Hill Climbing

- Simple, general idea:
  - Start wherever
  - Repeat: move to the best neighboring state
  - If no neighbors better than current, quit

- What’s bad about this approach?
  - Complete?
  - Optimal?

- What’s good about it?

Simulated Annealing

- Idea: Escape local maxima by allowing downhill moves
  - But make them rarer as time goes on

- Theoretical guarantee:
  - Stationary distribution: \( p(x) \propto e^{\frac{E(x)}{T}} \)

  - If \( T \) decreased slowly enough, will converge to optimal state!

  - Is this an interesting guarantee?

  - Sounds like magic, but reality is reality:
    - The more downhill steps you need to escape a local optimum, the less likely you are to ever make them all in a row
    - People think hard about ridge operators which let you jump around the space in better ways
Genetic Algorithms

- Genetic algorithms use a natural selection metaphor
  - Keep best N hypotheses at each step (selection) based on a fitness function
  - Also have pairwise crossover operators, with optional mutation to give variety
- Possibly the most misunderstood, misapplied (and even maligned) technique around

Example: N-Queens

- Why does crossover make sense here?
- When wouldn’t it make sense?
- What would mutation be?

Deterministic Two-Player

- E.g. tic-tac-toe, chess, checkers
- Zero-sum games
  - One player maximizes result
  - The other minimizes result
- Minimax search
  - A state-space search tree
  - Players alternate
  - Choose move to position with highest minimax value = best achievable utility against best play

Tic-tac-toe Game Tree

Minimax Example

Minimax Implementation

```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)

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

def min-value(state):
    initialize v = +∞
    for each successor of state: v = min(v, max-value(successor))
    return v
```
Worst-Case vs. Average Case

Idea: Uncertain outcomes controlled by chance!

Reminder: Probabilities

• A random variable represents an event whose outcome is unknown
• A probability distribution is an assignment of weights to outcomes
• Example: Traffic on freeway
  - Random variable: T = whether there’s traffic
  - Outcomes: T in {none, light, heavy}
  - Distribution: \( P(T=\text{none}) = 0.25, P(T=\text{light}) = 0.50, P(T=\text{heavy}) = 0.25 \)
• Some laws of probability (more later)
  - Probabilities are always non-negative
  - Probabilities over all possible outcomes sum to one

• As we get more evidence, probabilities may change
  - \( P(T=\text{heavy}) = 0.25 \)
  - \( P(T=\text{heavy} | \text{Hour=8am}) = 0.60 \)
  - More about methods for measuring and updating probabilities later

Reminder: 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 airport?

\begin{align*}
\text{Time:} & \quad 20 \text{ min} \quad + \quad 30 \text{ min} \quad + \quad 60 \text{ min} \\
\text{Probability:} & \quad 0.25 \quad + \quad 0.50 \quad + \quad 0.25 \\
\text{Result:} & \quad 35 \text{ min}
\end{align*}
Worst-Case vs. Average Case

Idea: Uncertain outcomes controlled by chance, not an adversary!

What Probabilities to Use?

- In expectimax search, we have a probabilistic model of how the opponent (or environment) will behave in any state.
  - Model could be a simple uniform distribution (e.g., roll a die)
  - Model could be sophisticated and require a great deal of computation
  - We have a chance node for any outcome out of our control: opponent or environment
  - The model might say that adversarial actions are likely!

- For now, assume each chance node magically comes along with probabilities that specify the distribution over its outcomes.

What Probabilities to Use?

Randomness?

- Why wouldn’t we know the results of an action?
  - Explicit randomness: rolling dice
  - Unpredictable opponents: the ghosts respond erratically
  - Actions can fail: when robot moves, its wheels might slip

Expectimax Search

- Values now reflect average-case (expected) outcomes, not worst-case (minimum) outcomes.
- Expectimax search: Compute average score under optimal play.
  - Max nodes as in minimax search
  - Chance nodes are like min nodes but the outcome is uncertain. Calculate their expected utilities
    - i.e., take weighted average (expectation) of children.

Expectimax 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 EXP:
        return exp-value(state)

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

def exp-value(state):
    initialize v = 0
    for each successor of state:
        v = p * v + p * value(successor)
    return v
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 EXP:
        return exp-value(state)
```
### Expectimax Pseudocode

**def exp-value(state):**

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

\[
v = \left(\frac{1}{2}\right)(8) + \left(\frac{1}{3}\right)(24) + \left(\frac{1}{6}\right)(-12) = 10
\]

### Utilities

**Maximum Expected Utility**

- Why should we average utilities?
- 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?

**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 preferences.
  - From them, any “rational” preferences can be summarized as a utility function.
- We hardwire utilities and let behaviors emerge.
  - Why don’t we let agent pick utilities?
  - Why don’t we prescribe behaviors?

**Utilities: Uncertain Outcomes**

- 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: \( A \sim B \)

**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: \( A \sim B \)
Rationality

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

- 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

- Continuity
- Transitivity
- Additivity
- Subadditivity
- Monotonicity

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:

- Maximum expected utility (MEU) principle:
  - Choose the action that maximizes expected utility
  - E.g., a lookup table for perfect tic-tac-toe, a reflex vacuum cleaner

MEU Principle

- Maximum expected utility (MEU) principle:
  - 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

Human Utilities

Playing Russian Roulette?
Playing Russian Roulette?
How much you would pay to avoid a 1 risk?
What value people would place on their own lives?

Perhaps tens of thousands of dollars...??

The actual human behavior reflects a much lower monetary value for a micromort!!!
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 \( L \) between
    - "best possible prize" \( u + \) with probability \( p \)
    - "worst possible catastrophe" \( u - \) with probability \( 1-p \)
  - Adjust lottery probability \( p \) until indifference: \( A \sim L \)
  - Resulting \( p \) is a utility in \([0,1]\)

Pay $30

Utility of Money

- Money plays a significant role in human utility functions
- Usually an agent prefers more money to less

The agent exhibits a monotonic preference for more money

But!

- This does not mean that money behaves as a utility function!
- This does not say anything about preferences between lotteries involving money!

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 \( EMV(L) = pX + (1-p)Y \)
  - Typically, \( U(L) < EMV(L) \)
    - In this sense, people are risk-averse
    - When deep in debt, people are risk-prone

Example:

- In a television game show:
  A) take $1,000,000 prize
  B) gamble on the flip of a coin:
    - If heads nothing
    - If tails get $2,500,000

Which one would you take? A or B?
Example:

• In a television game show:
  A) take $1,000,000 prize
  B) gamble on the flip of a coin:
  • If heads nothing
  • If tails get $2,500,000
• If coin is fair, Expected Monetary Value (EMV) of gamble is:
  \[ EMV = \frac{1}{2} (0) + \frac{1}{2} (2,500,000) = 1,250,000 \]
  \[ \text{more than } $1,000,000 \]

Would you choose B?

Example:

• In a television game show:
  A) take $1,000,000 prize
  B) gamble on the flip of a coin:
  • If heads nothing
  • If tails get $2,500,000
• If coin is fair, Expected Monetary Value (EMV) of gamble is:
  \[ EMV = \frac{1}{2} (0) + \frac{1}{2} (2,500,000) = 1,250,000 \]
  \[ \text{more than } $1,000,000 \]

Utility is not directly proportional to monetary value
Utility(first million) is very high!
Utility(additional million) is smaller!
\[ U(S_k) = 5, \]
\[ U(S_k + $1,000,000) = 8 \]
\[ U(S_k + $2,500,000) = 9 \]
Example: Human Rationality?

• Famous example of Allais (1953)
  • A: [0.8, $4k; 0.2, $0]
  • 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 > A, C > D

• But if U($0) = 0, then
  • B > A \Rightarrow U($3k) > 0.8 U($4k)
  • C > D \Rightarrow 0.8 U($4k) > U($3k)

Recommended

• Risks vs gains
• Probability estimates
• Cognitive architecture
• Much more