Rank Preservation and Reversal in Decision Making

Thomas L. Saaty*

Distinguished University Professor, University of Pittsburgh, USA

Abstract: There are numerous real life examples done by many people which show that the alternatives of a decision sometimes can reverse their original rank order when new alternatives are added or old ones deleted and without bringing in new criteria. There is no mathematical theorem which proves that rank must always be preserved and there cannot be because of real life and hypothetical counter examples in decision making methods. Rank preservation came to be accepted as the standard because of techniques that could only rate alternatives one at a time treating them as independent. Thus an alternative receives a score and it will not change when other alternatives are added or deleted. All methods that only rate alternatives one at a time, thus always preserving rank, may not lead to the right decision; even if they may be right in certain areas of application. In reality, to determine how good an alternative is on an intangible criterion needs experience and knowledge about other alternatives and hence in their evaluation, the alternatives cannot be completely considered as independent of one another.

Keywords: Decision Making, priorities, ranking, rank preservation, rank reversal.

1. INTRODUCTION

In decision making, ranking and prioritization are subjective through and through. They depend on human experience, understanding and interpretation. Occasionally data show that an event that occurs more often is important. But that too is subject to interpretation. Why do we want to rank things, like the alternatives of any decision? We do it to choose the best one or to admit or limit the number of candidates to a school, a hospital, award winners, search engine optimization and many others. Multicriteria ranking is a complicated process in which not only the alternatives are ranked, but also the criteria themselves must be prioritized and ranked and their priorities play a significant role in synthesizing the priorities of the alternatives to derive a final overall ranking for them. Because prioritization and ranking need judgments, and judgments are primarily subjective, one may argue that ranking is contingent on who provides those judgments and how well they are elicited and combined to produce a rational synthesis of the outcome that yields the priorities.

It appears that the outcome of ranking depends on which methods of ranking that are available one uses. But that is not always the case and there may be doubt about a method that is often at odds with the results obtained by other methods. A good decision also depends on the people who often must be experts in the subject of that decision.

Psychologists tell us that to evaluate things we must either compare them with one another or rate them one at a time with respect to a hypothetical ideal we have in mind on the property that they must be ranked. The ideal itself depends on shared experience to define it and is also subjective because one group may have a very different ideal as in different culture or religions.

Before the subject of comparisons was introduced, people thought that they must always compare the alternatives with respect to an ideal. By doing that, alternatives were thought to be independent of one another. However, that was a wrong assumption because, as we said before, the ideal itself is obtained from implicit understanding and comparison of many alternatives and thus indirectly formalizing the ideal required dependence among the alternatives with respect to a common property. In addition, by assigning numbers to each alternative separately assumes their independence which obviously is not the case because they collectively contribute to the formulation of the ideal. The conclusion is that this approach maintains the same numerical rankings for the alternatives no matter how many more alternatives are added to or deleted from the ranking process.

The other way is to compare the alternatives in pairs as to which one dominates the other with respect to a criterion or property they have in common. We note here that the criteria themselves need to always be compared according to their importance, preference or likelihood of their frequency of occurrence in that decision. There is no escaping the use of comparisons at this level of decision making. But comparisons of alternatives need not always be made by otherwise rating them one at a time.

How good an alternative is depends on what it is compared with. If one compares an old car with a new car, it may not look as good as it may be when it is

*Address correspondence to this author at the Distinguished University Professor, University of Pittsburgh, USA; Tel: 412 621 6546;
Fax: 412 648 1693; E-mail: saaty@katz.pitt.edu

compared with another used car. Thus the ranking of alternatives very much depends on what it is compared to. Since there are no absolutes as using an ideal assumes, comparisons are very realistic in producing more objective answers. But, as we shall see below, they can also cause some personal and social problems.

There is a long story behind the subject of always preserving rank when comparisons of criteria and of alternatives were not well understood to use them. It was assumed that using an ideal is the only possible way to rank alternatives and as a result the rank should not be influenced and reverse when new alternatives are added or old ones deleted. This was known as the principle of invariance.

The principle of invariance, sometimes known as the independence from irrelevant alternatives, encountered in decision making with utility theory has been found to be false for more than 40 years, when counterexamples were first published in the literature. Essentially, the principle of invariance says that the composite rank of a set of alternatives with respect to several criteria must stay the same if new alternatives are added or old ones deleted unless adding or deleting alternatives introduces or omits criteria. In the context of increasing the probability of changing one's choice, invariance is called *regularity*.

Why should one assume that such a principle would be true in the first place and what causes one to feel strongly that the principle should be made into a law of behavior? It is likely that the invariance principle is a consequence of one at a time kind of thinking. If one were to rate alternatives one at a time with respect to a set of criteria, each alternative would be examined by itself and obtains a score that is independent of other alternatives ranked before or after it. There would be no reason why a new alternative should affect the ranking of the old alternatives unless it adds new criteria in terms of which all the alternatives must now be additionally evaluated. But life is more complicated. We often rank alternatives by comparing them with each other on each criterion. Making comparisons is an intrinsic ability that all people have. How high we perceive an alternative to stand depends on what we already know about where other alternatives stand. We would never know how good an alternative is on an intangible criterion without having known about or experienced other alternatives on that criterion. This is particularly true when the criterion is not a physical property on which one can more or less measure the intensity with which every alternative possesses that criterion and then decide how desirable that intensity is. An intangible criterion requires judgment and judgment requires experience and knowledge about many alternatives. Rating alternatives one at a time is a special case of the process of paired comparisons because to create a scale of intensities for each criterion and use it to rate alternatives requires paired comparisons.

2. DISCUSSION

Here are some examples of unjustifiable rank reversal due to Corbin and Marley [1].

The first example concerns a lady in a small town, who wishes to buy a hat. She enters the only hat store in town, and finds two hats, A and B, that she likes equally well although she leans toward A. However, now suppose that the sales clerk discovers a third hat, A', identical to A. Then the lady may well choose hat B for sure (rather than risk the possibility of seeing someone wearing a hat just like hers), *a result that contradicts regularity*. The second example involves a guest being taken out to dinner, who, in deference to his host, refrains from selecting the most expensive and also most preferred dish and selects the second most expensive one, thereby increasing the chances of the second dish being chosen, *again contradicting regularity*.

Note that if instead of the hats the lady went shopping for a best PC computer, then she might well still buy the one she likes best even if there are many copies of it. Her preferences may be identical for the computers as they were for the hats. The same number of criteria may be used whose names may be different but whose priorities are the same as those used to choose the hats. In the end the numbers used are identical for the two examples but the labels are different. What should one do?

Some people have proposed that in the case of the desirability or the "uniqueness" of a hat, one can add such "fudge" criteria as uniqueness that would make the more preferred hat less desirable. Uniqueness and manyness are group properties and are not intrinsic attributes of any single hat. They require that one look at other hats to determine if the given hat is unique. But this implies that in ranking the hats one must assume that they are dependent on each other, violating another axiom of utility theory requiring independence among the alternatives. In passing we note that when

introducing alternatives which may introduce new criteria, if one keeps adding such alternatives one would eventually run out of new criteria (and words in the dictionary to describe them) and the number of copies wins out and alters preferences. There is no way to escape this fact. In general for any decision problem, the number of alternatives can far exceed the number of criteria. To capture the effect of manyness, a procedure is required that automatically tallies how many hats there are and how desirable the given hat is. That is precisely what relative measurement does.

In the field of marketing, the effect of phantom alternatives has been observed to cause rank reversal. A car manufacturer sells two types of car–one inexpensive and not as well made as the other. To induce people to shift from the cheaper car, the manufacturer advertises that a new car with the virtues of the better car will appear on the market but the price will be much higher than that of the better made car. People are now observed to change their mind and start buying the better car. The manufacturer in fact never makes the 'phantom'. It is an advertising gimmick. There are several other generic situations of this type that can lead to rank reversal. For many decision problems ranking is made not once but three times. Once with respect to benefits, once with respect to costs, and once with respect to risks and the overall ranks are obtained by dividing (more generally taking some function of) the benefits by the costs multiplied by the risks. Thus a decision problem consists of several phases and not just one. There are situations where only the benefits or only the costs determine the outcome because one may be negligible or insignificant when compared with the other. In this framework, adding alternatives to the three structures, benefits, costs and risks, would naturally cause rank reversal.

There are two lessons learned from these examples. One is that it is we who must decide in a particular decision problem whether for that problem, rank needs to be preserved or not. It is not automatically written in the abstract structure of the real life problem itself. The other lesson is that we cannot use one and only one procedure for aggregating preferences in a multicriteria decision process once and for all. We need one procedure to preserve rank and another to allow rank to change.

It is now clear that the outcome of a decision does depend on the procedure and mathematics used for the purpose. That is precisely the reason why the mathematics must derive from a deep and flexible understanding of decision making that emulates what decision making as a process in nature, rather than dictates what one should do to make a valid decision.

The Analytic Hierarchy Process (AHP) is a decision making theory based on relative measurement. It derives cardinal scales from paired comparisons. The AHP also has a procedure that preserves rank absolutely as needed. In the absolute measurement mode of the AHP, one derives a scale of importance priorities for different recognizable intensities of each criterion. One then divides by the priority of the largest intensity and thus the value of one is given to an ideal alternative, the best imaginable choice. In this case each alternative is assigned an intensity for each criterion. The intensity is multiplied (weighted) by the priority of the criterion and these weights are added to produce a ranking for the alternative. Adding a new alternative has no effect on the ranks of the other alternatives. Rank is categorically preserved. But the AHP with its pairwise comparisons of the alternatives is also concerned with those cases where rank can and should change. It turns out that the two methods used to preserve the rank of the most desired alternative or to allow it to reverse give different results only 8% of the time [2]. Similar results were obtained for the two top alternatives, two lowest ranked alternatives and so on. The AHP also has a procedure to shorten the effort in making comparisons by comparing each alternative only with one other alternative. All these comparisons yield a scale of priorities for the alternatives with respect to each criterion.

3. WHEN TO PRESERVE RANK AND WHEN NOT TO

John Paul Sartre said that choice precedes value. We believe that this is an excellent observation because, for example, our gold standard could have been made a silver standard or a diamond standard, or something else that is scarce, durable and attractive. Now we value gold because we chose it. People sometimes speak of frequency as indicative of priority. That is different; it is not always related to human purposes.

So under what social conditions should we preserve rank, and what other general conditions do we need to allow rank to reverse?

Unlike robots, people have feelings and an image of themselves. Their social and economic status matters much for both their survival and by putting forth their best effort. Thus we must ensure that rank is preserved

in organizations like the military where people are promoted upwards in their position as a reward for their effort. In faculty appointments at schools and universities, government posts are similar to the military and involve permanent positions. More transient examples are hospitals in which patients are admitted and rarely if ever sent home because of the need for the facilities by more urgent cases. Again in college admissions students are accepted and are not un-accepted when better students send applications later. Protocol in politics requires that certain actions be preserved and honored regardless of the circumstances. In 2015 the Boston Patriots won the Super Bowl in football, but it was learned that the ball was softer, and thus easier to grasp than it should have been, but in the end, they remained the number one team despite some people advocating that they should be penalized and the cup taken back because it would have been difficult to make their opponents, Seattle, the winner without replaying the game. In Japan, age is venerated and old people are kept in their status despite the fact that a new generation that is familiar with technology and progress is on the rise. Here rank reversal would be better for the entire society.

As significantly, in a creative society, rank should be allowed to change when new alternatives are added or deleted. That should be the case, for example, when there are many copies of an alternative which depreciate its uniqueness, such for example as many copies of the outstanding Ferrari or Maserati automobiles.

Rank may also be allowed to reverse when the weights of the criteria depend on the alternatives so that a criterion becomes more or less important depending on what alternatives there are. In design problems where new things are discovered that can replace old things, rank would naturally change. When we buy things by shopping around we may find things we have not known about that become more important than what we were accustomed to. When we look for candidates to occupy a position and interview several ones, their overall rank tends to change the more candidates we interview.

4. CONCLUSIONS

The philosopher, Arthur Schopenhauer, said, “Every truth is the reference of a judgment to something outside it, and intrinsic truth is a contradiction.” If the truth is the Ideal used that is outside, then an ideal is a composite of the qualities of all alternatives and thus can change intrinsically if new alternatives are added. Thus comparison with respect to an ideal implies dependence on other alternatives and hence alternatives cannot be evaluated one at a time, they must be compared as the criteria are always compared for their priorities. All methods that do not compare alternatives must and should be suspected about the answers they give.

To summarize, rank needs to be preserved in established systematic operations and allowed to change in exploratory and tentative kinds of decisions where one must explore larger and larger population to determine the best possible. The reader may have additional suggestions to these. Any decision theory must have two ways of dealing with rank. One way is to preserve it under the appropriate circumstances and another to allow it to change, particularly in exploratory decisions.

To maintain the established order we preserve rank. To allow the established order to change, we also allow rank to change.

REFERENCES

[1] Corbin R, Marley AAJ. Random Utility Models with Equality: An Apparent, but Not Actual, Generalization of Random Utility Models. J Math Psychol 1974; 11: 274-293. http://dx.doi.org/10.1016/0022-2496(74)90023-6

[2] Saaty TL, Vargas LG. Experiments on Rank Preservation and Reversal in Relative Measurement. Math Comp Model 1993; 17/4-5: 13-18. http://dx.doi.org/10.1016/0895-7177(93)90171-T

Received on 28-05-2015 Accepted on 16-06-2015 Published on 30-07-2015