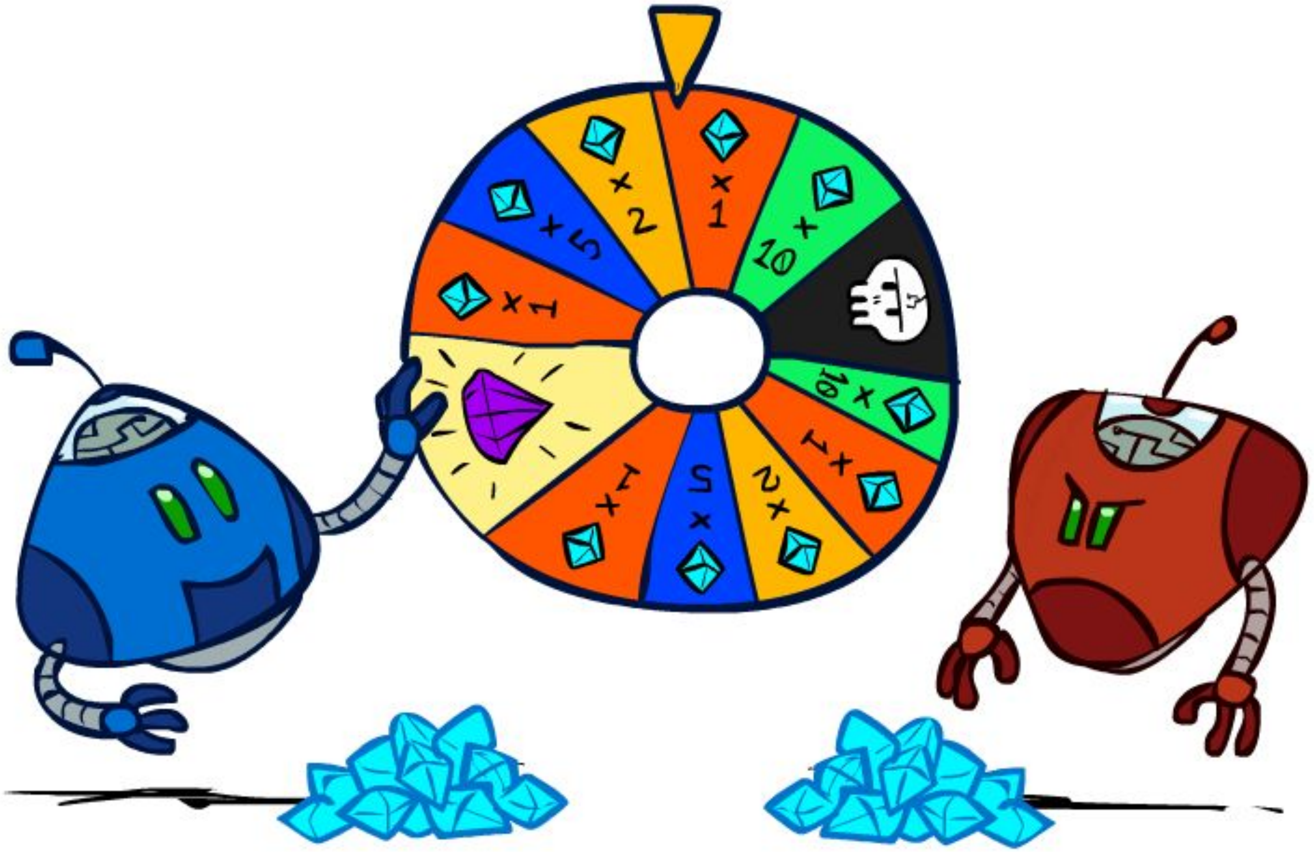
CS 188: Artificial Intelligence

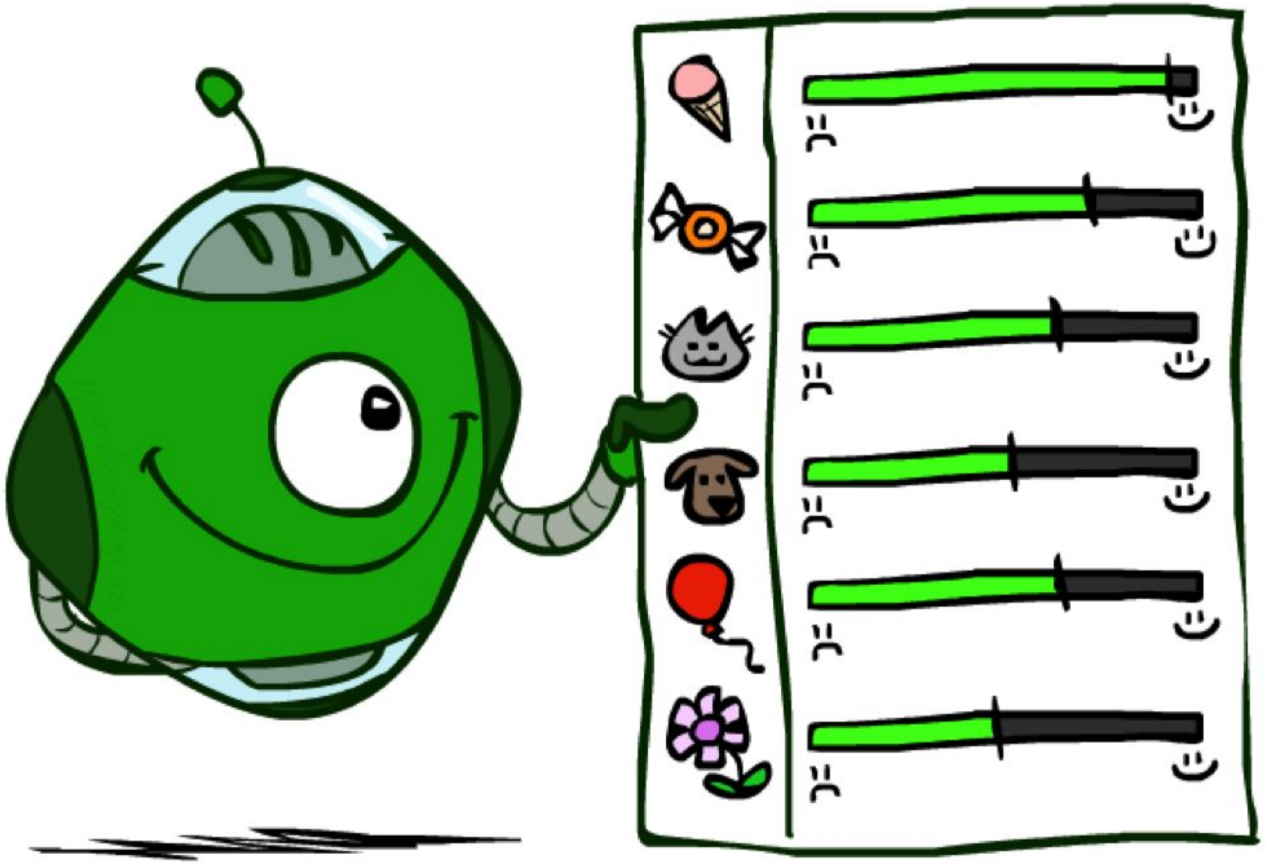
Rational Decisions



Instructor: Stuart Russell and Dawn Song

University of California, Berkeley

Utilities

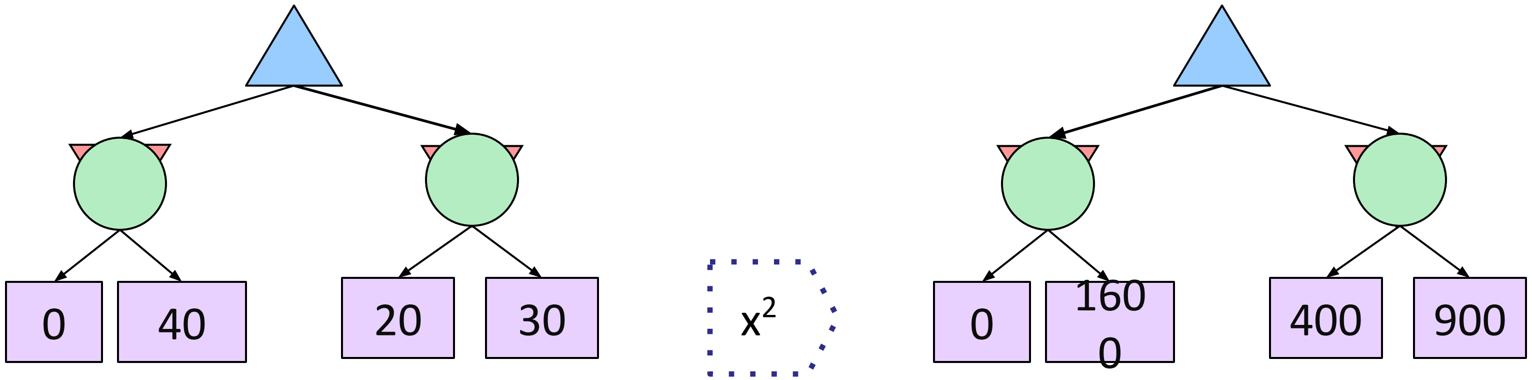


Maximum Expected Utility

- Principle of maximum expected utility:
  - A rational agent should choose the action that maximizes its expected utility, given its knowledge

- Questions:
  - Where do utilities come from?
  - How do we know such utilities even exist?
  - How do we know that averaging even makes sense?
  - What if our behavior (preferences) can't be described by utilities?

The need for numbers



- For worst-case minimax reasoning, terminal value scale doesn't matter
  - We just want better states to have higher evaluations (get the ordering right)
  - The optimal decision is invariant under any ***monotonic transformation***
- For average-case expectimax reasoning, we need ***magnitudes*** to be meaningful

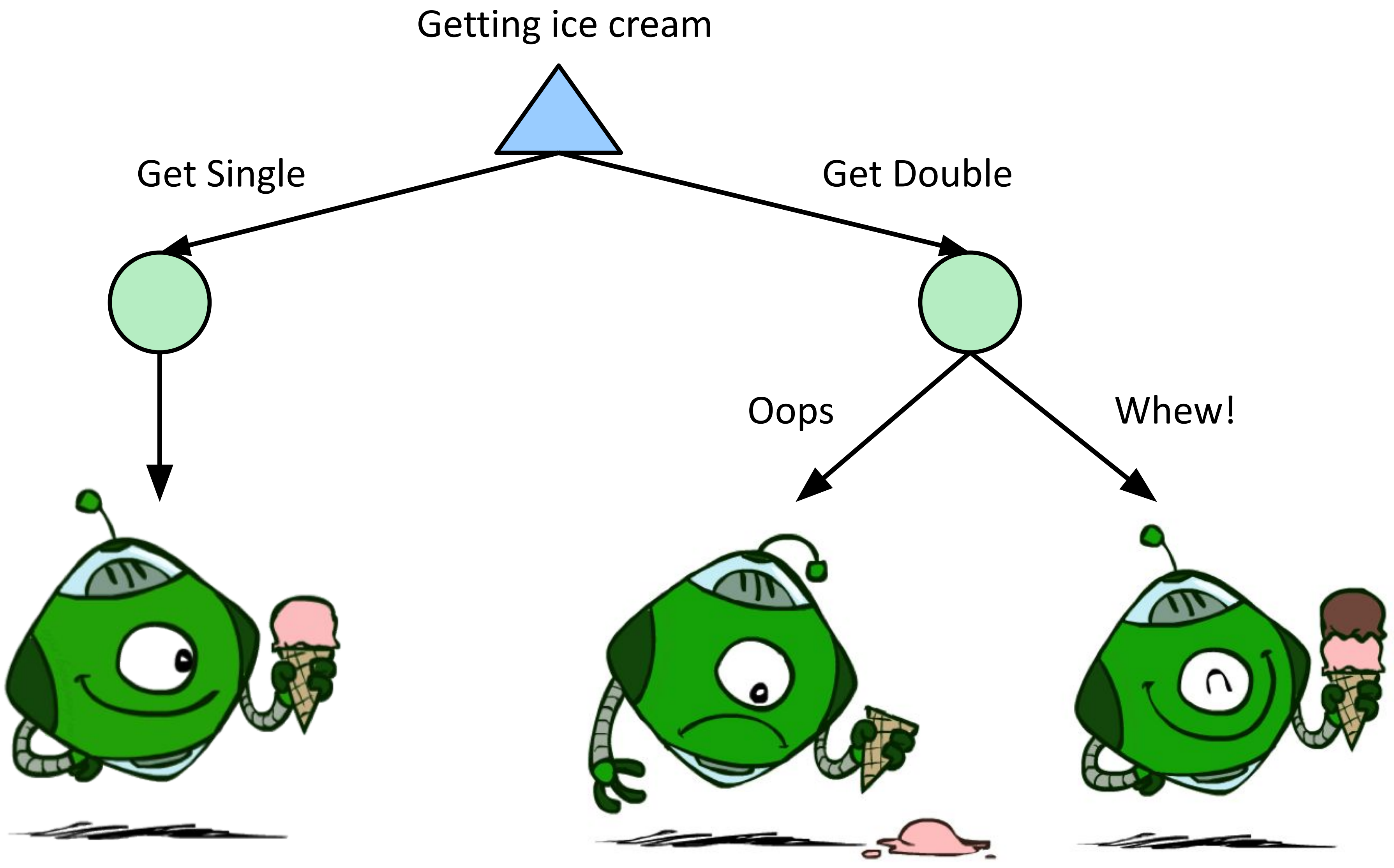
Utilities

- Utilities are functions from outcomes (states of the world) to real numbers that describe an agent's preferences

- Where do utilities come from?
  - In a game, may be simple (+1/-1)
  - Utilities summarize the agent's goals
  - Theorem: any "rational" preferences can be summarized as a utility function

- We hard-wire utilities and let behaviors emerge
  - Why don't we let agents pick utilities?
  - Why don't we prescribe behaviors?

Utilities: Uncertain Outcomes



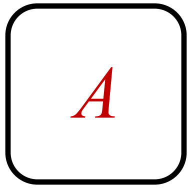
Preferences

- An agent must have preferences among:
  - Prizes: A, B, etc.
  - Lotteries: situations with uncertain prizes

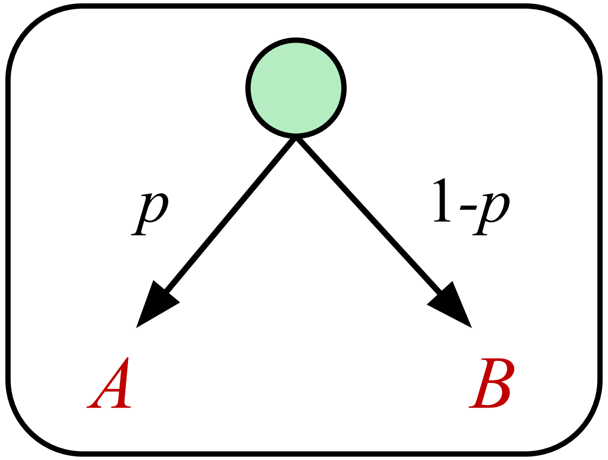
  $L = [p, A;\ (1-p), B]$

- Notation:
  - Preference: $A > B$
  - Indifference: $A \sim B$

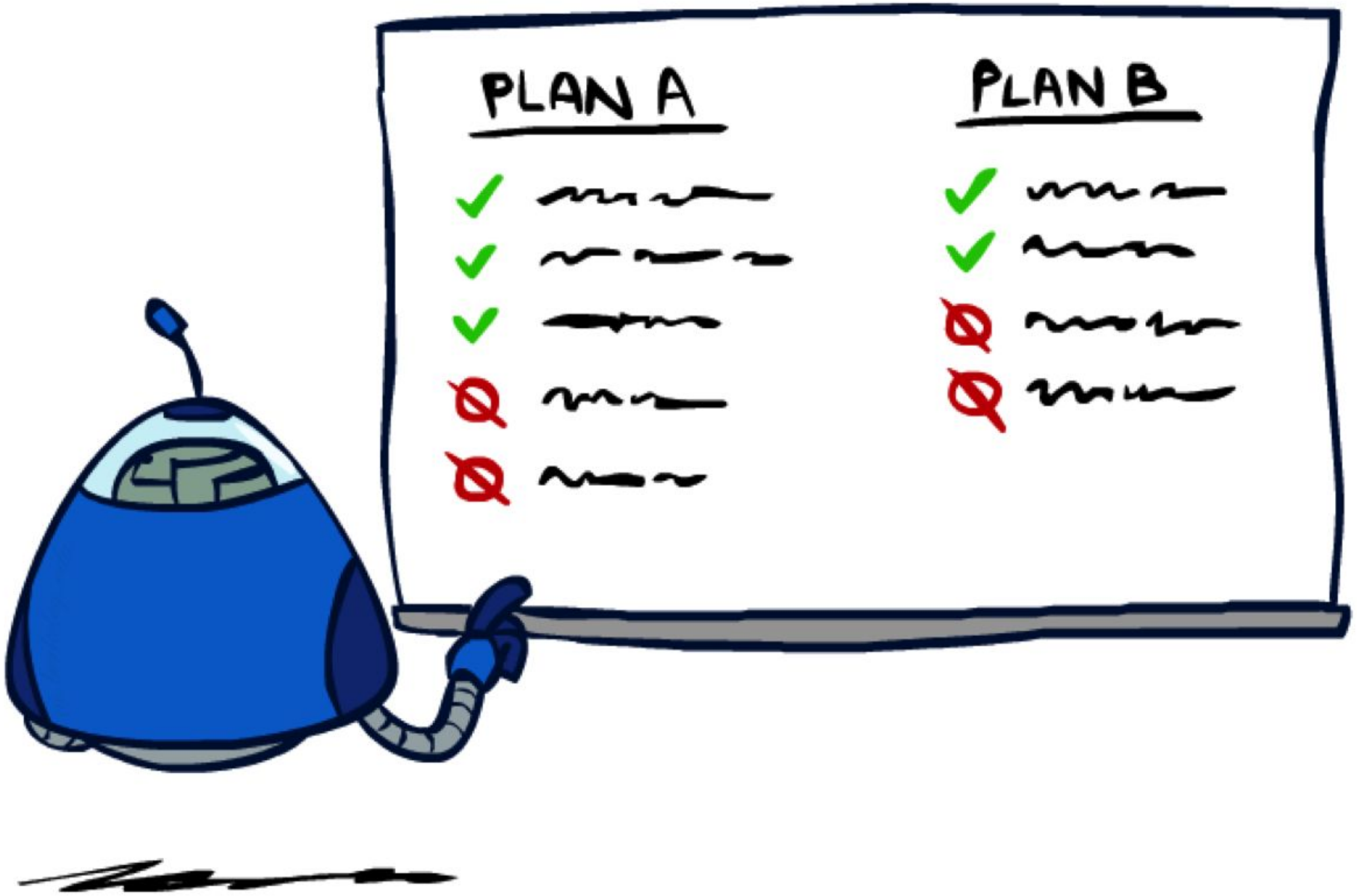
A Prize



A Lottery



Rationality

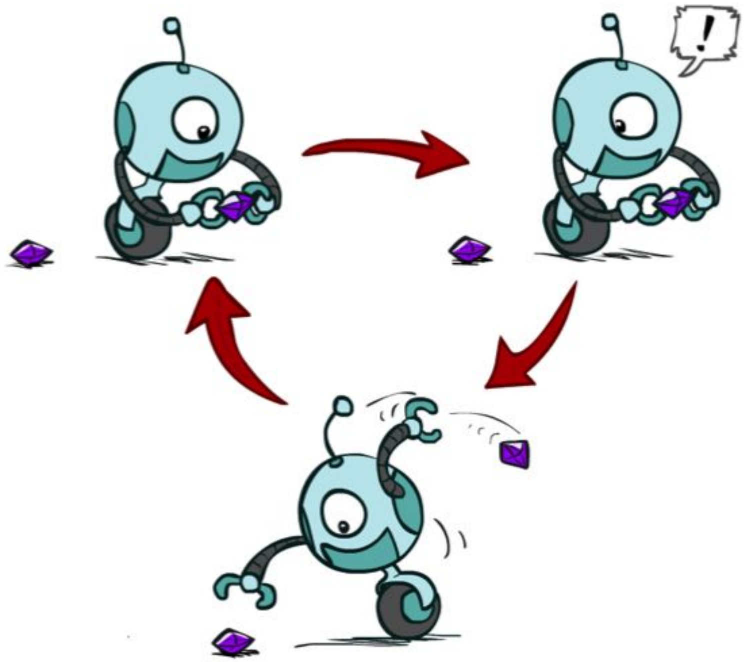


Rational Preferences

- We want some constraints on preferences before we call them rational, such as:

Axiom of Transitivity: $(A > B) \wedge (B > C) \Rightarrow (A > C)$

- For example: an agent with intransitive preferences can be induced to give away all of its money
  - If B > C, then an agent with C would pay (say) 1 cent to get B
  - If A > B, then an agent with B would pay (say) 1 cent to get A
  - If C > A, then an agent with A would pay (say) 1 cent to get C



Rational Preferences

The Axioms of Rationality

Orderability:

$(A > B) \lor (B > A) \lor (A \sim B)$

Transitivity:

$(A > B) \land (B > C) \Rightarrow (A > C)$

Continuity:

$(A > B > C) \Rightarrow \exists p\ [p, A;\ 1\text{-}p, C] \sim B$

Substitutability:

$(A \sim B) \Rightarrow [p, A;\ 1\text{-}p, C] \sim [p, B;\ 1\text{-}p, C]$

Monotonicity:

$(A > B) \Rightarrow$

$(p \geq q) \Leftrightarrow [p, A;\ 1\text{-}p, B] \geq [q, A;\ 1\text{-}q, B]$

Theorem: Rational preferences imply behavior describable as maximization of expected utility

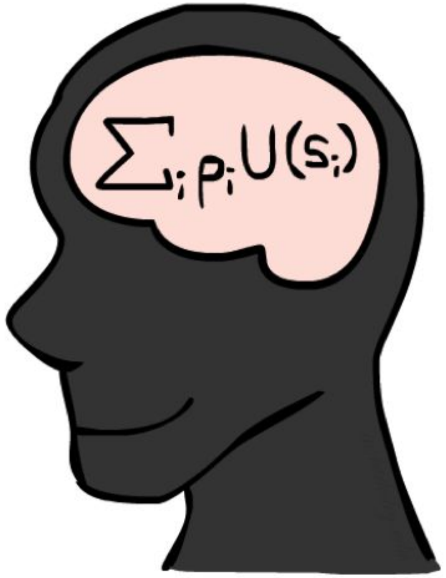
MEU Principle

- Theorem [Ramsey, 1931; von Neumann & Morgenstern, 1944]
  - Given any preferences satisfying these constraints, there exists a real-valued function U such that:

$$U(A) \geq U(B) \Leftrightarrow A \succeq B$$

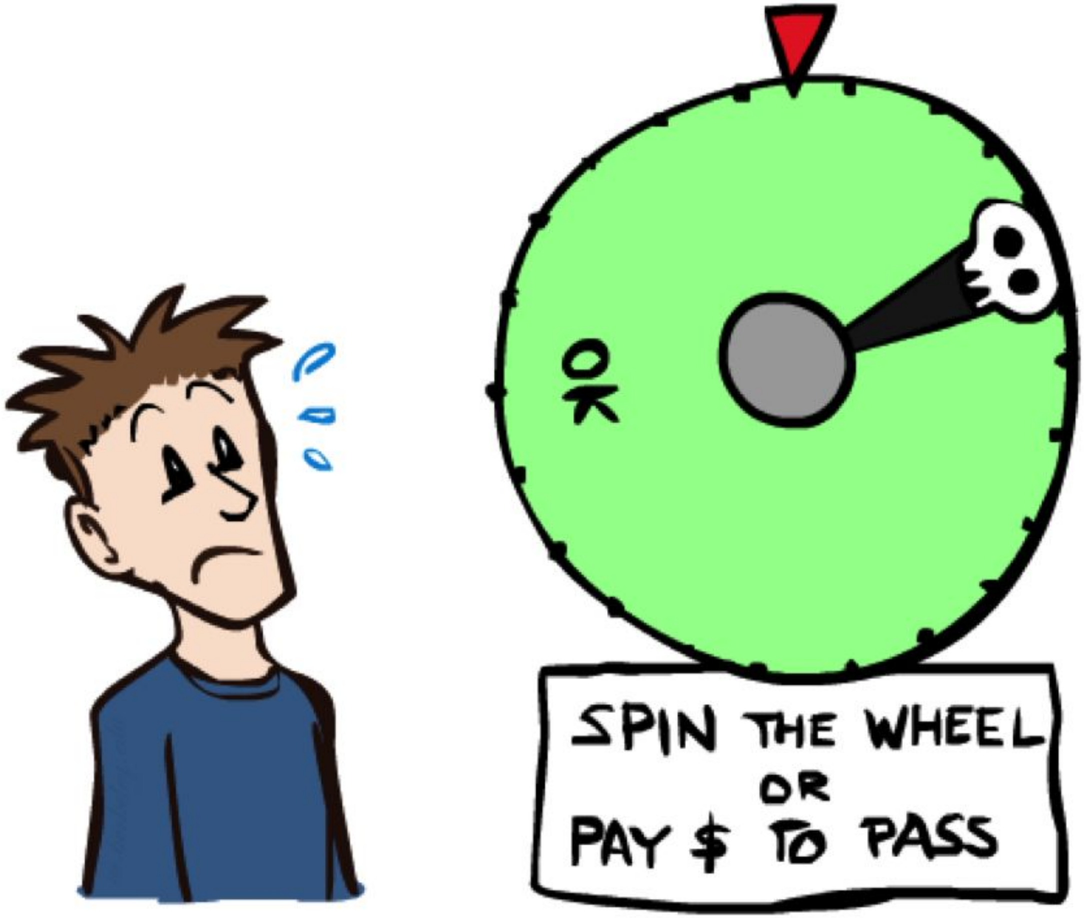
$$U([p_1,S_1; \ldots ; p_n,S_n]) = p_1 U(S_1) + \ldots + p_n U(S_n)$$

  - I.e. values assigned by U preserve preferences of both prizes and lotteries!
  - Optimal policy invariant under ***positive affine transformation*** $U' = aU+b$, $a>0$



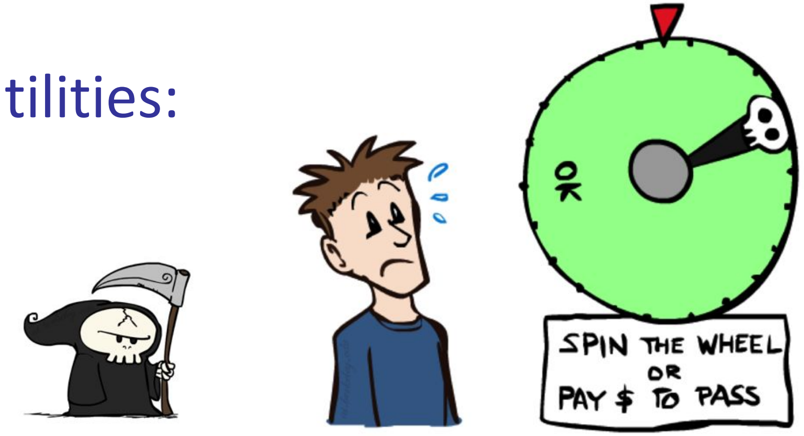
- Maximum expected utility (MEU) principle:
  - Choose the action that maximizes expected utility
  - Note: rationality does ***not*** require representing or manipulating utilities and probabilities
    - E.g., a lookup table for perfect tic-tac-toe

Human Utilities

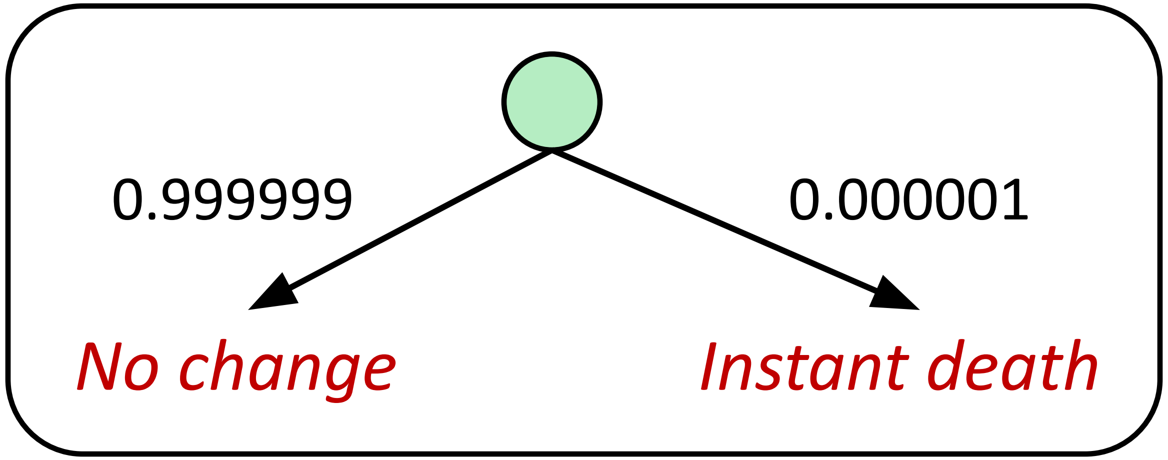


Human Utilities

- Utilities map states to real numbers. Which numbers?
- Standard approach to assessment (elicitation) of human utilities:
  - Compare a prize A to a ***standard lottery*** L_p between
    - "best possible prize" $u_\top$ with probability p
    - "worst possible catastrophe" $u_\bot$ with probability $1\text{-}p$
  - Adjust lottery probability p until indifference: $A \sim L_p$
  - Resulting p is a utility in $[0,1]$

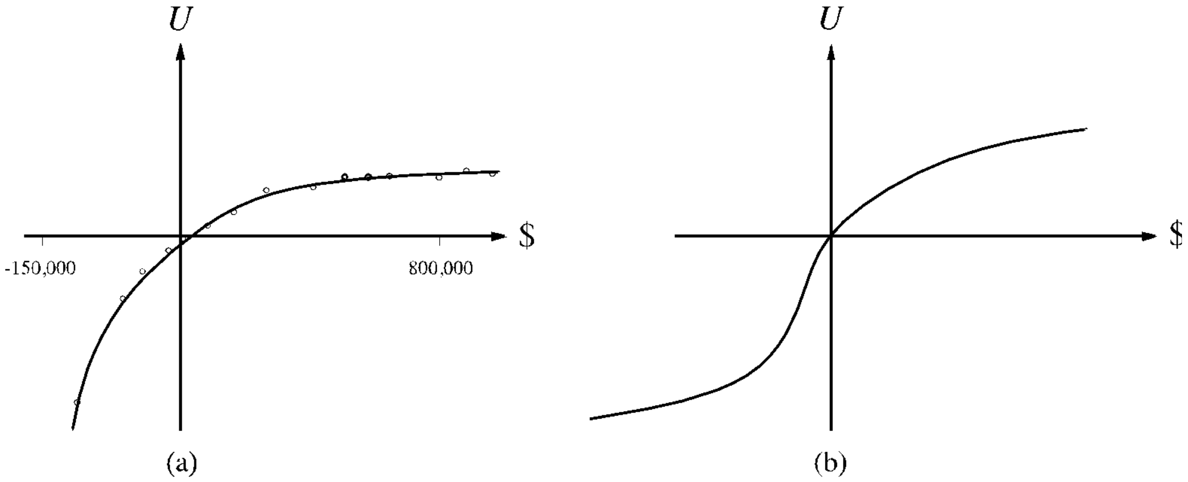


Pay $50 $\sim$



Money

- Money ***does not*** behave as a utility function, but we can talk about the utility of having money (or being in debt)
- Given a lottery L = [p, \$$X$; (1-$p$), \$$Y$]
  - The ***expected monetary value*** EMV(L) = $pX + (1-p)Y$
  - The utility is $U(L) = pU(\$X) + (1-p)U(\$Y)$
  - Typically, $U(L) < U($ EMV(L) $)$
  - In this sense, people are ***risk-averse***
  - E.g., how much would you pay for a lottery ticket L=[0.5, \$10,000; 0.5, \$0]?
  - The ***certainty equivalent*** of a lottery CE(L) is the cash amount such that CE(L) ~ L
  - The ***insurance premium*** is EMV(L) - CE(L)
  - If people were risk-neutral, this would be zero!



Post-decision Disappointment: the Optimizer's Curse

- Usually we don't have direct access to exact utilities, only ***estimates***
  - E.g., you could make one of *k* investments
  - An unbiased expert assesses their expected net profit $V_1,\ldots,V_k$
  - You choose the best one V^*
  - With high probability, ***its actual value is considerably less*** than V^*
- This is a serious problem in many areas:
  - Future performance of mutual funds
  - Efficacy of drugs measured by trials
  - Statistical significance in scientific papers
  - Winning an auction

Suppose true net profit is 0 and estimate ~ N(0,1);
Max of k estimates: