Uncertainty and Utilities in Search

Dr. Demetrios Glinos
University of Central Florida

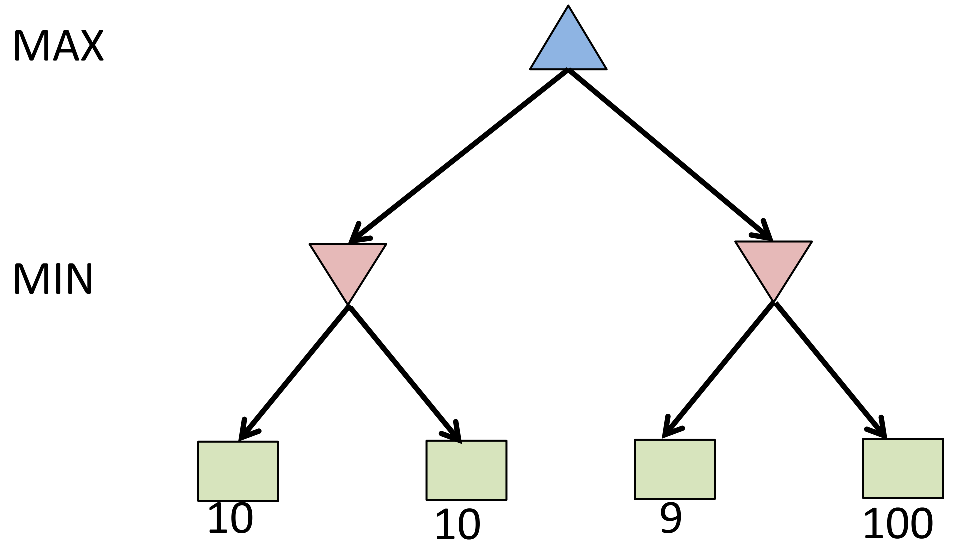
CAP4630 – Artificial Intelligence

Today

- Uncertainty in Search
  - Expected Value
  - Limiting Search
- Mixed and Multi-Player Games
- Utilities
  - Rational Preferences
  - Utility Functions

Topic 1: Uncertainty in Search

Last time: Minimax



Q: Minimax tells us to choose left branch, but what if opponent is less than perfect, or if the environment is just responding nondeterministically?

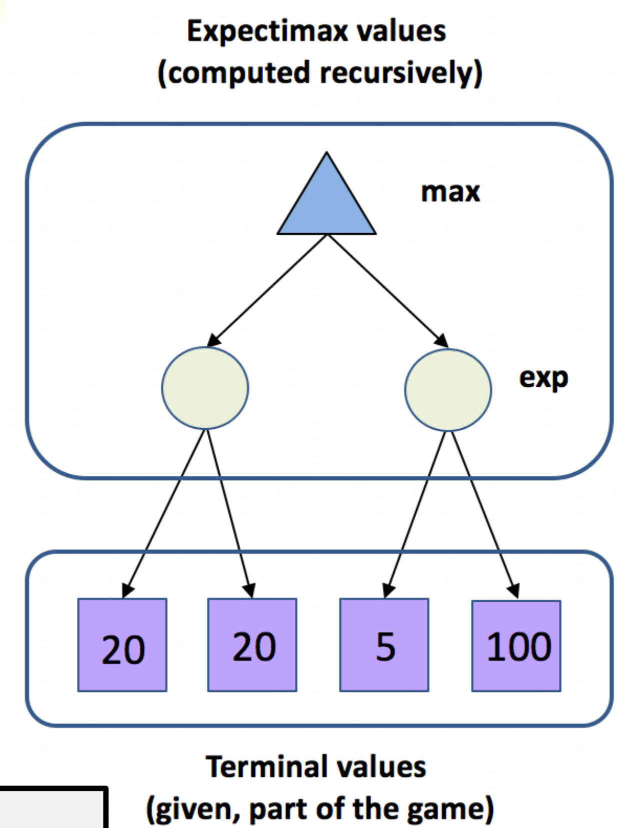
Expectimax

Minimax is worst-case analysis

"Chance" or ***"Expectimax"*** nodes allow us to consider average-case outcomes

Expectimax search: computes the average score under optimal play for Player (MAX)

- terminal states still have utilities
- max nodes same as for minimax
- chance nodes calculate *expected utility*



Expected utility = $\Sigma_i\, p_i U(s_i)$

where p_i is probability of successor state s_i

and all probabilities are positive and sum to 1.0

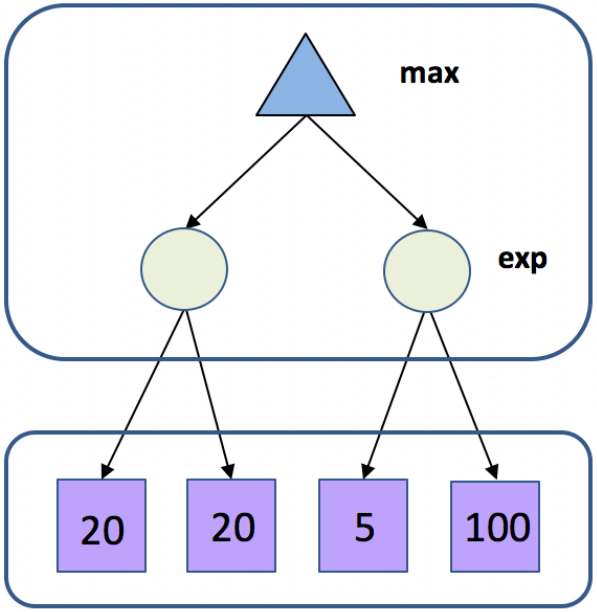
Expectimax Algorithm

Value(state) =
if terminal state, then return the state's utility
else if next agent is **MAX**, then return Max-Value(state)
else if next agent is **EXP**, then return Exp-Value(state)

```
Max-Value( state ):
    v ← - ∞
    for each successor s' of s {
        v = max( v, value( s' ) )
    }
    return v
```

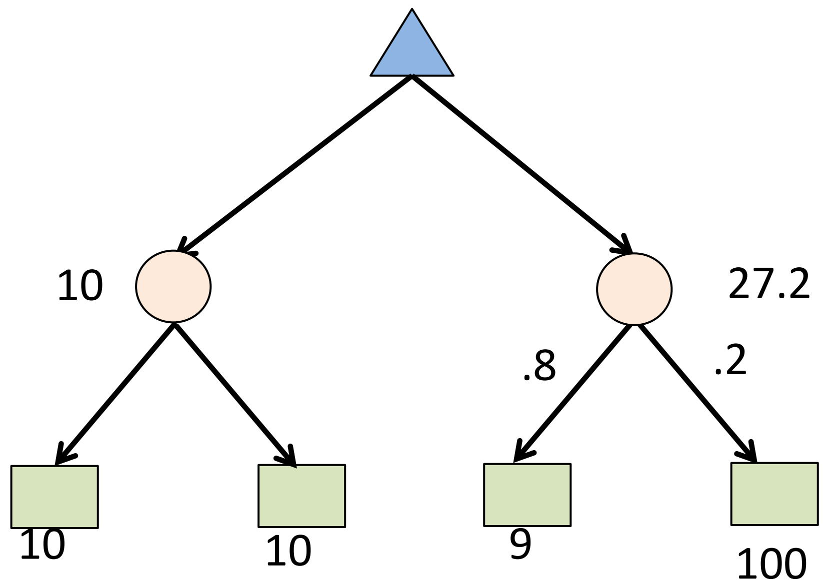
```
Exp-Value( state ):
    v ← 0
    for each successor s' of s {
        p = probability( s' )
        v += p*value( s' )
    }
    return v
```

Expectimax values (computed recursively)



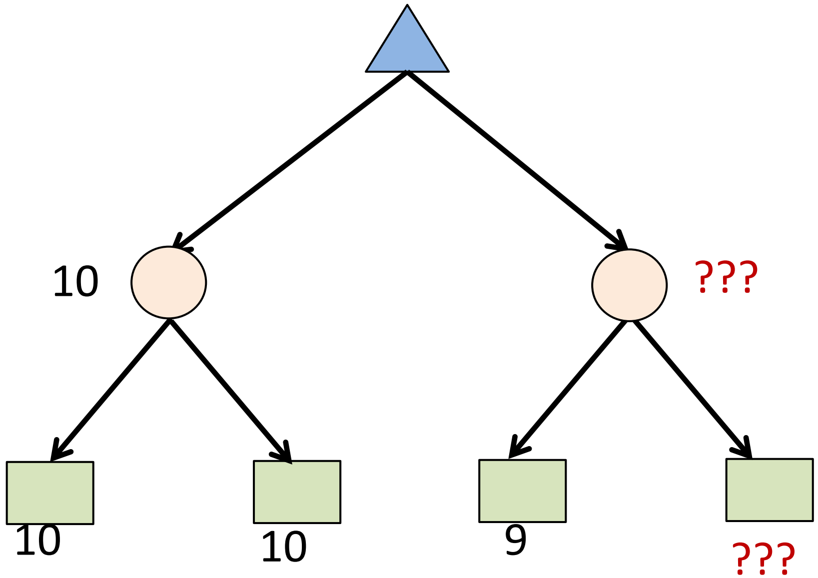
Terminal values (given, part of the game)

Applying Expectimax



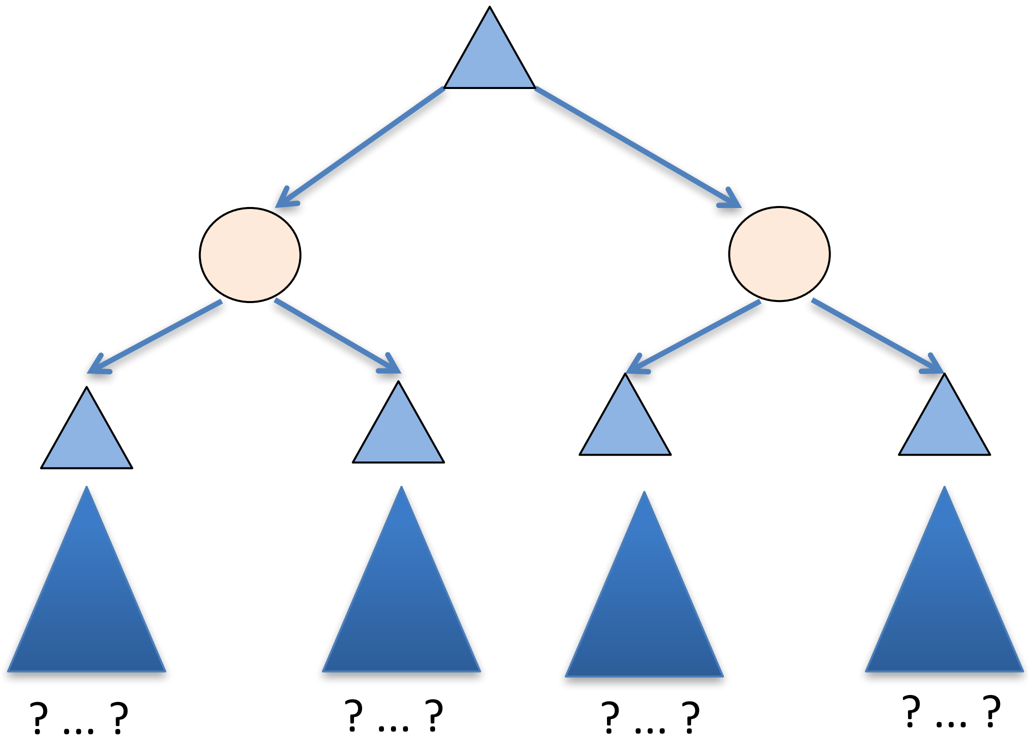
Here, expectimax chooses right branch

Expectimax and Pruning



Expectimax cannot prune: need all values to compute expected value

Depth-Limited Expectimax

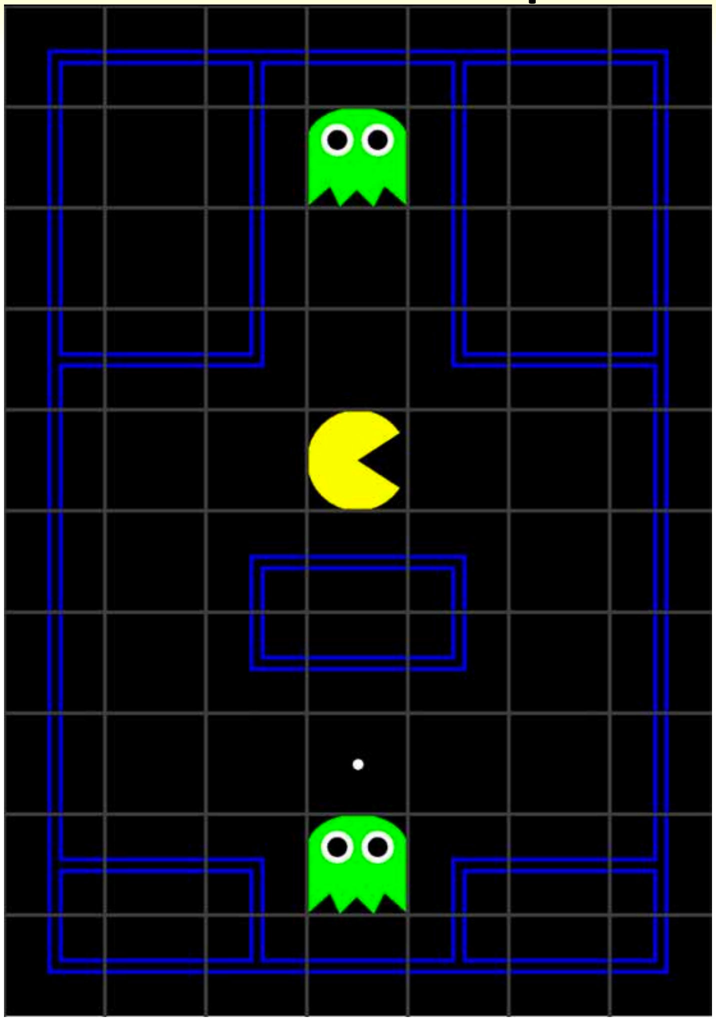


Use an evaluation function as before
Applies to both MAX and EXP nodes

About those Probabilities

- Recall:
  - random variable – an event whose outcome is uncertain
  - probability distribution – an assignment of probabilities (weights) to outcomes
    - probabilities are always nonnegative
    - probabilities in a distribution must sum to 1.0
- Outcome probabilities
  - can be based on a simple model of the environment or opponent (e.g., 6-sided die)
  - can be computed based on the state and/or the player's experience
  - can be just given
- Note
  - opponent or environment not necessarily rolling dice (e.g., the weather)
  - they are merely out of the Player's control

Impact of Strategies



demos: trapsmart, trapsmartminimax
demos: traprandom, traprandomminimax

Results for 5 trials:

	Ghosts: *Adversarial*	Ghosts: *Random*
Minimax Agent	Won 20/20	Won 20/20
Replan Agent	Won 0/20	Won 9/20

Q: Can you see why Replan Pac-Man always loses against smart ghosts on this maze?

Modeling the Opponent

It is important to model the opponent appropriately

	Opponent: *Adversarial*	Opponent: *Random*
Minimax Player	The best you can do against an adversary	Generally OK, but takes a little longer
Expectimax Player	Generally much worse than above	Better than adversarial

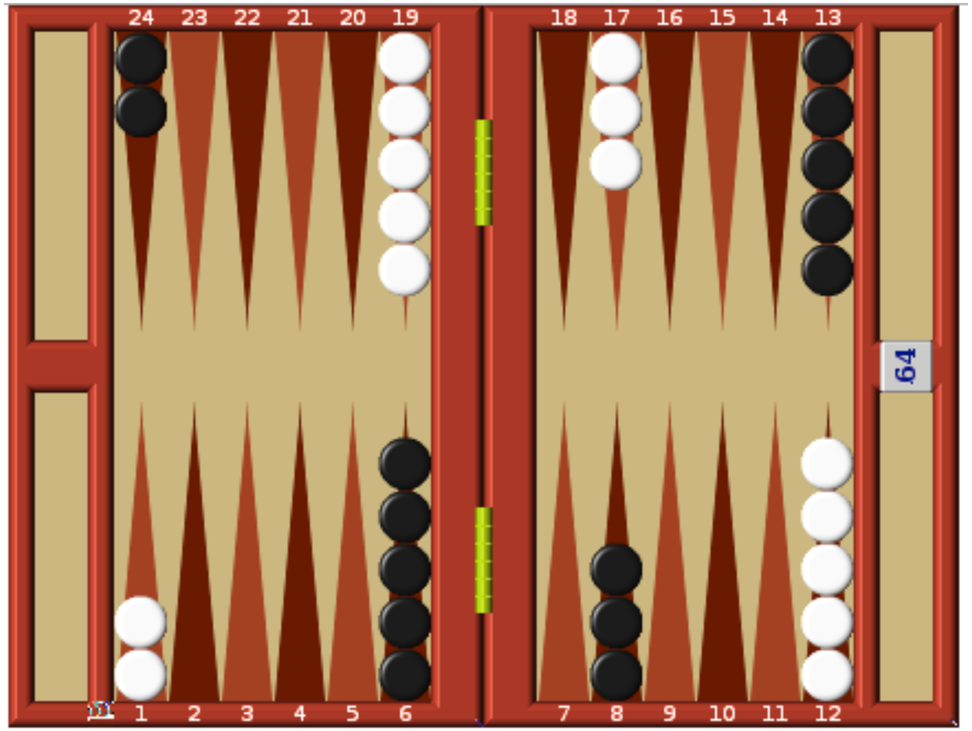
Unwarranted optimism (assuming probabilistic when adversarial) can be dangerous
Unwarranted pessimism (assuming adversarial when probabilistic) can be wasteful

Topic 2: Mixed and Multi-Player Games

Mixed Layer Game: Backgammon

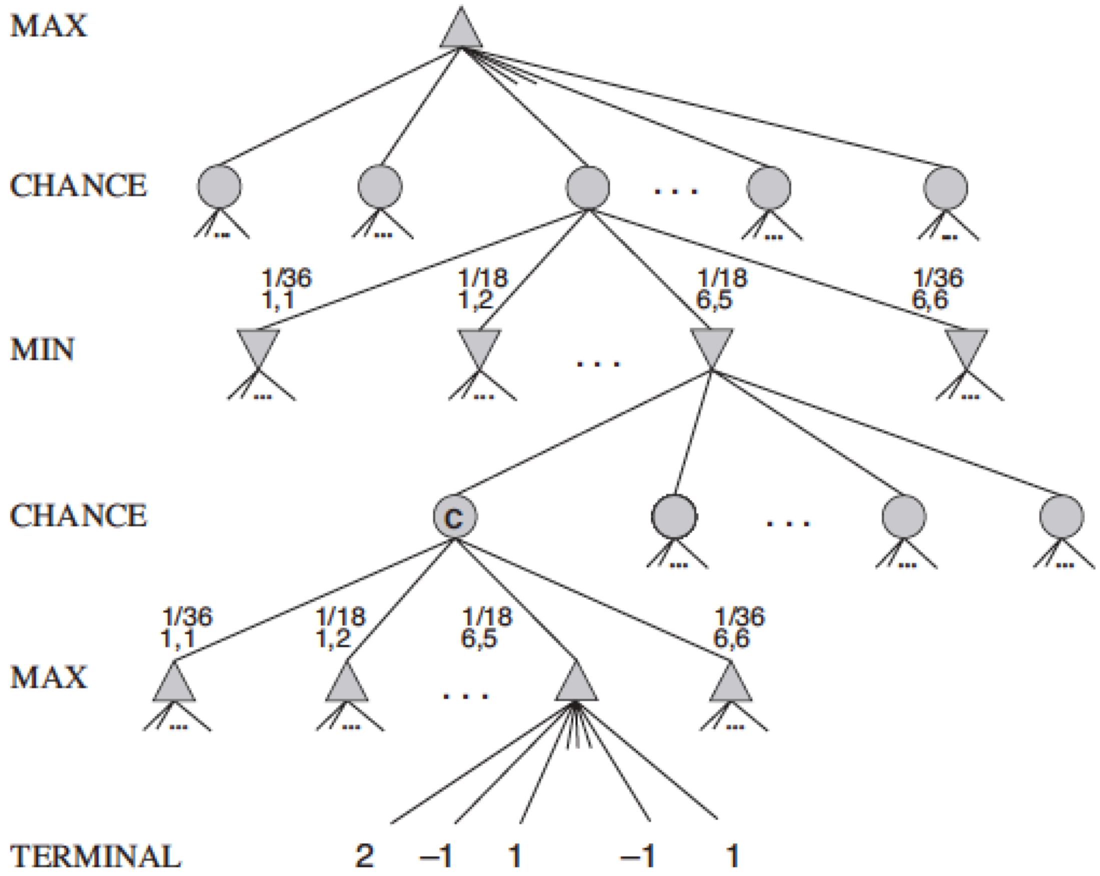
- Two-player zero-sum game with dice
- Dice rolls create large branching factor:
  - 21 possible rolls with 2 dice
  - about 20 legal moves
  - Depth 2: 1.2 x 10^9 states

- "TD-Gammon" (1992, IBM)
  - used depth-2 search + evaluation function trained on over 1 million games against itself
  - used emulator based on reinforcement learning and neural networks
  - achieved top-level play



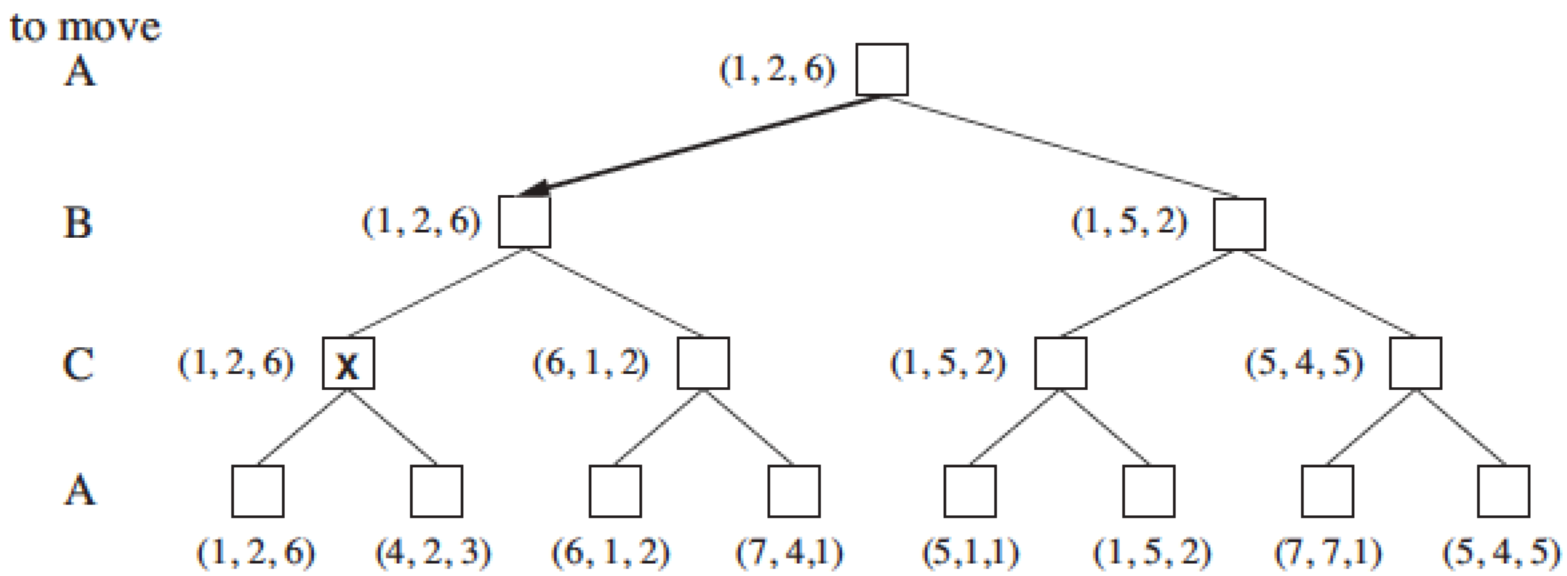
white moves counter-clockwise
black moves clockwise

Mixed Layers in Backgammon



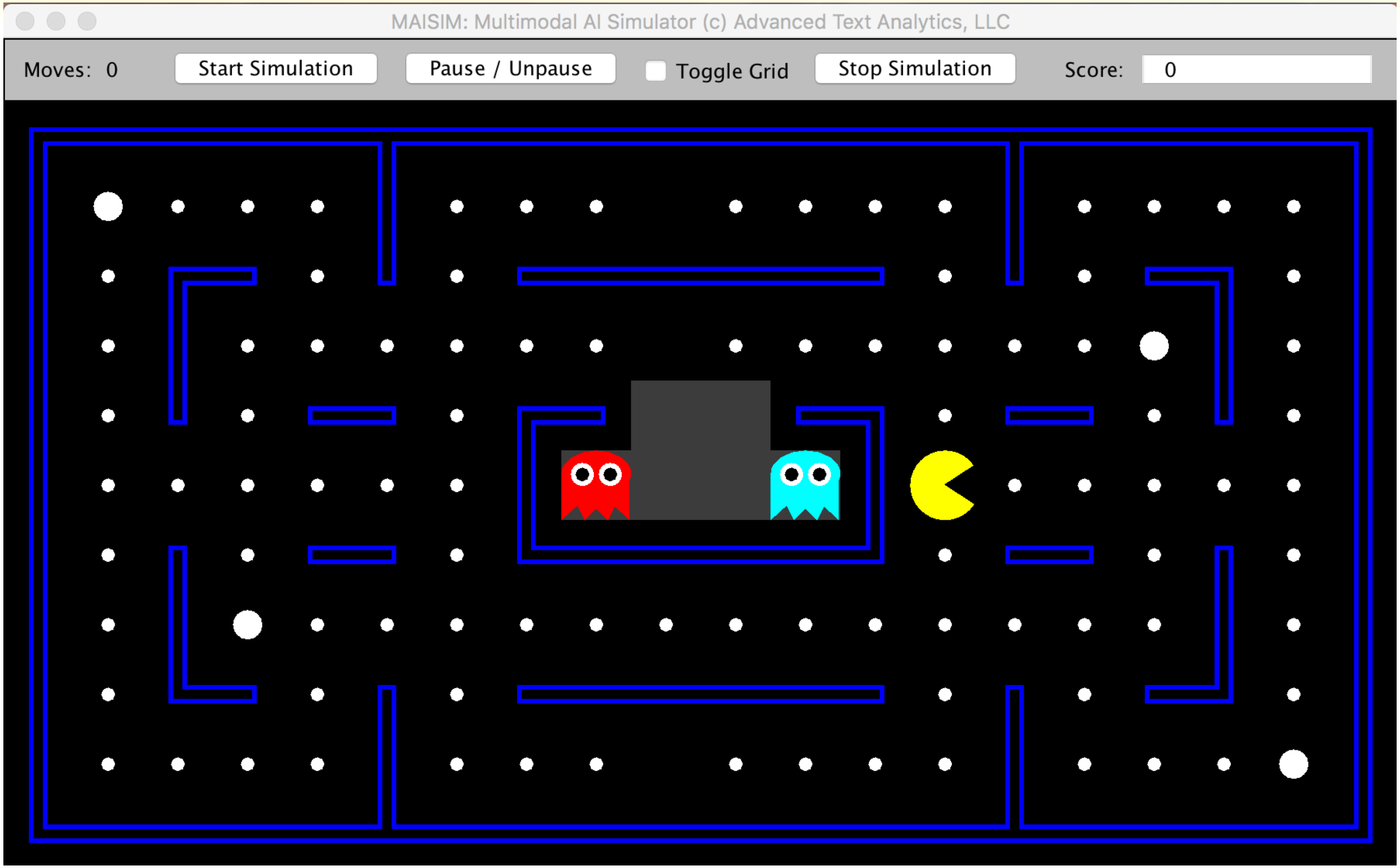
MAX, MIN, and EXP nodes compute appropriate combinations of their children

Multi-Player Games



Here, assume each node is a MAX for the player who moves
Utilities extended to show utility for each player

Mixed Mode Example



demo: minimax (minimax depth 2, partially stochastic opponents)

Topic 3: Utilities

Utility

- The measure of *value* to the player
- The basis for *rational choices*

Maximum Expected Utility (MEU):

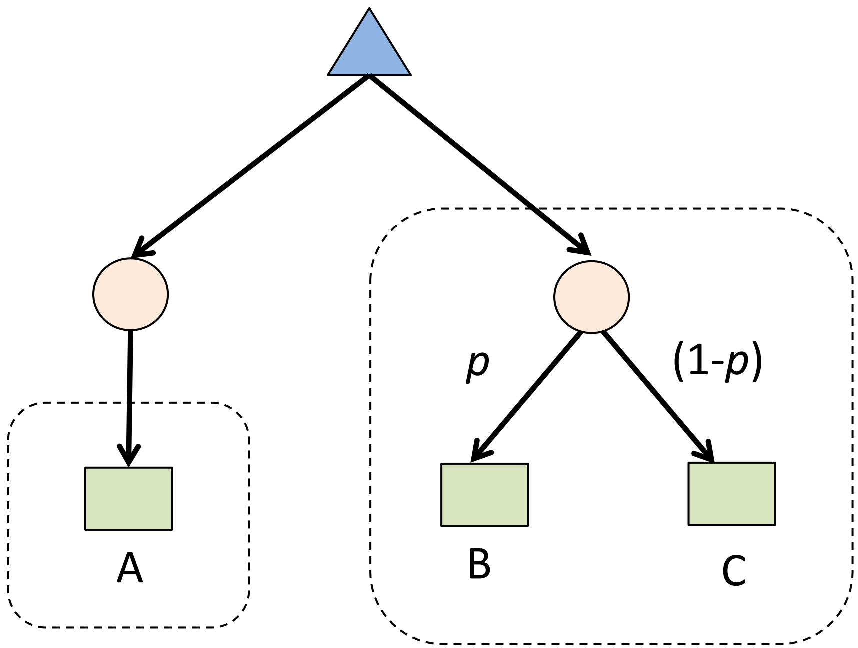
A rational agent should choose the action that maximizes the agent's expected utility, ***given its knowledge (i.e., what it knows now)***:

$$\textbf{action} = \textbf{argmax}_{\textbf{actions}}\ \textbf{ExpectedUtility(action | e)}$$

where e is the set of evidence observations (i.e., accumulated knowledge so far)

- This expresses a *preference* for the ***average***, as opposed to maximum or something else

Preferences Terminology



A is a ***prize***

This is a **lottery**: L = [p, B; (1-p), C]

where p is a probability

Notation:

Preference: $A \succ B$

Indifference: A ~ B

Rational Preferences

- To be useful for determining utilities, preferences must be rational
- To be rational, preferences must satisfy these axioms of rationality

Axioms of rationality:

Orderability: Exactly one of $(A \succ B)$, $(B \succ A)$ or $(A \sim B)$ holds

Transitivity: $(A \succ B) \land (B \succ C) \Rightarrow (A \succ C)$

Continuity: $A \succ B \succ C \Rightarrow \exists\, p \ni [\,p, A; (1-p), C\,] \sim B$

Substitutability: $A \sim B \Rightarrow [\,p, A; (1-p), C\,] \sim [\,p, B; (1-p), C\,]$

Monotonicity: $A \succ B \Rightarrow (\,p > q \Leftrightarrow [\,p, A; (1-p), B\,] \succ [\,q, A; (1-q), B\,]$

Decomposability: $[\,p, A; (1-p), [\,q, B; (1-q), C\,]] \sim [\,p, A; (1-p)q, B; (1-p)(1-q), C\,]$

Utilities from Preferences

Given rational preferences, von Neumann and Morgenstern (1944) proved:

Existence of Utility Function:

$U(A) > U(B) \Leftrightarrow A \succ B$

$U(A) = U(B) \Leftrightarrow A \sim B$

Expected Utility of a Lottery:

$U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i)$

Interpretation:

1. *values assigned by U preserve preferences for both prizes and lotteries*
2. *behavior is invariant under positive linear transformation:*

e.g., compare relative preferences for S and T under U and U', where
$U'(S) = a * U(S) + b$, and $U'(T) = a * U(T) + b$ where $a > 0$

Utility Scales

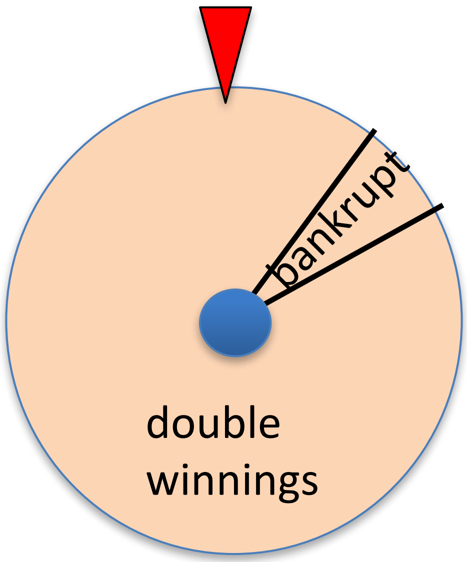
- Normalized utilities: $u_{worst} = 0.0$, $u_{best} = 1.0$
  - useful for comparing utilities
- "micromort": one-millionth chance of death
  - useful for cost-benefit of product improvements
- QALY ("Quality Adjusted Life Year"):
  - useful for choosing whether to undergo a medical procedure

Preference Elicitation

- Given a scale, we assess utility of a prize S and a "standard lottery" by adjusting p so that

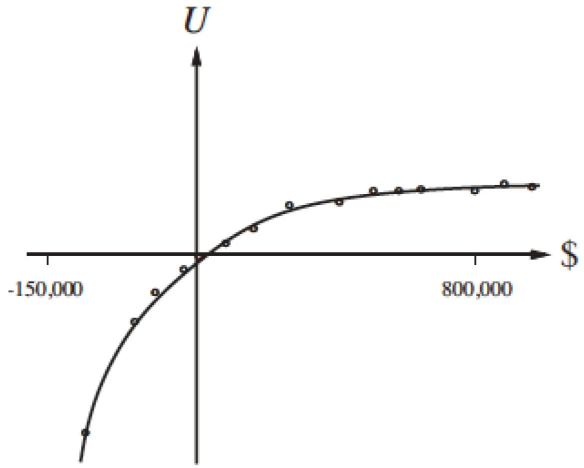
$$\boxed{S \sim [\, p,\ u_{worst},\ (1\text{-}p),\ u_{best}\ ]}$$

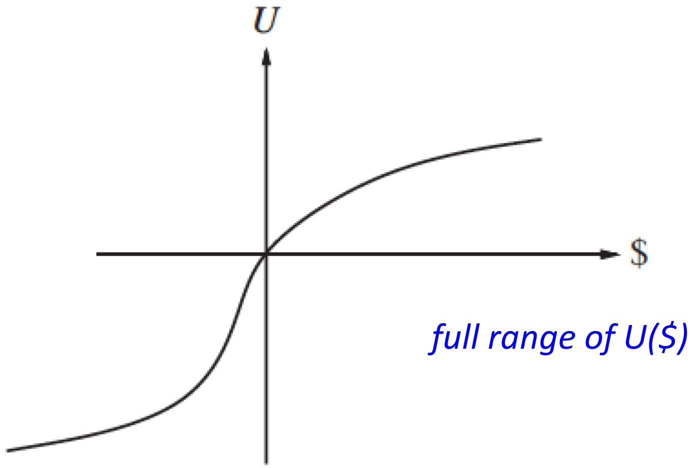
the value of p determines the size of the "bankrupt" wedge



Spin the wheel or pay to pass

Money





- **Money does not behave as a utility function**
- But we can talk about the utility of having money

- Consider a lottery: L = [p, $X; (1-p), $Y]
- Expected monetary value: EMV(L) = p*X + (1-p)*Y
- Utility of lottery: U(L) = p*U($X) + (1-p)*U($Y)
- Typically: U(L) < U[EMV(L)] ← people are generally risk-averse

  (opposite result when deeply in debt)

Insurance as Win-Win

- Consider the lottery: L = [0.5, $1000; 0.5, $0]
  - Expected monetary value: EMV(L) = $500
  - Certainty equivalent:
    - Amount acceptable instead of lottery
    - $400 for most people
  - The $100 difference is the **insurance premium**
    - what people will pay for the sure thing
    - the reason we have an insurance industry
  - Why this is “win-win”:
    - You would rather have the $400
    - The insrance company would rather have the premium and the lottery

Human Judgment and Irrationality

- Allais paradox (1953):
  - A = [0.8, $4K; 0.2 $0]
  - B = [1.0, $3K; 0.0 $0]
  - C = [0.2, $4K; 0.8 $0]
  - D = [.25, $3K; .75 $0]

Human Judgment and Irrationality

- Allais paradox (1953):
  - A = [0.8, $4K; 0.2 $0]
  - B = [1.0, $3K; 0.0 $0]

  - C = [0.2, $4K; 0.8 $0]
  - D = [.25, $3K; .75 $0]

- Most people prefer: B $\succ$ A and C $\succ$ D

- What does B $\succ$ A tell us?

Human Judgment and Irrationality

- Allais paradox (1953):
  - A = [0.8, $4K; 0.2 $0]
  - B = [1.0, $3K; 0.0 $0]
  - C = [0.2, $4K; 0.8 $0]
  - D = [.25, $3K; .75 $0]
- Most people prefer: $B \succ A$ and $C \succ D$
- But if U($0) = 0, then
  - $B \succ A$ implies U($3K) > 0.8 * U($4K)
- Now, What about $C \succ D$?

Human Judgment and Irrationality

- Allais paradox (1953):
  - A = [0.8, $4K; 0.2 $0]
  - B = [1.0, $3K; 0.0 $0]
  - C = [0.2, $4K; 0.8 $0]
  - D = [.25, $3K; .75 $0]
- Most people prefer: $B \succ A$ and $C \succ D$
- But if U($0) = 0, then
  - $B \succ A$ implies U($3K) > 0.8 * U($4K)
  - $C \succ D$ implies 0.8 * U($4K) > U($3K)

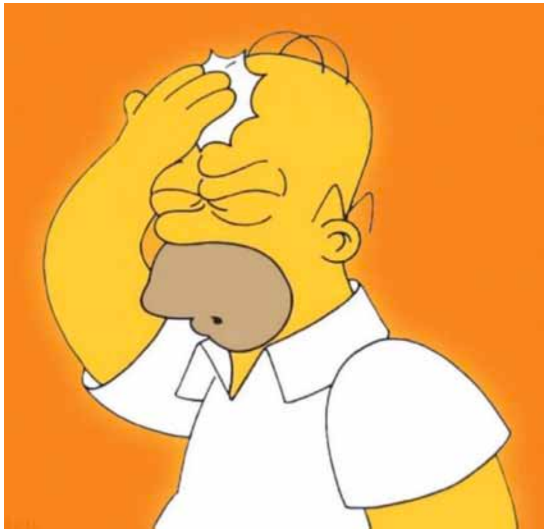
This is a contradiction !

➔ **Q: What are we missing, here?**

Human Judgment and Irrationality

- Allais paradox (1953):
  - A = [0.8, $4K; 0.2 $0]
  - B = [1.0, $3K; 0.0 $0]
  - C = [0.2, $4K; 0.8 $0]
  - D = [.25, $3K; .75 $0]
- Most people prefer: B $\succ$ A and C $\succ$ D
- But if U($0) = 0, then
  - B$\succ$A implies U($3K) > 0.8 * U($4K)
  - C$\succ$D implies 0.8 * U($4K) > U($3K)

➔ **Q: What are we missing, here?**



Answer: "**regret**", which is the feeling of being "stupid" if we choose A and lose
➔ This is why most people prefer B to A