Decision Making under Uncertainty: Algorithmic Approach (brief overview of related UTEP research)

Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
vladik@utep.edu

http://www.cs.utep.edu/vladik

1. Quantitative Approach to Decision Making: Misunderstandings

- Researchers and practitioners in computer science usually start with the *utility-based* approach.
- Many humanities researchers believe that the utility-based approach is oversimplified and long *discredited*.
- Main reason: they consider an easy-to-dismiss *caricature* instead of the actual utility approach.
- In view of this widely spread misunderstanding, we first start by explaining the *actual* utility-based approach.
- Our main area of research is how to add *uncertainty* to the traditional approach.
- We concentrate on *interval* and *fuzzy* uncert., emphasizing that "fuzzy" has a very precise meaning in CS.
- In this process, we provide examples of *applications*.

2. Decision Making: General Need and Traditional Approach

- To make a decision, we must:
  - find out the user's preference, and
  - help the user select an alternative which is the best – according to these preferences.
- Traditional approach is based on an assumption that for each two alternatives A' and A'', a user can tell:
  - whether the first alternative is better for him/her; we will denote this by $A'' < A'$;
  - or the second alternative is better; we will denote this by $A' < A''$;
  - or the two given alternatives are of equal value to the user; we will denote this by $A' = A''$.

3. The Notion of Utility

- Under the above assumption, we can form a natural numerical scale for describing preferences.
- Let us select a very bad alternative A_0 and a very good alternative A_1.
- Then, most other alternatives are better than A_0 but worse than A_1.
- For every prob. $p \in [0,1]$, we can form a lottery $L(p)$ in which we get A_1 w/prob. p and A_0 w/prob. $1-p$.
- When $p = 0$, this lottery simply coincides with the alternative A_0: $L(0) = A_0$.
- The larger the probability p of the positive outcome increases, the better the result:

$$p' < p'' \text{ implies } L(p') < L(p'').$$

4. The Notion of Utility (cont-d)

- Finally, for $p = 1$, the lottery coincides with the alternative A_1: $L(1) = A_1$.
- Thus, we have a continuous scale of alternatives $L(p)$ that monotonically goes from $L(0) = A_0$ to $L(1) = A_1$.
- Due to monotonicity, when p increases, we first have $L(p) < A$, then we have $L(p) > A$.
- The threshold value is called the *utility* of the alternative A:

$$u(A) \stackrel{\text{def}}{=} \sup\{p : L(p) < A\} = \inf\{p : L(p) > A\}.$$

- Then, for every $\varepsilon > 0$, we have

$$L(u(A) - \varepsilon) < A < L(u(A) + \varepsilon).$$

- We will describe such (almost) equivalence by $\equiv$, i.e., we will write that $A \equiv L(u(A))$.

5. Fast Iterative Process for Determining $u(A)$

- *Initially:* we know the values $\underline{u} = 0$ and $\overline{u} = 1$ such that $A \equiv L(u(A))$ for some $u(A) \in [\underline{u}, \overline{u}]$.
- *What we do:* we compute the midpoint u_{mid} of the interval $[\underline{u}, \overline{u}]$ and compare A with $L(u_{\text{mid}})$.
- *Possibilities:* $A \le L(u_{\text{mid}})$ and $L(u_{\text{mid}}) \le A$.
- *Case 1:* if $A \le L(u_{\text{mid}})$, then $u(A) \le u_{\text{mid}}$, so
$$u \in [\underline{u}, u_{\text{mid}}].$$
- *Case 2:* if $L(u_{\text{mid}}) \le A$, then $u_{\text{mid}} \le u(A)$, so
$$u \in [u_{\text{mid}}, \overline{u}].$$
- After each iteration, we decrease the width of the interval $[\underline{u}, \overline{u}]$ by half.
- After k iterations, we get an interval of width 2^{-k} which contains $u(A)$ – i.e., we get $u(A)$ w/accuracy 2^{-k}.

6. How to Make a Decision Based on Utility Values

- Suppose that we have found the utilities $u(A')$, $u(A'')$, …, of the alternatives A', A'', …
- Which of these alternatives should we choose?
- By definition of utility, we have:
  - $A \equiv L(u(A))$ for every alternative A, and
  - $L(p') < L(p'')$ if and only if $p' < p''$.
- We can thus conclude that A' is preferable to A'' if and only if $u(A') > u(A'')$.
- In other words, we should always select an alternative with the largest possible value of utility.
- So, to find the best solution, we must solve the corresponding optimization problem.

7. Before We Go Further: Caution

- We are *not* claiming that people estimate probabilities when they make decisions: we know they often don't.
- *Our claim:* when people make *definite* and *consistent* choices, these choices *can* be described by probabilities.
- *Example:* a falling rock does not solve equations but follows Newton's equations $ma = m\frac{d^2x}{dt^2} = -mg$.
- In practice, decisions are often *not* definite (uncertain) and *not* consistent.
- *Inconsistency* is one of the reasons why people make bad decisions (drugs, health hazards, speeding).
- People do choose $A > B > C > A$; we need *psychologists* and *sociologists* to study and solve this problem.
- *Uncertainty* is what we concentrate on; see below.

8. How to Estimate Utility of an Action

- For each action, we usually know possible outcomes $S_1, \ldots, S_n$.
- We can often estimate the prob. $p_1, \ldots, p_n$ of these outcomes.
- By definition of utility, each situation S_i is equiv. to a lottery $L(u(S_i))$ in which we get:
  - A_1 with probability $u(S_i)$ and
  - A_0 with the remaining probability $1 - u(S_i)$.
- Thus, the action is equivalent to a complex lottery in which:
  - first, we select one of the situations S_i with probability p_i: $P(S_i) = p_i$;
  - then, depending on S_i, we get A_1 with probability $P(A_1 \mid S_i) = u(S_i)$ and A_0 w/probability $1 - u(S_i)$.

9. How to Estimate Utility of an Action (cont-d)

- *Reminder:*
  - first, we select one of the situations S_i with probability p_i: $P(S_i) = p_i$;
  - then, depending on S_i, we get A_1 with probability $P(A_1 \mid S_i) = u(S_i)$ and A_0 w/probability $1 - u(S_i)$.
- The prob. of getting A_1 in this complex lottery is:

$$P(A_1) = \sum_{i=1}^{n} P(A_1 \mid S_i) \cdot P(S_i) = \sum_{i=1}^{n} u(S_i) \cdot p_i.$$

- In the complex lottery, we get:
  - A_1 with prob. $u = \sum\limits_{i=1}^{n} p_i \cdot u(S_i)$, and
  - A_0 w/prob. $1 - u$.
- So, we should select the action with the largest value of expected utility $u = \sum p_i \cdot u(S_i)$.

10. Subjective Probabilities

- In practice, we often do not know the probabilities p_i of different outcomes.
- For each event E, a natural way to estimate its subjective probability is to fix a prize (e.g., \$1) and compare:
  - the lottery ℓ_E in which we get the fixed prize if the event E occurs and 0 is it does not occur, with
  - a lottery $\ell(p)$ in which we get the same amount with probability p.
- Here, similarly to the utility case, we get a value $ps(E)$ for which, for every $\varepsilon > 0$:

$$\ell(ps(E) - \varepsilon) < \ell_E < \ell(ps(E) + \varepsilon).$$

- Then, the utility of an action with possible outcomes $S_1, \ldots, S_n$ is equal to $u = \sum\limits_{i=1}^{n} ps(E_i) \cdot u(S_i)$.

11. Auxiliary Issue: Almost-Uniqueness of Utility

- The above definition of utility u depends on A_0, A_1.
- What if we use different alternatives A'_0 and A'_1?
- Every A is equivalent to a lottery $L(u(A))$ in which we get A_1 w/prob. $u(A)$ and A_0 w/prob. $1-u(A)$.
- For simplicity, let us assume that $A'_0 < A_0 < A_1 < A'_1$.
- Then, $A_0 \equiv L'(u'(A_0))$ and $A_1 \equiv L'(u'(A_1))$.
- So, A is equivalent to a complex lottery in which:
  1) we select A_1 w/prob. $u(A)$ and A_0 w/prob. $1-u(A)$;
  2) depending on A_i, we get A'_1 w/prob. $u'(A_i)$ and A'_0 w/prob. $1-u'(A_i)$.
- In this complex lottery, we get A'_1 with probability $u'(A) = u(A)\cdot(u'(A_1)-u'(A_0))+u'(A_0)$.
- So, in general, utility is defined modulo an (increasing) linear transformation $u' = a\cdot u + b$, with $a > 0$.

12. Traditional Approach Summarized

- Traditional approach summarized:
  - we assume that we know possible actions, and
  - we assume that we know the exact consequences of each action;
  - then we should select an action with the largest value of expected utility.
- Similarly, when we have several participants:
  - we assume that we know the preferences of each participant,
  - then game theory provides us with reasonable solutions:
    - maximin for zero-sum games,
    - Nash bargaining solution, Nash equilibrium, or Shapley vector for cooperative games, etc.

13. Traditional Approach: Algorithmic Challenges

- In all these cases, we have a *well-defined* mathematical *problem* (e.g., an *optimization* problem).
- *Problem:* the existing algorithms run *too long* when the number of parameters increase.
- The first algorithmic challenge is to find *feasible* algorithms for solving these problems.
- *Case study:* security-related problems:
  - assigning air marshals to flights,
  - assigning security personnel to airport terminals, etc.
- Mathematically, solutions are known, but for thousands of flight, existing algorithms are inadequate.
- For these problems, Chris Kiekintveld developed new efficient algorithms, used by Homeland Security.

14. Need for Distributed Decision Making and the Resulting Algorithmic Challenges

- *Traditional approach:* we have a central decision maker.
- *In practice:* decisions are often made locally.
- *Challenge:* to operate efficiently, a distributed system needs a stable self-healing self-adjusting control.
- *Example:* Internet became possible only when Transmission Control Protocol (TCP) was invented.
- *Research direction* (E. Freudenthal): develop similar solutions for other systems.
- *Example 1:* transfer of medical information from patient-side sensors to patient-monitoring systems.
- *Example 2:* peer-to-peer communications, how to make sure that everyone contributes.

15. Need to Take Uncertainty into Account

- In the traditional approach, we assume that:
  - we know exactly which actions are possible,
  - we know the exact preferences of each participant,
  - we know the exact consequences of each action.
- Then, we have a constraint optimization problem.
- In reality:
  - we may not know exactly which actions are possible (i.e., we have "soft" constraints);
  - we only have partial information about the preferences; and
  - we only have partial information about consequences of each action.
- In this case, we face a problem of optimization and decision making under uncertainty.

16. Types of Uncertainty

- Ideally, for each quantity, we need to know:
  - which values are possible, and
  - how frequent are different possible values.
- So ideally, we should have *probabilistic* uncertainty.
- Sometimes, we only know the range $[\underline{x}, \overline{x}]$ of possible values; in this case, we have *interval* uncertainty.
- Sometimes, we also know narrower bounds $[\underline{x}(\alpha), \overline{x}(\alpha)]$ valid with some degree of certainty α.
- Such family of nested intervals is known as a *fuzzy set.*
- The degree of certainty can be described, e.g., by a Likert scale.
- Sometimes, we also know a *range* $[\underline{p}, \overline{p}]$ of probabilities p (or of mean or variance).

17. Privacy-Motivated Additional Uncertainty

- *Problem:* we often do not know what causes different diseases, which treatment is most efficient.
- *Solution:* collect data about patients, look for patterns.
- *Specifics:* since we do not know a priori which patterns to look for, we need to try various hypotheses.
- *Problem:* if we allow arbitrary queries, we may be able to reveal individual records – thus violating privacy.
- *Example:* how far influence from Asarco?
- We try average until 1001 Robinson and until 1003 Robinson, so we get the exact data re 1003 Robinson.
- *Solution:* instead of storing the original data, store *ranges*, e.g., for age, 0 to 10, 10 to 20, etc.
- *Challenge* (L. Longpré) we need to process data and make decisions under this interval uncertainty.

18. Uncertainty Leads to Soft Constraints: Toy Example

- *Objective:* come to school on time.
- *At first glance:* precisely formulated problem.
- *Fact:* traffic jams happen.
- *In rare cases:* traffic jams can be up to an hour long.
- *Guaranteed solution:* leave home an hour earlier.
- *Problem:* wasting an hour every day.
- *Solution:* realize that "on time" is a soft constraint.
- *Specifically:* it is OK to be late one day a year–when everyone is late due to a traffic jam.

19. Uncertainty Leads to Soft Constraints

- *Case study* (Martine Ceberio): researchers design an innovative water filtering system.
- *Objective:* minimize energy use.
- *Constraints:* lower bound on the output, and physics-based constraints relating parameters.
- *At first glance:* there is no uncertainty, all physics-motivated constraints seem exact.
- *Surprise:* the constraints turned out to be inconsistent.
- *Reason:* relations are approximate (similar to using 3.14 instead of π).
- *Solution:* relax constraints, i.e., replace equalities with approximate equalities.
- *Algorithmic challenge:* to simplify computations, we need to minimize the number of relaxed constraints.

20. Uncertainty in Objective Function: A Problem

- *General case:* utility depends on the parameters $x_1, \ldots, x_n$: $u = u(x_1, \ldots, x_n)$.
- *First approximation:* assume that the dependence is linear $u = \sum_{i=1}^{n} c_i \cdot x_i$.
- *In practice:* linear dependencies are usually only approximate ones.
- *Seemingly natural idea:* add quadratic (and higher order) terms $u = \sum_{i=1}^{n} c_i \cdot x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \cdot x_i \cdot x_j$.
- *Fact:* the situation is often scale-invariant.
- *Example:* x_i are money, and preferences should not change if we use not dollars but Euros.
- *Problem:* quadratic preferences are not scale-invariant.

21. Uncertainty in Objective Function Leads to Non-Additive (Fuzzy) Measures

- *Problem* (reminder): quadratic preferences are not scale-invariant.

- *First idea:* use scale-invariant ordinal statistics

$$x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)},$$

$$x_{(1)} = \min(x_1, \ldots, x_n), \ldots, x_{(1)} = \max(x_1, \ldots, x_n).$$

- *Resulting solution:* take $u = \sum\limits_{i=1}^{n} c_i \cdot x_{(i)}$.

- *General scale-invariant expression:* can be described as an integral over a non-additive ("fuzzy") measure.

- *Successful case study* (M. Ceberio, X. Wang): how to describe software quality.

- *Result:* fuzzy measure-based approach better describes expert preferences.

22. Uncertainty in System Dynamics: Interval-Related Approach

- *Traditional approach:* dynamics is described by differential equations, like Newton's equations

$$\frac{d^2x}{dt^2} = \frac{F}{m}.$$

- *Fact:* usually, we do not know the exact equations

$$\dot{x} = f(x).$$

- *Possibility:* we only know the approximate equations, i.e., we know the ranges $\left[\underline{f}(x), \overline{f}(x)\right]$ for which

$$\dot{x} \in \left[\underline{f}(x), \overline{f}(x)\right].$$

- *Solution* (B. Djafari-Rouhani): analyze such differential inequalities.

23. Uncertainty in System Dynamics: Symmetry Approach

- *One of the main objectives of science:* prediction.
- *Basis for prediction:* we observed *similar* situations in the past, and we expect similar outcomes.
- *In mathematical terms:* similarity corresponds to *symmetry*, and similarity of outcomes – to *invariance.*
- *Example:* we dropped the ball, it fall down.
- *Symmetries:* shift, rotation, etc.
- Symmetries are ubiquitous in *modern physics*:
  - starting with quarks, new theories are represented in terms of symmetries;
  - traditional physical theories (GRT, QM, Electrodynamics, etc.) can be described in symmetry terms.

24. Beyond Traditional Decision Making: Towards a More Realistic Description

- Previously, we assumed that a user can always decide which of the two alternatives A' and A'' is better:
  - either $A' < A''$,
  - or $A'' < A'$,
  - or $A' \equiv A''$.
- In practice, a user is sometimes unable to meaningfully decide between the two alternatives; denoted $A' \parallel A''$.
- In mathematical terms, this means that the preference relation:
  - is no longer a *total* (linear) order,
  - it can be a *partial* order.

25. From Utility to Interval-Valued Utility

- Similarly to the traditional decision making approach:
  - we select two alternatives $A_0 < A_1$ and
  - we compare each alternative A which is better than A_0 and worse than A_1 with lotteries $L(p)$.
- Since preference is a *partial* order, in general:

$$\underline{u}(A) \stackrel{\text{def}}{=} \sup\{p : L(p) < A\} < \overline{u}(A) \stackrel{\text{def}}{=} \inf\{p : L(p) > A\}.$$

- For each alternative A, instead of a single value $u(A)$ of the utility, we now have an *interval* $[\underline{u}(A), \overline{u}(A)]$ s.t.:
  - if $p < \underline{u}(A)$, then $L(p) < A$;
  - if $p > \overline{u}(A)$, then $A < L(p)$; and
  - if $\underline{u}(A) < p < \overline{u}(A)$, then $A \parallel L(p)$.
- We will call this interval the *utility* of the alternative A.

26. Interval-Valued Utility: Practical Consequences

- *Idea:* select alternative A with largest $u(A)$.
- As situation changes, we may change our selection.
- *Interval case:* for each alternative, we know the utility with some uncertainty Δ, i.e., we know $\widetilde{u}(A)$ for which
$$u(A) \in [\widetilde{u}(A) - \Delta, \widetilde{u}(A) + \Delta].$$
- *Additional aspect:* there is usually a cost in change (e.g., a cost in reinvesting in different stocks).
- *Conclusion:* we only change from A to B if we are sure that $u(A) < u(B)$, i.e., when $\widetilde{u}(A) + \Delta < \widetilde{u}(B) + \Delta$.
- *Problem:* it is difficult to estimate Δ exactly.
- If we *underestimate* Δ, we make a lot of unnecessary changes ("mania").
- If we *overestimate* Δ, we miss good opportunities ("depression").

27. Symmetry Approach to decision Making Under Uncertainty: Examples

- What are the best locations of radiotelescopes forming a Very Large Baseline Interferometer (VLBI)?
- *Fact:* the optimal location depends on what objects we will observe.
- *Challenge:* we do not know what objects we will observe with the new VLBI system.
- *Environmental sciences:* what is the best location of a meteorological tower?
- *Fact:* the optimal location depends on subtle details of local weather patterns.
- *Challenge:* these patterns are exactly what we plan to determine with the new tower.
- In all these cases, *symmetry* helps.

Thanks for your attention!

28. Case Study

- *Objective:* select the best location of a sophisticated multi-sensor meteorological tower.
- *Constraints:* we have several criteria to satisfy.
- *Example:* the station should not be located too close to a road.
- *Motivation:* the gas flux generated by the cars do not influence our measurements of atmospheric fluxes.
- *Formalization:* the distance x_1 to the road should be larger than a threshold t_1: $x_1 > t_1$, or $y_1 \stackrel{\text{def}}{=} x_1 - t_1 > 0$.
- *Example:* the inclination x_2 at the tower's location should be smaller than a threshold t_2: $x_2 < t_2$.
- *Motivation:* otherwise, the flux determined by this inclination and not by atmospheric processes.

29. General Case

- *In general:* we have several differences $y_1, \ldots, y_n$ all of which have to be non-negative.
- For each of the differences y_i, the larger its value, the better.
- Our problem is a typical setting for *multi-criteria optimization*.
- A most widely used approach to multi-criteria optimization is *weighted average*, where
  - we assign weights $w_1, \ldots, w_n > 0$ to different criteria y_i and
  - select an alternative for which the weighted average

$$w_1 \cdot y_1 + \ldots + w_n \cdot y_n$$

attains the largest possible value.

30. Limitations of the Weighted Average Approach

- *In general:* the weighted average approach often leads to reasonable solutions of the multi-criteria problem.
- *In our problem:* we have an additional requirement – that all the values y_i must be positive. So:
  - when selecting an alternative with the largest possible value of the weighted average,
  - we must only compare solutions with $y_i > 0$.
- *We will show:* under the requirement $y_i > 0$, the weighted average approach is not fully satisfactory.
- *Conclusion:* we need to find a more adequate solution.

31. Limitations of the Weighted Average Approach: Details

- The values y_i come from measurements, and measurements are never absolutely accurate.
- The results $\widetilde{y}_i$ of the measurements are not exactly equal to the actual (unknown) values y_i.
- *If:* for some alternative $y = (y_1, \ldots, y_n)$
  - we measure the values y_i with higher and higher accuracy and,
  - based on the measurement results $\widetilde{y}_i$, we conclude that y is better than some other alternative y'.
- *Then:* we expect that the actual alternative y is indeed better than y' (or at least of the same quality).
- Otherwise, we will not be able to make any meaningful conclusions based on real-life measurements.

32. The Above Natural Requirement Is Not Always Satisfied for Weighted Average

- *Simplest case:* two criteria y_1 and y_2, w/weights $w_i > 0$.
- If $y_1, y_2, y'_1, y'_2 > 0$, and $w_1 \cdot y_1 + w_2 \cdot y_2 > w_1 \cdot y'_1 + w_2 \cdot y'_2$, then $y = (y_1, y_2) \succ y' = (y'_1, y'_2)$.
- If $y_1 > 0$, $y_2 > 0$, and at least one of the values y'_1 and y'_2 is non-positive, then $y = (y_1, y_2) \succ y' = (y'_1, y'_2)$.
- Let us consider, for every $\varepsilon > 0$, the tuple $y(\varepsilon) \stackrel{\text{def}}{=} (\varepsilon, 1 + w_1/w_2)$, and $y' = (1, 1)$.
- In this case, for every $\varepsilon > 0$, we have
$$w_1 \cdot y_1(\varepsilon) + w_2 \cdot y_2(\varepsilon) = w_1 \cdot \varepsilon + w_2 + w_2 \cdot \frac{w_1}{w_2} = w_1 \cdot (1 + \varepsilon) + w_2$$
and $w_1 \cdot y'_1 + w_2 \cdot y'_2 = w_1 + w_2$, hence $y(\varepsilon) \succ y'$.
- However, in the limit $\varepsilon \to 0$, we have $y(0) = \left(0, 1 + \dfrac{w_1}{w_2}\right)$, with $y(0)_1 = 0$ and thus, $y(0) \prec y'$.

33. Towards a Precise Description

- Each alternative is characterized by a tuple of n positive values $y = (y_1, \ldots, y_n)$.
- Thus, the set of all alternatives is the set $(R^+)^n$ of all the tuples of positive numbers.
- For each two alternatives y and y', we want to tell whether
  - y is better than y' (we will denote it by $y \succ y'$ or $y' \prec y$),
  - or y' is better than y ($y' \succ y$),
  - or y and y' are equally good ($y' \sim y$).
- *Natural requirement:* if y is better than y' and y' is better than y'', then y is better than y''.
- The relation $\succ$ must be transitive.

34. Towards a Precise Description (cont-d)

- *Reminder:* the relation $\succ$ must be transitive.
- Similarly, the relation $\sim$ must be transitive, symmetric, and reflexive ($y \sim y$), i.e., be an *equivalence relation.*
- *An alternative description:* a transitive *pre-ordering* relation $a \succeq b \Leftrightarrow (a \succ b \vee a \sim b)$ s.t. $a \succeq b \vee b \succeq a$.
- Then, $a \sim b \Leftrightarrow (a \succeq b) \,\&\, (b \succeq a)$, and

$$a \succ b \Leftrightarrow (a \succeq b) \,\&\, (b \not\succeq a).$$

- *Additional requirement:*
  - *if* each criterion is better,
  - *then* the alternative is better as well.
- *Formalization:* if $y_i > y'_i$ for all i, then $y \succ y'$.

35. Scale Invariance: Motivation

- *Fact:* quantities y_i describe completely different physical notions, measured in completely different units.
- *Examples:* wind velocities measured in m/s, km/h, mi/h; elevations in m, km, ft.
- Each of these quantities can be described in many different units.
- A priori, we do not know which units match each other.
- Units used for measuring different quantities may not be exactly matched.
- It is reasonable to require that:
  - if we simply change the units in which we measure each of the corresponding n quantities,
  - the relations $\succ$ and $\sim$ between the alternatives $y = (y_1, \ldots, y_n)$ and $y' = (y'_1, \ldots, y'_n)$ do not change.

36. Scale Invariance: Towards a Precise Description

- *Situation:* we replace:
    - a unit in which we measure a certain quantity q
    - by a new measuring unit which is $\lambda > 0$ times smaller.
- *Result:* the numerical values of this quantity increase by a factor of λ: $q \to \lambda \cdot q$.
- *Example:* 1 cm is $\lambda = 100$ times smaller than 1 m, so the length $q = 2$ becomes $\lambda \cdot q = 2 \cdot 100 = 200$ cm.
- Then, scale-invariance means that for all $y, y' \in (R^+)^n$ and for all $\lambda_i > 0$, we have
    - $y = (y_1, \ldots, y_n) \succ y' = (y'_1, \ldots, y'_n)$ implies $(\lambda_1 \cdot y_1, \ldots, \lambda_n \cdot y_n) \succ (\lambda_1 \cdot y'_1, \ldots, \lambda_n \cdot y'_n)$,
    - $y = (y_1, \ldots, y_n) \sim y' = (y'_1, \ldots, y'_n)$ implies $(\lambda_1 \cdot y_1, \ldots, \lambda_n \cdot y_n) \sim (\lambda_1 \cdot y'_1, \ldots, \lambda_n \cdot y'_n)$.

37. Formal Description

- By a *total pre-ordering relation* on a set Y, we mean
  - a pair of a transitive relation $\succ$ and an equivalence relation $\sim$ for which,
  - for every $y, y' \in Y$, exactly one of the following relations hold: $y \succ y'$, $y' \succ y$, or $y \sim y'$.
- We say that a total pre-ordering is *non-trivial* if there exist y and y' for which $y \succ y'$.
- We say that a total pre-ordering relation on $(R^+)^n$ is:
  - *monotonic* if $y'_i > y_i$ for all i implies $y' \succ y$;
  - *continuous* if
    - ∗ whenever we have a sequence $y^{(k)}$ of tuples for which $y^{(k)} \succeq y'$ for some tuple y', and
    - ∗ the sequence $y^{(k)}$ tends to a limit y,
    - ∗ then $y \succeq y'$.

38. Main Result

Theorem. *Every non-trivial monotonic scale-inv. continuous total pre-ordering relation on* $(R^+)^n$ *has the form:*

$$y' = (y'_1, \ldots, y'_n) \succ y = (y_1, \ldots, y_n) \Leftrightarrow \prod_{i=1}^{n} (y'_i)^{\alpha_i} > \prod_{i=1}^{n} y_i^{\alpha_i};$$

$$y' = (y'_1, \ldots, y'_n) \sim y = (y_1, \ldots, y_n) \Leftrightarrow \prod_{i=1}^{n} (y'_i)^{\alpha_i} = \prod_{i=1}^{n} y_i^{\alpha_i},$$

for some constants $\alpha_i > 0$.

Comment: Vice versa,

- for each set of values $\alpha_1 > 0, \ldots, \alpha_n > 0$,
- the above formulas define a monotonic scale-invariant continuous pre-ordering relation on $(R^+)^n$.

39. Practical Conclusion

- *Situation:*
  - we need to select an alternative;
  - each alternative is characterized by characteristics $y_1, \ldots, y_n$.
- *Traditional approach:*
  - we assign the weights w_i to different characteristics;
  - we select the alternative with the largest value of $\sum_{i=1}^{n} w_i \cdot y_i$.
- *New result:* it is better to select an alternative with the largest value of $\prod_{i=1}^{n} y_i^{w_i}$.
- *Equivalent reformulation:* select an alternative with the largest value of $\sum_{i=1}^{n} w_i \cdot \ln(y_i)$.

40. Multi-Agent Cooperative Decision Making

- *How to describe preferences:* for each participant P_i, we can determine the utility $u_{ij} \stackrel{\text{def}}{=} u_i(A_j)$ of all A_j.
- *Question:* how to transform these utilities into a reasonable group decision rule?
- *Solution:* was provided by another future Nobelist John Nash.
- *Nash's assumptions:*
  - symmetry,
  - independence from irrelevant alternatives, and
  - *scale invariance* – under replacing function $u_i(A)$ with an equivalent function $a \cdot u_i(A)$,

41. Nash's Bargaining Solution (cont-d)

- *Nash's assumptions (reminder):*
  - symmetry,
  - independence from irrelevant alternatives, and
  - scale invariance.
- *Nash's result:*
  - the only group decision rule satisfying all these assumptions
  - is selecting an alternative A for which the product $\prod\limits_{i=1}^{n} u_i(A)$ is the largest possible.
- *Comment.* the utility functions must be "scaled" s.t. the "status quo" situation $A^{(0)}$ has utility 0:

$$u_i(A) \to u_i'(A) \stackrel{\text{def}}{=} u_i(A) - u_i(A^{(0)}).$$

42. Interval-Valued Utilities and Interval-Valued Subjective Probabilities

- To feasibly elicit the values $\underline{u}(A)$ and $\overline{u}(A)$, we:
  1) starting w/$[\underline{u}, \overline{u}] = [0, 1]$, bisect an interval s.t. $L(\underline{u}) < A < L(\overline{u})$ until we find u_0 s.t. $A \parallel L(u_0)$;
  2) by bisecting an interval $[\underline{u}, u_0]$ for which $L(\underline{u}) < A \parallel L(u_0)$, we find $\underline{u}(A)$;
  3) by bisecting an interval $[u_0, \overline{u}]$ for which $L(u_0) \parallel A < L(\overline{u})$, we find $\overline{u}(A)$.
- Similarly, when we estimate the probability of an event E:
  – we no longer get a single value $ps(E)$;
  – we get an *interval* $\left[\underline{ps}(E), \overline{ps}(E)\right]$ of possible values of probability.
- By using bisection, we can feasibly elicit the values $\underline{ps}(E)$ and $\overline{ps}(E)$.

43. Decision Making Under Interval Uncertainty

- *Situation:* for each possible decision d, we know the interval $[\underline{u}(d), \overline{u}(d)]$ of possible values of utility.
- *Questions:* which decision shall we select?
- *Natural idea:* select all decisions d_0 that *may* be optimal, i.e., which are optimal for some function

$$u(d) \in [\underline{u}(d), \overline{u}(d)].$$

- *Problem:* checking all possible functions is not feasible.
- *Solution:* the above condition is equivalent to an easier-to-check one:

$$\overline{u}(d_0) \ge \max_d \underline{u}(d).$$

- *Interval computations* can help in describing the range of all such d_0.
- *Remaining problem:* in practice, we would like to select *one* decision; which one should be select?

44. Need for Definite Decision Making

- *At first glance:* if $A' \parallel A''$, it does not matter whether we recommend alternative A' or alternative A''.
- Let us show that this is *not* a good recommendation.
- E.g., let A be an alternative about which we know nothing, i.e., $[\underline{u}(A), \overline{u}(A)] = [0, 1]$.
- In this case, A is indistinguishable both from a "good" lottery $L(0.999)$ and a "bad" lottery $L(0.001)$.
- Suppose that we recommend, to the user, that A is equivalent both to $L(0.999)$ and to $L(0.001)$.
- Then this user will feel comfortable:
  - first, exchanging $L(0.999)$ with A, and
  - then, exchanging A with $L(0.001)$.
- So, following our recommendations, the user switches from a very good alternative to a very bad one.

45. Need for Definite Decision Making (cont-d)

- The above argument does not depend on the fact that we assumed complete ignorance about A:
  - every time we recommend that the alternative A is "equivalent" both to $L(p)$ and to $L(p')$ $(p < p')$,
  - we make the user vulnerable to a similar switch from a better alternative $L(p')$ to a worse one $L(p)$.
- Thus, there should be only a single value p for which A can be reasonably exchanged with $L(p)$.
- In precise terms:
  - we start with the utility interval $[\underline{u}(A), \overline{u}(A)]$, and
  - we need to select a single $u(A)$ for which it is reasonable to exchange A with a lottery $L(u)$.
- How can we find this value $u(A)$?

46. Decisions under Interval Uncertainty: Hurwicz Optimism-Pessimism Criterion

- *Reminder:* we need to assign, to each interval $[\underline{u}, \overline{u}]$, a utility value $u(\underline{u}, \overline{u}) \in [\underline{u}, \overline{u}]$.
- *History:* this problem was first handled in 1951, by the future Nobelist Leonid Hurwicz.
- *Notation:* let us denote $\alpha_H \stackrel{\text{def}}{=} u(0, 1)$.
- *Reminder:* utility is determined modulo a linear transformation $u' = a \cdot u + b$.
- *Reasonable to require:* the equivalent utility does not change with re-scaling: for $a > 0$ and b,

$$u(a \cdot u^- + b, a \cdot u^+ + b) = a \cdot u(u^-, u^+) + b.$$

- For $u^- = 0$, $u^+ = 1$, $a = \overline{u} - \underline{u}$, and $b = \underline{u}$, we get

$$u(\underline{u}, \overline{u}) = \alpha_H \cdot (\overline{u} - \underline{u}) + \underline{u} = \alpha_H \cdot \overline{u} + (1 - \alpha_H) \cdot \underline{u}.$$

47. Hurwicz Optimism-Pessimism Criterion (cont)

- The expression $\alpha_H \cdot \overline{u} + (1 - \alpha_H) \cdot \underline{u}$ is called *optimism-pessimism criterion*, because:
  - when $\alpha_H = 1$, we make a decision based on the most optimistic possible values $u = \overline{u}$;
  - when $\alpha_H = 0$, we make a decision based on the most pessimistic possible values $u = \underline{u}$;
  - for intermediate values $\alpha_H \in (0, 1)$, we take a weighted average of the optimistic and pessimistic values.
- According to this criterion:
  - if we have several alternatives $A', \ldots$, with interval-valued utilities $[\underline{u}(A'), \overline{u}(A')], \ldots$,
  - we recommend an alternative A that maximizes

$$\alpha_H \cdot \overline{u}(A) + (1 - \alpha_H) \cdot \underline{u}(A).$$

48. Which Value α_H Should We Choose? An Argument in Favor of $\alpha_H = 0.5$

- Let us take an event E about which we know nothing.
- For a lottery L^+ in which we get A_1 if E and A_0 otherwise, the utility interval is $[0, 1]$.
- Thus, the equiv. utility of L^+ is $\alpha_H \cdot 1 + (1 - \alpha_H) \cdot 0 = \alpha_H$.
- For a lottery L^- in which we get A_0 if E and A_1 otherwise, the utility is $[0, 1]$, so equiv. utility is also α_H.
- For a complex lottery L in which we select either L^+ or L^- with probability 0.5, the equiv. utility is still α_H.
- On the other hand, in L, we get A_1 with probability 0.5 and A_0 with probability 0.5.
- Thus, $L = L(0.5)$ and hence, $u(L) = 0.5$.
- So, we conclude that $\alpha_H = 0.5$.

49. Which Action Should We Choose?

- Suppose that an action has n possible outcomes $S_1, \ldots, S_n$, with utilities $[\underline{u}(S_i), \overline{u}(S_i)]$, and probabilities $[\underline{p}_i, \overline{p}_i]$.
- We know that each alternative is equivalent to a simple lottery with utility $u_i = \alpha_H \cdot \overline{u}(S_i) + (1 - \alpha_H) \cdot \underline{u}(S_i)$.
- We know that for each i, the i-th event is equivalent to $p_i = \alpha_H \cdot \overline{p}_i + (1 - \alpha_H) \cdot \underline{p}_i$.
- Thus, this action is equivalent to a situation in which we get utility u_i with probability p_i.
- The utility of such a situation is equal to $\sum\limits_{i=1}^{n} p_i \cdot u_i$.
- Thus, the equivalent utility of the original action is equivalent to

$$\sum_{i=1}^{n} \left(\alpha_H \cdot \overline{p}_i + (1 - \alpha_H) \cdot \underline{p}_i\right) \cdot (\alpha_H \cdot \overline{u}(S_i) + (1 - \alpha_H) \cdot \underline{u}(S_i)) .$$

50. Observation: the Resulting Decision Depends on the Level of Detail

- Let us consider a situation in which, with some prob. p, we gain a utility u, else we get 0.
- The expected utility is $p \cdot u + (1 - p) \cdot 0 = p \cdot u$.
- Suppose that we only know the intervals $[\underline{u}, \overline{u}]$ and $[\underline{p}, \overline{p}]$.
- The equivalent utility u_k (k for $k$now) is

$$u_k = (\alpha_H \cdot \overline{p} + (1 - \alpha_H) \cdot \underline{p}) \cdot (\alpha_H \cdot \overline{u} + (1 - \alpha_H) \cdot \underline{u}).$$

- If we only know that utility is from $[\underline{p} \cdot \underline{u}, \overline{p} \cdot \overline{u}]$, then:

$$u_d = \alpha_H \cdot \overline{p} \cdot \overline{u} + (1 - \alpha_H) \cdot \underline{p} \cdot \underline{u} \text{ (d for $d$on't know).}$$

- Here, additional knowledge decreases utility:

$$u_d - u_k = \alpha_H \cdot (1 - \alpha_H) \cdot (\overline{p} - \underline{p}) \cdot (\overline{u} - \underline{u}) > 0.$$

- (This is maybe what the Book of Ecclesiastes meant by "For with much wisdom comes much sorrow"?)

51. Beyond Interval Uncertainty: Partial Info about Probabilities

- *Frequent situation:*
  - in addition to $\mathbf{x}_i$,
  - we may also have *partial* information about the probabilities of different values $x_i \in \mathbf{x}_i$.
- An *exact* probability distribution can be described, e.g., by its cumulative distribution function

$$F_i(z) = \text{Prob}(x_i \le z).$$

- A *partial* information means that instead of a single cdf, we have a *class* $\mathcal{F}$ of possible cdfs.
- *p-box* (Scott Ferson):
  - for every z, we know an interval $\mathbf{F}(z) = [\underline{F}(z), \overline{F}(z)]$;
  - we consider all possible distributions for which, for all z, we have $F(z) \in \mathbf{F}(z)$.

52. Describing Partial Info about Probabilities: Decision Making Viewpoint

- *Problem:* there are many ways to represent a probability distribution.
- *Idea:* look for an objective.
- *Objective:* make decisions $E_x[u(x,a)] \to \max\limits_a$.
- *Case 1:* smooth $u(x)$.
- *Analysis:* we have $u(x) = u(x_0) + (x - x_0) \cdot u'(x_0) + \ldots$
- *Conclusion:* we must know moments to estimate $E[u]$.
- *Case of uncertainty:* interval bounds on moments.
- *Case 2:* threshold-type $u(x)$ (e.g., regulations).
- *Conclusion:* we need cdf $F(x) = \text{Prob}(\xi \le x)$.
- *Case of uncertainty:* p-box $[\underline{F}(x), \overline{F}(x)]$.

53. Multi-Agent Decision Making under Interval Uncertainty

- *Reminder:* if we set utility of status quo to 0, then we select an alternative A that maximizes
$$u(A)=\prod_{i=1}^n u_i(A).$$
- *Case of interval uncertainty:* we only know intervals $[\underline{u}_i(A),\overline{u}_i(A)]$.
- *First idea:* find all A_0 for which $\overline{u}(A_0)\ge \max\limits_A \underline{u}(A)$, where
$$[\underline{u}(A),\overline{u}(A)]\stackrel{\text{def}}{=}\prod_{i=1}^n [\underline{u}_i(A),\overline{u}_i(A)].$$
- *Second idea:* maximize $u^{\text{equiv}}(A)\stackrel{\text{def}}{=}\prod\limits_{i=1}^n u_i^{\text{equiv}}(A)$.
- *Interesting aspect:* when we have a conflict situation (e.g., in security games).

54. Beyond Optimization

- *Traditional interval computations:*
  - we know the intervals $X_1, \ldots, X_n$ containing $x_1, \ldots, x_n$;
  - we know that a quantity z depends on $x = (x_1, \ldots, x_n)$:

$$z = f(x_1, \ldots, x_n);$$

  - we want to find the range Z of possible values of z:

$$Z = \left[\min_{x \in X} f(x), \max_{x \in X} f(x)\right].$$

- *Control situations:*
  - the value $z = f(x, u)$ also depends on the control variables $u = (u_1, \ldots, u_m)$;
  - we want to find Z for which, for every $x_i \in X_i$, we can get $z \in Z$ by selecting appropriate $u_j \in U_j$:

$$\forall x\, \exists u\, (z = f(x, u) \in Z).$$

55. Reformulation in Logical Terms – of Modal Intervals

- *Reminder:* we want $\forall x_{\in X}\, \exists u_{\in U}\, (f(x,u) \in Z)$.
- There is a logical difference between intervals X and U.
- The property $f(x,u) \in Z$ must hold
  - *for all* possible values $x_i \in X_i$, but
  - *for some* values $u_j \in U_j$.
- We can thus consider pairs of intervals and quantifiers (*modal intervals*):
  - each original interval X_i is a pair $\langle X_i, \forall\rangle$, while
  - controlled interval is a pair $\langle U_j, \exists\rangle$.
- We can treat the resulting interval Z as the range defined over modal intervals:

$$Z = f(\langle X_1, \forall\rangle, \ldots, \langle X_n, \forall\rangle, \langle U_1, \exists\rangle, \ldots, \langle U_m, \exists\rangle).$$

56. Even Further Beyond Optimization

- In more complex situations, we need to go beyond control.
- For example, in the presence of an adversary, we want to make a decision x such that:
  - for every possible reaction y of an adversary,
  - we will be able to make a next decision x' (depending on y)
  - so that after every possible next decision y' of an adversary,
  - the resulting state $s(x, y, x', y')$ will be in the desired set:

$$\forall y\, \exists x'\, \forall y'\, (s(x, y, x', y') \in S).$$

- In this case, we arrive at general Shary's classes.

57. Proof of Symmetry Result: Part 1

- Due to scale-invariance, for every $y_1, \ldots, y_n, y'_1, \ldots, y'_n$, we can take $\lambda_i = \dfrac{1}{y_i}$ and conclude that

$$(y'_1, \ldots, y'_n) \sim (y_1, \ldots, y_n) \Leftrightarrow \left(\frac{y'_1}{y_1}, \ldots, \frac{y'_n}{y_n}\right) \sim (1, \ldots, 1).$$

- Thus, to describe the equivalence relation $\sim$, it is sufficient to describe $\{z = (z_1, \ldots, z_n) : z \sim (1, \ldots, 1)\}$.
- Similarly,

$$(y'_1, \ldots, y'_n) \succ (y_1, \ldots, y_n) \Leftrightarrow \left(\frac{y'_1}{y_1}, \ldots, \frac{y'_n}{y_n}\right) \succ (1, \ldots, 1).$$

- Thus, to describe the ordering relation $\succ$, it is sufficient to describe the set $\{z = (z_1, \ldots, z_n) : z \succ (1, \ldots, 1)\}$.
- Similarly, it is also sufficient to describe the set

$$\{z = (z_1, \ldots, z_n) : (1, \ldots, 1) \succ z\}.$$

58. Proof of Symmetry Result: Part 2

- *To simplify:* take logarithms $Y_i = \ln(y_i)$, and sets

$$S_\sim = \{Z : z = (\exp(Z_1), \ldots, \exp(Z_n)) \sim (1, \ldots, 1)\},$$
$$S_\succ = \{Z : z = (\exp(Z_1), \ldots, \exp(Z_n)) \succ (1, \ldots, 1)\};$$
$$S_\prec = \{Z : (1, \ldots, 1) \succ z = (\exp(Z_1), \ldots, \exp(Z_n))\}.$$

- Since the pre-ordering relation is total, for Z, either $Z \in S_\sim$ or $Z \in S_\succ$ or $Z \in S_\prec$.
- *Lemma:* $S_\sim$ is closed under addition:
  - $Z \in S_\sim$ means $(\exp(Z_1), \ldots, \exp(Z_n)) \sim (1, \ldots, 1)$;
  - due to scale-invariance, we have

$$(\exp(Z_1+Z_1'), \ldots) = (\exp(Z_1)\cdot\exp(Z_1'), \ldots) \sim (\exp(Z_1'), \ldots);$$

  - also, $Z' \in S_\sim$ means $(\exp(Z_1'), \ldots) \sim (1, \ldots, 1)$;
  - since $\sim$ is transitive,

$$(\exp(Z_1 + Z_1'), \ldots) \sim (1, \ldots) \text{ so } Z + Z' \in S_\sim.$$

59. Proof of Symmetry Result: Part 3

- *Reminder:* the set $S_\sim$ is closed under addition;
- Similarly, $S_\prec$ and $S_\succ$ are closed under addition.
- *Conclusion:* for every integer $q > 0$:
  - if $Z \in S_\sim$, then $q \cdot Z \in S_\sim$;
  - if $Z \in S_\succ$, then $q \cdot Z \in S_\succ$;
  - if $Z \in S_\prec$, then $q \cdot Z \in S_\prec$.
- Thus, if $Z \in S_\sim$ and $q \in N$, then $(1/q) \cdot Z \in S_\sim$.
- We can also prove that $S_\sim$ is closed under $Z \to -Z$:
  - $Z = (Z_1, \ldots) \in S_\sim$ means $(\exp(Z_1), \ldots) \sim (1, \ldots)$;
  - by scale invariance, $(1, \ldots) \sim (\exp(-Z_1), \ldots)$, i.e., $-Z \in S_\sim$.
- Similarly, $Z \in S_\succ \Leftrightarrow -Z \in S_\prec$.
- So $Z \in S_\sim \Rightarrow (p/q) \cdot Z \in S_\sim$; in the limit, $x \cdot Z \in S_\sim$.

60. Proof of Symmetry Result: Final Part

- *Reminder:* $S_\sim$ is closed under addition and multiplication by a scalar, so it is a linear space.
- *Fact:* $S_\sim$ cannot have full dimension n, since then all alternatives will be equivalent to each other.
- *Fact:* $S_\sim$ cannot have dimension $< n-1$, since then:
  - we can select an arbitrary $Z \in S_\prec$;
  - connect it w/$-Z \in S_\succ$ by a path γ that avoids $S_\sim$;
  - due to closeness, $\exists \gamma(t^*)$ in the limit of $S_\succ$ and $S_\prec$;
  - thus, $\gamma(t^*) \in S_\sim$ – a contradiction.
- Every $(n-1)$-dim lin. space has the form $\sum\limits_{i=1}^{n} \alpha_i \cdot Y_i = 0$.
- Thus, $Y \in S_\succ \Leftrightarrow \sum \alpha_i \cdot Y_i > 0$, and

$$y \succ y' \Leftrightarrow \sum \alpha_i \cdot \ln(y_i/y'_i) > 0 \Leftrightarrow \prod y_i^{\alpha_i} > \prod {y'_i}^{\alpha_i}.$$