Dominance-Based Decision Theory

Kenny Easwaran

Group in Logic and the Methodology of Science
University of California, Berkeley

Introduction

- Classical decision theory decides between actions by calculating their expected utilities.
- But various paradoxes involving infinities and undefined probabilities have shown some of the limits of expected utility reasoning.
- I would like to develop a new decision theory centered instead on dominance reasoning - this is the beginnings of that project.

Decision Theory

- The goal of decision theory is to provide some sort of means for an agent to choose from among the available options.
- Much traditional decision theory has aimed to map actions to real numbers, but this seems to do much more than we should expect.
- I will suggest that we should only hope to get some linear ordering among the actions, and if certain types of infinitary actions are available, then perhaps even only a partial ordering with some incomparable actions.
- I will treat actions as functions from a state space (intuitively, the set of all relevant alternatives for how the world might actually be) to the set of possible outcomes.
- The set of outcomes will have a (linear or partial) pre-ordering given by $>, \geq, \equiv$

Not real numbers

- You have just entered heaven, and have to decide how you will spend your time there. The joy one experiences on each day can be represented by a positive natural number, but any number is available.
- Because you will have countably many days in heaven, if one sequence of natural numbers is larger than another on all but finitely many days, then the first sequence represents a better choice.
- However, among the choices of infinite sequences, one can construct an ascending chain of length ω_1, so in particular the values of the sequences can't be represented by real numbers, since there is no uncountable increasing sequence of reals.

St. Petersburg and Pasadena

- I will repeatedly flip a (fair) coin until it comes up heads - the payoff will depend only on n, the number of flips needed.
- In St. Petersburg, the payoff is $\$2^n$, while in Pasadena it is $\$\frac{(-2)^n}{n}$. [Nover and Hájek, "Vexing Expectations"]
- The expected utility of St. Petersburg is thus infinite, while that of Pasadena is undefined. (In one order, the sum comes to $\log 2$, but it can sum to any value in $[-\infty, +\infty]$.)

Leningrad and Altadena

- Despite these problematic expectations, it seems that we can still make some decisions.
- The Leningrad game [Colyvan, "Relative Expectation Theory"] has payoffs \$$(2^n + 1)$ while the Altadena game has payoffs \$$\left(\frac{(-2)^n}{n} + 1\right)$.
- Although the expectations are respectively still ∞ and undefined, it seems clear that one should prefer Leningrad to St. Petersburg and Altadena to Pasadena, because the same state always leads to a greater payoff in the former cases.
- This is dominance reasoning. In most cases it is much weaker than expected utility, but here it speaks where expected utility can't.

Partial Orderings

- I want to characterize some preference ordering on available actions for an agent.
- I will use the symbols $\succ, \succeq, \approx$ for strong and weak preference, and indifference respectively.
- I will call a partial ordering using these symbols *adequate* iff it meets the following conditions:
  1. $A \succeq A$
  2. If $A \succeq B$ and $B \succeq C$ then $A \succeq C$
  3. If $A \succ B$ then $A \succeq B$
  4. If $A \succ B$ then $B \not\succeq A$
  5. If $A \succ B$ and $B \succeq C$, or if $A \succeq B$ and $B \succ C$, then $A \succ C$
  6. $A \approx B$ iff $A \succeq B$ and $B \succeq A$
- Note that I do not require that $A \succ B$ iff $A \succeq B$ and $A \not\approx B$.

Compatible Extensions

- If $(\succ_1, \succeq_1, \approx_1)$ is an adequate ordering, then I will say that an equivalence relation $\approx_2$ is *compatible* with it iff the following conditions hold:
  1. If $A \succ_1 B$ then $A \not\approx_2 B$
  2. There is B such that $A \succ_1 B$ and $B \approx_2 C$ iff there is B' such that $A \approx_2 B'$ and $B' \succ_1 C$. (In this case I will say that $A \succ_2 C$.)
  3. There is B such that $A \succeq_1 B$ and $B \approx_2 C$ iff there is B' such that $A \approx_2 B'$ and $B' \succeq_1 C$. (In this case I will say that $A \succeq_2 C$.)

Compatible Extensions

- If $(\succ_1, \succeq_1, \approx_1)$ is an adequate ordering, then I will say that an equivalence relation $\approx_2$ is *compatible* with it iff the following conditions hold:
  1. If $A \succ_1 B$ then $A \not\approx_2 B$
  2. There is B such that $A \succ_1 B$ and $B \approx_2 C$ iff there is B' such that $A \approx_2 B'$ and $B' \succ_1 C$. (In this case I will say that $A \succ_2 C$.)
  3. There is B such that $A \succeq_1 B$ and $B \approx_2 C$ iff there is B' such that $A \approx_2 B'$ and $B' \succeq_1 C$. (In this case I will say that $A \succeq_2 C$.)
- Theorem: If $\approx_2$ is compatible, then $\succ_2$ and $\succeq_2$, together with the naturally defined equivalence relation, is an adequate ordering.
- (Note that the equivalence relation $\approx_2$ is not necessarily the same as the relation generated by $A \succeq_2 B$ and $B \succeq_2 A$.)

Dominance

- My basic overall strategy will be to start with an adequate ordering given by dominance reasoning, and find compatible equivalence relations to extend it, hopefully getting closer to a linear ordering.
- I define the relations $\succ, \succeq, \approx$ between actions on the same state space as follows:
  1. $A \succ B$ iff for every state x, $A(x) > B(x)$
  2. $A \succeq B$ iff for every state x, $A(x) \geq B(x)$
  3. $A \approx B$ iff for every state x, $A(x) \equiv B(x)$
- One might have wanted to define $A \succ B$ iff $A \succeq B$ and $A \not\approx B$, but if I have time I will point out some problems with that.
- On this account, Leningrad beats St. Petersburg, and Altadena beats Pasadena, as hoped.

Permutations of States

- Note that not every two actions have the same state space - the coin might not be flipped if you choose a different action.
- Following a suggestion in Chapter 2 of [Schick, *Ambiguity and Logic*], I will consider some re-identification of states in different state spaces.
- The only constraint I propose for this re-identification is that identified states must have the same probability.
- This suggests an equivalence relation $\approx_D$ ("equivalence in distribution"), which two actions bear to one another iff there is a probabilistic isomorphism of the state spaces, such that corresponding outcomes of the two actions are equally good.
- This equivalence is compatible with the dominance ordering.

Future Plans

- Having found one compatible equivalence relation by allowing permutations of states with equal probability, I would like to consider another equivalence relation allowing movement of utility between outcomes, provided that the utility scale has some additive structure.
- Another possible equivalence relation (which isn't compatible unless measure 0 events are ignored) considers two actions equivalent if their expected difference in utility is 0.
- Either of these extensions would give an ordering extending the ordering given by traditional expected utility-based decision theory.
- However, I conjecture that no extension will ever give a total ordering, because of examples like the Pasadena game, which seems in many ways strongly incomparable to the status quo.