Game, Expectiminimax, Utilities

CS 4804 Fall 2020
Virginia Tech
Today’s Topics

• Stochastic Game, Expeciminimax
• Utilities
Recap: Game Tree
Recap: Minimax and Alpha-Beta Pruning
Recap: Minimax and Alpha-Beta Pruning

- Minimax and alpha-beta are both assume that players are playing against an adversary who makes optimal decisions.
Stochastic Game

- Games that combine luck and skill
- Uncertain outcomes
Stochastic Game Tree

MAX

CHANCE

MIN

CHANCE

MAX

TERMINAL

2  -1  1  -1  1
$$\text{EXPECTINIMIMAX}(s) =$$

$$\begin{cases} 
\text{UTILITY}(s) & \text{if } \text{TERMINAL-TEST}(s) \\
\max_a \text{EXPECTINIMIMAX}(\text{RESULT}(s, a)) & \text{if } \text{To-Move } (s) = \text{MAX} \\
\min_a \text{EXPECTINIMIMAX}(\text{RESULT}(s, a)) & \text{if } \text{To-Move } (s) = \text{MIN} \\
\sum_r \Pr(r) \text{EXPECTINIMIMAX}(\text{RESULT}(s, r)) & \text{if } \text{To-Move } (s) = \text{CHANCE}
\end{cases}$$
Expectiminimax Search

- Explicit randomness
- Expectminimax value for game with chance nodes
- A chance node is an expected value. Sum of the value over all outcomes, weighted by the probability of each chance action
- Compute the average score under optimal play
Expected value \( V = \frac{1}{3} \times 24 + \frac{1}{2} \times 8 + \frac{1}{6} \times -12 = 10 \)
Probabilities Basic

• A random variable represents an event whose outcome is unknown
• A probability distribution is an assignment of weights to outcomes
• Probabilities are always positive
• Probabilities over all possible outcome sum to one
• Expectations: The expected value of a function of a random variable is the average, weighted by the probability distribution over outcomes
  – Example: How long to get to the ROA airport?
  – $45(25\%) + 55(55\%) + 60(20\%) = 53.5 \text{ (mins)}$
Expectimax Example

Assume each chance node have a probability of occurrence of 1/3

B: $3 \times \frac{1}{3} + 12 \times \frac{1}{3} + 9 \times \frac{1}{3} = 8$
Probability Basic

- $P(a, b) = P(a \mid b) \ P(b)$
- $P(a \mid b) = \frac{P(a, b)}{P(b)}$
- $P(b \mid a) = \frac{P(a \mid b) \ P(b)}{P(a)}$
Probability Quiz

1. \( P(X_1=1, X_2=0) = \)
2. \( P(X_3 =0) = \)
3. \( P(X_1 =0|X_2 =1, X_3 =1) = \)
4. \( P(X_2 = 1|X_3 = 1) = \)
5. \( P(X_1 = 0, X_2 = 1|X_3 = 1) = \)

<table>
<thead>
<tr>
<th>(X_1)</th>
<th>(X_2)</th>
<th>(X_3)</th>
<th>(P(X_1, X_2, X_3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Utilities

- Utilities are functions from outcomes (state of the world) to real numbers that describe an agent’s preferences.
- Utilities summarize the agent’s goals.
- Theorem: any “rational” preferences can be summarized as a utility function.
- Example: Game (+1 / -1)
Decisions

• Agents make actions based on the desirability of their immediate outcomes (state)
  – Buy a lottery
  – Pick a Movie to watch
  – Take a course
  – …

• An agent chooses the action based on its preferences
• Agents’ preference are captured by a utility function, \( U(s) \)
• \( U(s) \) assigns a value to express the desirability of a state
Maximum Expected Utility (MEU)

- $EU(a)$: Expected utility of an action $a$ given the evidence
- $P(\text{Result}(a)=s')$: The probability of reaching $s'$ by doing action $a$ in the current state
- $EU(a) = \sum_{s'} P(\text{Result}(a) = s')U(s')$
- Principle: A rational agent should choose the action that maximizes the agent’s expected utility

$$action = \arg\max_a EU(a)$$
Preferences

- $A > B$: Agent prefers A over B
- $A \sim B$: Agent is indifferent between A and B
- $A \succeq B$

An agent must have preferences among:
- Prizes: A, B, etc.
- Lotteries: uncertain prizes (outcomes)
Constraints

- **Orderability:** \((A \succ B) \lor (B \succ A) \lor (A \sim B)\)
  
  A rational agent must either prefer one of A or B, or be indifferent between the two.

- **Transitivity:** \((A \succ B) \land (B \succ C) \Rightarrow (A \succ C)\)
  
  If a rational agent prefers A to B and B to C, then it prefers A to C.

- **Continuity:** \(A \succ B \succ C \Rightarrow \exists p \left[ p, A; (1-p), C \right] \sim B\)
  
  If a rational agent prefers A to B but B to C, then it’s possible to construct a lottery \(L\) between A and C such that the agent is indifferent between \(L\) and B with appropriate selection of \(p\).

- **Substitutability:** \(A \sim B \Rightarrow [p, A; (1-p), C] \sim [p, B; (1-p), C]\)
  
  A rational agent indifferent between two prizes A and B is also indifferent between any two lotteries which only differ in substitutions of A for B or B for A.

- **Monotonicity:** \(A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; (1-p), B] \succeq [q, A; (1-q), B]\)
  
  If a rational agent prefers A over B, then given a choice between lotteries involving only A and B, the agent prefers the lottery assigning the highest probability to A.
Constraints

- Decomposability
- Axioms of utility theory

\[
[p, A; 1 - p, [q, B; 1 - q, C]] \sim [p, A; (1 - p)q, B; (1 - p)(1 - q), C]
\]
Irrational Behavior

- Nontransitive preferences: $A > B > C > A$
Rational Preferences lead to Utility

• Existence of Utility Function:

\[ U(A) > U(B) \iff A \succ B \]
\[ U(A) = U(B) \iff A \sim B \]

• Expected Utility of a Lottery (EUL):

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

• Agent choose the action that maximizes expected utility
EU and EUL

- \( EUL: U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i) \)
- \( EU(a) = \sum_{s'} P(\text{Result}(a) = s') U(s') \)
Utility Scales

• Normalized utilities: \( u_\perp = 0 \) and \( u_\top = 1 \)
• Micromorts: one-millionth chance of death, useful for paying to reduce product risks, etc.
• QALYs: quality-adjusted life years, useful for medical decisions involving substantial risk
• Expected monetary value (EMV)
  – Gamble: \( \frac{1}{2} ($0) + \frac{1}{2} ($2,500,000) \)

\[
EU(Accept) = \frac{1}{2} U(S_k) + \frac{1}{2} U(S_{k+2,500,000}),
\]
\[
EU(Decline) = U(S_{k+1,000,000}).
\]
Human Rationality

• Allais paradox (Allais, 1953)
  – A: 80% chance of $4000. B: 100% chance of $3000
  – C: 20% chance of $4000. D: 25% chance of $3000
• 90% survival rate sounds better than 10% death rate
• Emotional state, financial state, etc.
Recap

• Player is maximizing expected estimated value
  1. What is Player’s expected value if she takes the expectimax optimal action?
  2. What is the worst possible payoff she could see from that action?

• Player now only considers actions whose worst-case outcome is 10 or better
  3. Which action does the Player choose for this tree?
  4. What is the expected value for that action?
  5. What is the worst value possible for that action?
Reading and Next Class

- Expectimax: AIMA 5.5
- Utilities: AIMA 16.1-16.3
- Next: CSP 6.1-6.2

Search Game

(a) Atomic
(b) Factored
(c) Structured

CSP
Project & Midterm & Final exam

• Project 1: http://courses.cs.vt.edu/cs4804/Fall20/projects/project1.html

• **Mark your calendar!**
  – 10/15 (Thur) Midterm 3:30pm – 5:40pm (Online)
  – 12/16 (Wed) Final exam 1:05pm – 3:05pm (Online)