Just Decisions Using Multiple Criteria
or: Who Gets the Porsche?
An Application of Ronald R. Yager's Fuzzy Logic Method

Lothar Philipps
Institute of Philosophy of Law and Computers and Law,
Ludwig-Maximilians-University, Munich

Abstract: The article describes a decision making procedure which uses multiple criteria of differing rank. The criteria are fulfilled to varying degrees. The creation of cardinal scales is not necessary. A distribution problem is used to exemplify the procedure.

1. The Candidates

Those facing decisions must often take into account multiple criteria. This is true also for decisions that are meant to be just. Let us use the following example:
A sports fan donates a Porsche before an important international soccer match. The German Soccer Association is to award it to the player who "deserves it the most."
After the match - which the team won - the situation is this:
(a) Wing W scored both goals including the winning goal. He skilfully exploited two goal opportunities; otherwise his play was rather weak.
(b) Sweeper S applied himself in a creative and committed way, right from the start. He showed the greatest spirit and determination of all players. Most importantly, he cleared some dangerous situations in front of the goal.
(c) Centre-forward F also applied himself in a self-sacrificing manner: quite literally, for he was badly fouled and seriously injured after only ten minutes. He probably won't be able to play for a year.
d) Right half R only showed average performance; but it was his 75th international match. All other players have played in considerably fewer international matches.

There are no other likely candidates for the prize. Who should get the Porsche?

2. The Ranking of Criteria

Arranging all relevant criteria in ranked order would facilitate the decision. The following ranking suggests itself:

(a) In today's sports success is what counts. In our case it is essential that W scored two goals, including the winning goal, and that W cleared dangerous situations, preventing goals.
(b) Playing performance and a fighting spirit, here characterised by S, are also very important. But even great and willing application might not always lead to success, thus this criterion should be given rank two.
(c) To receive comfort and compassion for hardship and suffering is a view that is not alien to sports. However, realistically speaking it should be given fourth and last rank.
(d) So the remaining criterion, a player's "status" (here: the 75th international match) will receive third rank.

This is the resulting order:

Success	**4**
Application	**3**
Status	**2**
Hardship	**1**

3. What is the Use of this Ranking?

The ranking of criteria in itself does not yet yield definite results. Someone might fulfil a low-ranking criterion to the fullest extent and a high-ranking criterion only slightly. How should that person be placed?
There would be an easy answer to this question if the order of criteria and degree of fulfilment could be taken to be cardinal (i. e. ranked and showing magnitude), rather than ordinal (i.e. only ranked). A cardinal scale would not only give the ranks of the criteria but also distance between ranks. The same would apply to the degree to which each criterion is

fulfilled. The rank of a criterion could be multiplied by its degree of fulfilment [1].
Under the conditions of cardinality wing W fulfilled the highest criterion of success (rank 4) to the fullest extent (degree of fulfilment 4), giving him a score of 4 x 4 = 16. In addition, since he fulfilled the criterion "application" (rank 3), if only to a slight degree (1), his total score would be 16 + 3, or 19 points.
Sweeper S fulfilled the criterion "application" (rank 3) to the maximum (degree 4). In repeatedly clearing dangerous situations before the goal, he fulfilled the highest criterion "success" (rank 4), but not to the highest possible degree: the opposing player might have missed or the goalie might have saved the ball (degree 3). His total score would be 12 in both areas, or a total of 24 points.
S would thus get the Porsche (leaving F and R out of consideration for the moment).
He would even get the Porsche if W had played with moderate application (degree 2); but his lead would have shrunk to 2 points (24 over 22).
These are neat calculations, but are they realistic? We are allowing statements such as the following: the criterion of success is exactly twice as important as a player's status. Application in play counts three times as much as hardship endured. How can such statements be justified? How can the value of a criterion be measured so exactly? This is much more problematic than the postulation of a mere ordering. No judge would venture such definite views.
A ranking (ordinal scale) of criteria on the other hand would be possible in many cases. It could occur in court rulings [2]. But it appears that ordinal scales are not sufficient: multiplication, for example, is not possible and therefore the different degree to which each criterion is fulfilled could not fully be taken into account. The highest ranking criterion would decide, even if present to only slight degree; only if absent would the next highest criterion replace it as the single deciding factor. Such an procedure would seem very rough [3].
We sense instinctively that there are instances in which the strong prevalence of a lower-ranking criterion eclipses a weak case of a higher-criterion. Such decisions are encountered in judicial sentencing of criminal cases or in grading exams.

4. A Fuzzy Logic Look at Rankings

The American computer scientist Ronald R. Yager suggested a procedure in which decisions could be made using multiple criteria in a non-cardinal ranking [4]. The procedure uses fuzzy logic. In several cases, this logic has made it possible to grasp indeterminate concepts that guide us in every-day life [5].

According to classic logic and set theory an element is part of a set or it is not - totally or not at all. According to fuzzy logic an element may be part of a set to a certain degree - more or less. Yager holds that this is the case for evaluation criteria and the phenomena that are being evaluated. His fuzzy logic interpretation of the degree to which a phenomenon fulfils a criterion is that the phenomenon belongs to a set to a smaller or greater degree. This degree of membership Yager expresses in a simple scale such as: "none (missing, insignificant), low (small, little, weak), medium, high (large, great, strong), perfect (full, complete)". In numbers: 0,1,2,3,4.
The only claim made for Yager´s scale is that its steps are linear, from lowest to highest (ordinal scale), not that the steps are of a certain size (cardinal scale). The numbers that replace the names of the steps express only that 4 is more than 3 and 3 more than 2 (and thus also that 4 is more than 2), but not that 4 means twice as much as 2. This is comparable to the grading of exams: when I give a grade of 2 points (a medium "unsatisfactory") I am not saying that it is half as good as a test I graded with 4 points (a low "sufficient") or 1/9th as good as an 18 (upper "excellent"). Frankly, I would not be able to grade with such exactness.
The same interpretation applies to criteria. Their different ranks mean that they belong to the set "importance" to varying degrees, in our case from 0 (meaningless) to 4 (of highest importance).
It is possible for several criteria to be of the same rank, and several phenomena can fulfil a criterion to the same degree.
The degree to which a criterion is fulfilled can only affect the decision if the criterion's rank on the importance scale is taken into account. This should be interpreted as a conditional (implication): if the criterion has such-and-such importance, then the degree to which it is fulfilled has such-and-such effect.
Several suggestions have been made for the representation of implication in fuzzy logic. They commonly share the view that if only the "classical" values of 1 and 0 are used (representing either complete or no membership of the set) the classical truth table representing implication can be used.
Yager chooses a definition for the implication which directly corresponds to classical logic: $X' \cup Y$ (X is missing or Y is given - in fuzzy logic, elements can also be given to varying degrees). In Yager´s procedure the complementary value of a criterion's importance must first be determined; then a union is formed with the degree of its fulfilment [6].
To establish complementary values in an ordinal scale the ranking is reversed: the first will be the last

and the last will be the first, but those in the middle stay where they are.

perfect	4	=>	0	none
high	3	=>	1	low
medium	2	=>	2	medium
low	1	=>	3	high
none	0	=>	4	perfect

The complementary value can also be determined by subtracting the value to be negated from the highest value: 4-4=0, 4-3=1, and so forth.
To establish the value of the union of several membership values the maximum value is chosen.

5. Evaluation and Decision

We are now ready to look at our factual example, focusing on a criterion and its degree of fulfilment. We will form a union of the complement of the criterion's rank and the degree of the criterion's fulfilment ($C_n' \cup F_n$).
Sweeper S e. g. applied himself to the highest degree: 4. The complementary value of the second highest criterion "application" is 1 (4 minus 3). The greater of the two values is 4. S' performance would thus be given the value of 4. In this manner all relevant phenomena are evaluated.
To draw a conclusion for each player, an intersection is formed of all the values for the criteria. In fuzzy logic, the intersection equals the lowest affiliation to any set. The final evaluation is determined by the lowest individual value. Compare it to a convoy: in order to stay together all ships have to sail at the slowest ship's speed [7].
The decision will be made in the favour of the man whose evaluation is the highest.
In our example, this would be the picture, given certain additions that I have made beyond those facts given initially: R's 75 international matches received only a high (3), not the highest value; after all, Lothar Matthäus has played in more than 100 international matches. Also, the fact that F will be on the sick list for a year was only given a value of 3: he could have been completely incapacitated. A sports injury shouldn't produce undue sentimentality.

Sweeper S has the highest total value. He should get the Porsche. This result is satisfactory but by no means obvious. It is remarkable in that we reached it without the use of a cardinal scale or the multiplication of criterion and its degree of fulfilment, even though W, B's toughest competitor, fulfilled the top-ranking criterion "success" to the highest degree.

6. Some Questions

6.1 Why does Yagers's method work? At first glance the procedure seems to be paradoxical: a low-ranking criterion yields a high value when its complement is formed, a high-ranking criterion yields a low value. But that is just the point: a criterion's low rank is turned into a high threshold that must be come by with a high degree of fulfilment. Only from that point on fulfilling the criterion will pay off. this should be considered as well: the points given to each criterion as such have no effects on the result, since they form the minimum value for each candidate.
Conversely, a high-ranking criterion will yield a low complement and thus a low threshold: fulfilling it will start the counting early [8].
6.2 What happens in the case of a draw? In our example, both W and S received a final score of 1. How could the runner-up be determined for second or even third prizes? To further differentiate, a kind of tie-break is used. For each candidate the minimum value is eliminated (only once of course, should the same value appear more than once). The resulting minimum might be higher for one candidate who will then become runner-up. If there is still a draw the procedure can be repeated. In our example the values 1,2, and 3 have to be eliminated before W prevails

	CRITERION		Wing		Sweeper		Injured Forward		Playing Record	
	C	C'	W	C'∪W	S	C'∪S	C	C'∪F	R	C'∪R
Success	4	0	4	4	3	3	0	0	1	1
Application	3	1	1	1	4	4	2	2	2	2
Status	2	2	0	2	0	2	0	2	3	3
Hardship	1	3	0	3	0	3	3	3	0	3
TOTAL				1		2		0		1

over R with a score of 4 to 3.

Should this tie-break yield no result, further differentiation may still be possible, namely in cases in which the same degree of fulfilment is attained in criteria of differing ranks. An example: if in our case W had played with "medium " (degree 2) instead of "low" application (1), S would still win, But the necessity for a tie-break shows how close the decision has become. If we take things a step further and assume that W applied himself "greatly" (3), a draw is achieved between W and S that cannot be resolved by the tie-break procedure. But W scored the highest value in the top-ranking criterion, while S did so in the second-highest criterion. Therefore W should be given the Porsche.

6.3 How can the certainty of such decisions be determined? If one is uncertain about an individual value, it could be experimentally changed. In our example, doubts might be felt about whether F really did not contribute to the team's success even though he left the game after ten minutes.
If we experimentally raise the value for "success" to one, F would equal R (and W) with a score of 1. In a tie-break R would still prevail 3 to 2.Therefore, we see that the uncertainty concerning that particular value does not affect the result.
In relation to W no tie-break is necessary since W has already won the tie-break over R. (In Yager's method, results are in "partial order".)
Uncertainty concerning a particular value can also be a reason to refine the scales of the criteria and their fulfilment. In our case, the five-step scale could be replaced by a seven-step scale and F's effort could be termed "very little" rather than "insignificant". Refining the scale is also often a way to resolve a draw between candidates.

6.4 And again: Why not stay with multiplication of the criterion's importance by its degree of fulfilment, as demonstrated at the beginning of this paper? Multiplication seems more natural and familiar, and both methods yielded the same results, even when wing W's application was assumed to be "medium".
This may be true in our case but would not be true in every case. Using multiplication, errors and inconsistencies in assigning the ranks and degrees would cause significant distortion: they would literally multiply. Yager's method avoids this amplification of errors: values remain the same and are merely compared to one another. Thus there is a solid practical reason, not methodological purism, for using this method.

7. Outlook

The question I have used to demonstrate Yager's decision making procedure, "Which of several highly-paid soccer players should get a sports car?" is not particularly important. But the procedure is no different if it concerns not a car but a heart for transplantation, desperately needed by several patients. Criteria such as tissue toleration, waiting period, or presence of crisis might play a role and be fulfilled to varying degrees. One doesn't begin general and basic discussions, however, with life and death subjects.

8. Appendix

So let's have a look at a less ambitious legal case. That national soccer game I discussed had a parallel incident in the clubhouse of a fan club, where a group of young people were sitting in front of a large TV-set. Immediately after the victory goal, the TV-set shattered and some of the viewers were hit by splinters from the screen; one nearly bled to death since his artery was cut. The day before a TV-mechanic had replaced the vacuum tube of the TV-set incorrectly.
This is a case of negligently caused bodily harm, of course (§ 230 Strafgesetzbuch). But the German criminal code has still another section which might be relevant, § 311: "He who produces an explosion and thus endangers body and life of another person ..." An explosion? Literally speaking it was an implosion. Can we apply the statute all the same ?

There are three main criteria of statutory interpretation, which are often ranked as follows [9]:
(a) Literal interpretation, according to the letter of the statute
(b) Subjective interpretation, according to the intention of the legislator
(c) Objective interpretation, according to the aim and object of the statute

Many authors doubt the possibility of ranking these criteria, but this could be due to the fact that they have no rational technique to cope with the phenomenon of a high criterion fulfilled to a low degree and a low criterion fulfilled to a high degree. (The Anglo-Saxon doctrine ranks the objective before the subjective interpretation, or it totally denies the relevancy of the legislator's intentions; some German scholars follow them.)

Let us evaluate the case in the light of the three criteria:
(a) Literal interpretation: For a physicist there would be a fundamental difference between explosion and implosion. But a general penal code is not addressing

itself to physicists, but rather to the man in the street. E. g. if someone from the soccer fan club would have spoken of an "explosion" he would not have been corrected by his peers. (They would have corrected him, no doubt, if he spoke of a Ferrari instead of a Porsche.) So "exploding" may be not the perfect term for describing a bursting TV-set, but it is not inappropriate either. We could grade it as "highly acceptable" - or, more conservatively, which I would prefer - as "acceptable" (medium): 2.

(b) Subjective interpretation: The legal drafters of § 311 were not unaware of the possibility of an implosion. The section had a predecessor: there exists a draft of a proposed criminal code (vintage 1962) that had never become a statute, but for the first time (in the penal code itself) provided for the punishment of someone who produces a dangerous explosion. The official explanation of the draft was that phenomena such as implosions should also be considered.. The drafters of § 311 did not dissociate themselves from this interpretation. So we can assume that they shared it. Let us grade the subjective interpretation not as cogent (4) but as very plausible: 3.

(c) Objective interpretation: Whether an object auses an explosion or an implosion, a bursting occurs which makes no difference to a potential victim. If there is no difference in effect there is no reason for differently evaluating the causes. So we have a strong reason to grade the objective interpretation as cogent: 4.

Since the fulfilment grade of every single criterion surpasses the corresponding thresholds (0, 1, 2), the literal interpretation, which received the minimum grade, determines the score: 2. This is exactly the middle of the scale and does not look too impressive: as if it were a touch and go whether to apply the statute. But with Yager's method everything is comparison: so we have to carry out a cross-check and to experimentally take an opposing course of not to apply the statute. Now the grades of fulfilment are (2, 1, 2) and the subjective interpretation has the lowest value (identical with the height of the threshold of the criterion's complement) and determines the score: thus the bias is definitely towards the decision to apply the norm.

Every judge can immediately use Yager's simple method; no mathematical skills are needed. The law is full of values, principles, criteria, which are already ordered in a commonly accepted way or are waiting to be ordered. Computers could store the rankings and decisions. That would be a relatively modest task for a computer, more in the line of a database than of an expert system. But since law is something organic, historically grown, the use of computers in law should resemble rather the use of a machine in a garden than in a factory.

Notes:

1) Cf. B. Schlink, Abwägung im Verfassungsrecht, Berlin 1976, p. 131 ff.

2) At least if it does not contain too many steps. I have therefore used a five-step scale. Yager uses a seven-step scale that includes the steps "very little" and "very high". A seven-step scale has the advantage of being the most differentiated scale that can be grasped as a whole - as psychologist assure us. (The grading scale in German legal state exams includes seven nominal steps: from "insufficient" to "excellent", which are - with the sole exception of "insufficient" - numerically divided in groups of three.) For most legal applications a five-step scale seems preferable to me. In law instances of more than three or four aspects occurring at the same time are rare (with the possible exception of criminal sentencing). Besides, the reluctance of the courts and legal writers to give a definite ranking will grow with the number of steps in a scale, since the number of possible combinations will grow rapidly (n!). Not every constitutional law practitioner will be as decisive as Rosenberg in John Grisham's The Pelican Brief (London, 1992): "His ideology was simple; government over business, the individual over government, the environment over everything. And the Indians, give them whatever they want."

3) With reference to civil rights as a system of values this is also the criticism of R. Alexy (Theorie der Grundrechte, Frankfurt a.M. 1986, p. 138 ff.): cardinal scales are unrealistic since they are too demanding, and ordinal scales lead to a "tyranny" of the top value since the degrees of injury cannot be taken into account. Thus Alexy insists on individually evaluating a legal case - at least as long as there is no model of a comparable case.
The technique of comparing cases according to their strength is highly elaborated by the well-known HYPO conception; e.g. see K.D. Ashley, Modeling Legal Argument: Reasoning with Cases and Hypotheticals, Cambridge, MA, 1990; D. B. Skalak and E. L. Rissland, Arguments and Cases: A Inevitable Intertwining, Artificial Intelligence and Law 1 (1992) pp. 3-44.
From German writing an analysis of sit-down blockades by L. Kuhlen is worth mentioning: Regel und Fall in der juristischen Methodenlehre, ARSP-Beiheft 45 (1992), pp. 101-128.

4) On the subject of the Yager method, see R. R. Yager, Concepts, Theory, and Techniques: A New Methodology for Ordinal Multiobjective Decisions Based on Fuzzy Sets, Decision Sciences 12 (4), Oct. 1981, 598-600; and M. Caudill, Fuzzy Decisions, AI Expert, April 1990, pp. 59-64. Some criticism of Yager´s decision procedures based on ordinal scales is offered by H. Rommelfanger, Fuzzy Decision-Support-Systeme - Entscheiden bei Unschärfe, 2nd ed., Berlin-Heidelberg 1993, p. 144: "grobes Raster." The beginner in decision theory should start by reading Caudill's text.

5) An introduction to the fuzzy logic way of thinking rather than techniques can be found in: Bart Kosko, FUZZY THINKING. The New Science of Fuzzy Logic, New York 1993. One might love or hate that very personal book - in any case it is most worthwhile. More literature can be found in the following articles on legal application of fuzzy logic: L. Philipps, Unbestimmte Rechtsbegriffe und Fuzzy Logic, Festschrift für Arthur Kaufmann, Fr. Haft, W. Hassemer, U. Neuman, W. Schild, U. Schroth (eds.), Heidelberg 1993, pp. 265-280; the English version: Vague Legal Concepts and Fuzzy Logic. An Attempt to Determine the Required Period of Waiting after Traffic Accidents, INFORMATICA E DIRITTO vol. 2 (1993) pp. 37-51; L. Philipps, Kompensatorische Verknüpfungen in der Rechtsanwendung - ein Fall für Fuzzy Logic, Festschrift für Günther Jahr, M. Martinek, J. Schmidt, E Wadle (eds.), Tübingen 1993, pp. 169-180; J. Heithecker, Fuzzy Logic und der "Tierhalter", KI 1993, pp. 7-10; L. Philipps, Ein bißchen Fuzzy Logic für Juristen, Institutionen und Einzelne im Zeitalter der Informationsthechnik, M.-T. Tinnefeld, L. Philipps, K. Weis (eds.), München 1994, pp. 219-224; Kl. Köhler / J. Laeverenz, Moderne Technologien und das Haftungsrisiko des Arbeitnehmers. Ein Fuzzy-Logic-Expertensystem zur Ermittlung des Haftungsanteils, Institutionen und Einzelne im Zeitalter der Informationsthechnik (s.above), pp. 225-248; L. Philipps, Eine Theorie der unscharfen Subsumtion - Die Subsumtionsschwelle im Lichte der Fuzzy Logic, will soon be published in the Archiv für Rechts- und Sozialphilosophie (ARSP).

6) I also experimented with other variations of the conditional, but the results did not convince me. Cf. Th. Tilli, Fuzzy-Logik, Munich 1991, p. 167 ff. However, the following complementary formulas, which are inspired by the "Gödel Implication". could be worth mentioning:
(a) $V_a(x,y) = 0$ if $x' > y$, else y
(b) $V_b(x,y) = 1$ if $y \geq x'$, else y
In (a) all instances that do not qualify are indiscriminately excluded; as for the winners, a fine tuning is done.
In (b) all that qualify are indiscriminately accepted; if necessary, a fine tuning is done on the losers' side.

7) Other forms of AND and OR connections are used in fuzzy logic, too. However those require multiplication and are thus not useable with ordinal scales.

8) I suspect another problem at this point: There are decisions in which one might be ready to accept a low degree of fulfilment for other compensation, but not that the criterion is totally missing. (This was not the case in our example.) Yager would - if the example in his article (annotation 4) can be taken to be personal - buy a car of lower comfort because the price and fuel consumption are more important to him. But would this still be true if the car's comfort would be literally "null"? Yager's decision-making method might suggest that the difference between little or no comfort would disappear when set beside a medium level criterion. Perhaps the reasonable solution would be to include criteria that must not be completely absent in the description of the decision situation rather than try to represent them by improving the formula.

9) Cf. H.-J. Koch / H. Rüßmann, Juristische Begründungslehre, Munich 1982, p. 176 ff. ; K. Larenz, Methodenlehre der Rechtswissenschaft, 6th ed., Berlin Heidelberg 1991, chap. 5, 2.