way to find them is simply to compute the scores. However, the v_i values give us more information about the players than just their order of strength; it also reveals how much stronger player i is when compared with player j.

[28] The differences between Zermelo's and Goddard's approaches are cogently analyzed by Stob (1985). Much of what is said in this and the next paragraph is based on Stob's brief note.

Leaving aside now the game context and looking at Zermelo's method from the point of view of fuzzy systems, it is not difficult to envision a new interpretation whereby the outcome matrix expresses the individual's choice between pairs of alternatives. The values V_i and their estimates v_i can be viewed as values of *desirability* of alternatives. A ranking based on desirability of alternatives is certainly a worthy goal of inquiry and Zermelo's approach gives us plausible way to achieve it.[29]

One more point on tournaments is worth making, viz. the individual's preference relation may be incomplete and we are still able to construct a tournament matrix. The score of alternatives not included in the preference relation is determined solely on the basis of comparisons with those alternatives which are included in the preference relation. Hence, tournament methods are capable of handling incomplete tournaments as well. In fact, incomplete tournaments are the starting point of Zermelo's analysis.

The above remarks pertain to situations where we are given an individual preference tournament (complete or incomplete) and, for one reason or another, are looking for a ranking that would best approximate it. It is, however, quite easy to envision situations where no ranking at all is required, but rather choice of a subset of "best alternatives".

Tournament solutions

Since tournaments are basically sets of pairwise comparisons, it is customary to suggest that Condorcet winners be elected whenever they exist. In other words, all tournament solutions ought to be Condorcet extensions. Starting from this requirement leaves us open to the criticism stemming from Saari's (1995) important results regarding the instability of Condorcet winners under certain transformations of preference rankings (esp. adding or subtracting Condorcet components, that is, groups of voters whose rankings form fully symmetric Condorcet paradox profiles). Yet, basically all currently discussed tournament solutions are Condorcet extensions and, assuming that no ordinal information about alternatives is at hand, this makes sense.

Since all Condorcet extensions result in the same outcome when a Condorcet winner exists, the variations come about when this is not the case, i.e. when no alternative defeats all the others in pairwise contests with a majority of "votes". In individual decision making contexts, the non-existence of a Condorcet winner means that for each alternative there is another that is preferred to it on a majority of criteria. What would be the plausible choice, then, under such circumstances?

A plausible subset of winning alternatives can be formed by using the following dominance relation over the alternatives: alternative x dominates alternative y if it defeats by a majority of criteria not only y but also all those alternatives that y defeats. Obviously, such a dominance relation is asymmetric and transitive, but not complete. Nonetheless, it enables us to define the set of uncovered alternatives, *i.e.* a set of those

[29] We shall here ignore the ties in pairwise comparisons. These can certainly be dealt with in fuzzy systems theory. Also the tournament literature referred to here is capable of handling them. Ties are typically considered as half-victories, i.e. given a value 1/2 in the tournament matrices.

alternatives that are not dominated by any other alternative. This set can, however, be quite large. Indeed, it may include all alternatives as in the case of Condorcet's paradox.

On the other hand it may also be a relatively small subset of the alternatives and when a Condorcet winner exists, it collapses into this single alternative. Two important subsets of the uncovered set can be defined: the Banks set and the set of Copeland winners (Banks 1986; Moulin 1988). The latter represents the long tradition of tournament solutions, while the former can be used to characterize all outcomes of sophisticated - as opposed to sincere – voting in amendment systems that consist of pairwise comparisons of alternatives. The Copeland winners have a practical advantage over the Banks set: the scores underlying it can be used to form a ranking. Both sets are always nonempty and, hence, genuine Condorcet extensions.

Conclusion: institutions based on tournaments

We have attempted to show above that there are quite plausible reasons for individuals to deviate from the behavior dictated by preference-based utility maximization theory. Indeed, behavior based on reasons would seem to be particularly prone to these kinds of deviations. Rankings being the basic concept underlying the maximization theory, our main conclusion is that alternatives to ranking assumption already exist. One of these, individual preference tournament, has been discussed above.

Of particular interest is the re-discovery of Zermelo's approach to tournaments since it provides a natural link between directly observable pairwise choices and the underlying notion of desirability. It thus provides a method for estimating preference degrees for observational data. Replacing individual preference rankings with tournaments enables us to deal with more general choice settings than those underlying the maximization theory. However, this comes with a cost. To wit, if the individuals possess complete and transitive preference relations over the alternatives (*i.e.* have rankings), this ordinal information is essentially lost in the aggregation based on the induced tournaments. In particular, when the tournaments are used first in aggregating individual opinions into collective tournaments and then into collective preference rankings, the outcome may be quite different from one obtained by aggregating the original preference rankings e.g by Kemeny's rule. As pointed out above, however, there are the conditions underlying preference rankings that are not always satisfied. What we have aimed to suggest is that reasonable choices and even rankings can be made in their absence. Hence, there is a case to be made for institutions aggregating individual preference tournaments.

References

Baigent, N. (1987). Preference proximity and anonymous social choice. *The Quarterly Journal of Economics* 102(1): 161-169.

Baigent, N. & Eckert, D. (2004). Abstract aggregations and proximity preservation: an impossibility result. *Theory and Decision* 56(4): 359-366.

Baigent, N. & Klamler, C. (2004). Transitive closure, proximity and intransitivities. *Economic Theory* 23(1): 175-181.

Banks, J. (1985). Sophisticated voting outcomes and agenda control. *Social Choice and Welfare* 1(4): 295-306.

Bar-Hillel, M. & Margalit, A. (1988). How vicious are cycles of intransitive choice? *Theory and Decision* 24(2): 119-145.

Brunk, H. (1960). Mathematical models for ranking from paired comparisons. *Journal of the American Statistical Association* 55(291): 503-520.

Dietrich, F. & List, C. (2013). A reason-based theory of rational choice. *Nous* 47(1): 104-134.

Dushnik, B. & Miller, E. (1941). Partially ordered sets. *American Mathematical Monthly* 63: 600-610.

Eckert, D. & Lane, B. (2002). Anonymity, ordinal preference proximity and imposed social choices. *Social Choice and Welfare* 19(3): 681-684.

Fishburn, P. C. (1970). The irrationality of transitivity in social choice. *Behavioral Science* 15(1): 119-123.

Goddard, S. (1983). Ranking in tournaments and group decision making. *Management Science* 29(12): 1384-1392.

Hägele, G. & Pukelsheim, F. (2001). Lull's writings on electoral systems. *Studia Lulliana* 41: 3-38.

Harsanyi. J. C. (1977). *Rational Behavior and Bargaining Equilibrium in Games and Social Situations.* Cambridge: Cambridge University Press.

Kemeny, J. (1959). Mathematics without numbers. *Daedalus* 88(4): 571-591.

Kendall, M. (1955). Further contributions to the theory of paired comparisons. *Biometrics* 11(1): 43-62.

May, K. O. (1954). Intransitivity, utility and aggregation of preference patterns. *Econometrica* 22(1): 1-13.

Mongin, P. (2000). Does optimization imply rationality? *Synthese* 124(1-2): 73-111.

Moulin, H. (1986). Choosing from a tournament. *Social Choice and Welfare* 3(4): 271-291.

Nurmi, H. (2012). On the relevance of theoretical results to voting system choice. In D. Felsenthal & M. Machover (eds), *Electoral Systems. Paradoxes, Assumptions, and Procedures*. Berlin- Heidelberg: Springer Verlag, 255-274.

Rubinstein, A. (1980). Ranking the participants in a tournament. *SIAM Journal of Applied Mathematics* 38(1): 108-111.

Simon, H. (1983). *Models of Bounded Rationality*. Cambridge, MA: MIT Press.

Slater, P. (1961). Inconsistencies in a schedule of paired comparisons. *Biometrica* 48(3/4): 303-312.

Stob, M. (1985). Rankings from round-robin tournaments. *Management Science* 31(9): 1191-1195.

Szpilrajn, E. (1930). Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae* 16(1): 386-389.

Szpiro, G. (2010). *Numbers Rule: The Vexing Mathematics of Democracy from Plato to the Present*. Princeton: Princeton University Press.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review* 76(1): 31-48.

Von Neumann, J. & Morgenstern, O. (2004). *Theory of Games and Economic Behavior*. 60'th Anniversary Edition. Princeton: Princeton University Press.

Zermelo, E. (1929). Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift* 29(1): 436-460.