Other Game Types

- E.g. Backgammon
- Expectiminimax
  - Environment is an extra "random agent" player that moves after each min/max agent
  - Each node computes the appropriate combination of its children

Mixed Layer Types

- Mixed layer types
  - E.g. Backgammon
  - Expectiminimax
  - Environment is an extra "random agent" player that moves after each min/max agent
  - Each node computes the appropriate combination of its children

Expectiminimax Pseudocode

```python
def value(state):
    if the state is a terminal state: return the state's utility
    if the next agent is MAX: return max-value(state)
    if the next agent is MIN: return min-value(state)
    if the next agent is EXP: return exp-value(state)

def exp-value(state):
    initialize v = 0
    for each successor of state:
        p = probability (successor)
        v += p * value(successor)
    return v

def max-value(state):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor))
    return v

def min-value(state):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor))
    return v
```

Example: Backgammon

- Dice rolls increase b: 21 possible rolls with 2 dice
- Backgammon = 22 legal moves
- Depth 2 x 20 x (21 x 20) = 1.3 x 10^6
- As depth increases, probability of reaching a given search node shrinks
- Usefulness of search is diminished
- Trimming depth is less damaging
- But pruning is tricker...
- Historic AI: TDGammon uses depth-2 search + very good evaluation function + reinforcement learning: world-champion level play
- 1st AI world champion in any game!

Multi-Agent Utilities

- What if the game is not zero-sum, or has multiple players?
- Generalization of minimax:
  - Terminals have utility tuples
  - Node values are also utility tuples
  - Each player maximizes its own component
  - Can give rise to cooperation and competition dynamically...

Games Summary

- Games require decisions
  - Optimality is impossible (for most games/problems)
  - Bounded-depth search and evaluation functions
  - Alpha-beta pruning
  - Important advances (from game playing)
    - Reinforcement learning
    - Iterative deepening
    - Monte Carlo tree search
    - Video games
  - Greater challenges
Maximum Expected Utility

- Why should we average utilities? Why not minimax?
- 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?

What Utilities to Use?

- For worst-case minimax reasoning, terminal function scale doesn’t matter
- We just want better states to have higher evaluations (get the ordering right)
- We call this insensitivity to monotonic transformations
- 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

- Getting ice cream
  - Get Single
  - Get Double

Preferences

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

Rationality

- We want some constraints on preferences before we call them rational, such as:
  - Axiom of Transitivity: \((A > B) \land (B > C) \Rightarrow (A > C)\)
- 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 \(C\)

The Axioms of Rationality

- Orderliness
  - \((A > B) \lor (B > A) \lor (A = B)\)
- Transitivity
  - \((A > B) \land (B > C) \Rightarrow (A > C)\)
- Continuity
  - \(A > B > C \Rightarrow \exists \alpha, \beta, \gamma; A > \alpha, \beta > C, \gamma > B\)
  - Substitutability
  - \((A > B) \Rightarrow [p, A, 1-p, C] \sim [p, B, 1-p, C]\)
  - Monotonicity
  - \(A > B \Rightarrow \exists \alpha > \beta > \gamma, A > \alpha, B > \beta, C > \gamma\)

Theorem: Rational preferences imply behavior describable as maximization of expected utility
Theorem [Ramsey, 1931; von Neumann & Morgenstern, 1944]
Given any preferences satisfying these constraints, there exists a real-valued function \( U \) such that:
\[
U(A) \geq U(B) \iff A \succeq B
\]
\[
U([p_1, S_1; \ldots; p_n, S_n]) = \sum p_i U(S_i)
\]
 i.e. values assigned by \( U \) preserve preferences of both prizes and lotteries.

Maximum expected utility (MEU) principle:
- Choose the action that maximizes expected utility
- Note: an agent can be entirely rational (consistent with MEU) without ever representing or manipulating utilities and probabilities
- E.g., a lookup table for perfect tic-tac-toe, a reflex vacuum cleaner

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 \( L = [p, S; (1-p), S'] \)
  \[ U(L) = pU(S) + (1-p)U(S') \]
- Typically, \( U(L) > U(\text{EMV}(L)) \)
- In this sense, people are risk-averse
- When deep in debt, people are risk-prone

Example: Insurance
- Consider the lottery \([0.5, 1000; 0.5, 0]\)
- What is its expected monetary value? \($500\)
- What is its certainty equivalent?
  - Monetary value acceptable in lieu of lottery
  - \$400 for most people
- Difference of \$100 is the insurance premium
- There's an insurance industry because people will pay to reduce their risk
- If everyone were risk-neutral, no insurance needed!
- It's win-win: you'd rather have the \$400 and the insurance company would rather have the lottery (their utility curve is flat and they have many lotteries)

Example: Human Rationality?
- Famous example of Allais (1953)
  - A: \([0.8, 4k; 0.2, 0]\) \(\succ\)
  - B: \([1.0, 3k; 0.0, 0]\)
  - C: \([0.2, 4k; 0.8, 0]\)
  - D: \([0.25, 3k; 0.75, 0]\)
- Most people prefer B \(\succ\) A, C \(\succ\) D
- But if \( U(0) = 0\), then
  - B \(\succ\) A \(\iff\) \(U(3k) + 0.8U(4k) > U(3k)\)
  - C \(\succ\) D \(\iff\) \(U(4k) > 0.8U(4k)\)