Recap: HMMs

- Hidden Markov models (HMMs)
  - Underlying Markov chain over states $X$
  - You only observe outputs (effects) $E$ at each time step
  - Want to reason about the hidden states $X$ given observations $E$

$$P(x_{1:T}, e_{1:T}) = P(x_1)P(e_1|x_1) \prod_{t=2}^{T} P(x_t|x_{t-1})P(e_t|x_t).$$

Recap: Speech Recognition

- Observations are acoustic measurements
  - Real systems:
    - 39 MFCC coefficients
    - Real numbers, modeled with mixtures of multidimensional Gaussians
  - Your projects:
    - 2 real numbers (formant frequencies)
    - Discretized values, discrete conditional probs

Speech Recognition

- States indicate which part of which word we’re speaking
  - Each word broken into phonemes
  - Real systems: context-dependent sub-phonemes
  - Your projects: just one state per phoneme
- Example: Yes/No recognizer

$$P(x | x')$$

- $P(x|x_0) = 0.5$ if $x = x_1$, 0.5 if $x = x_4$, 0 otherwise
- $P(x|x_1) = 0.8$ if $x = x_1$, 0.2 if $x = x_2$, 0 otherwise

Speech Recognition

- Emission probs: distribution over acoustic observations for each phoneme
  - How to learn these? See project 3!

$$P(e | "0") = \begin{bmatrix} 0.1 & 0.2 \end{bmatrix} \ldots$$

$$P(e | "s") = \begin{bmatrix} 0.6 & 0.2 & 0.1 \end{bmatrix} \ldots$$

Example of Hidden Sequences

- For the yes/no recognizer, imagine we hear “yynoo" 
- What are the scores of possible labelings?

$$E$$

- $X$ Low, best?
- VV Low
- V Low
- ZERO
The Viterbi Algorithm

- The Viterbi algorithm computes the best labeling for an observation sequence
  - Incrementally computes best scores for subsequences
  - Recurrence:
    \[
    m_t[x_t] = \max_{x_{t-1}} P(x_{t-1}, x_t, \epsilon_{t-1}) \\
    = \max_{x_{t-1}} P(x_{t-1}, \epsilon_{t-1}) P(x_t | x_{t-1}) P(\epsilon_t | x_t) \\
    = P(\epsilon_t | x_t) \max_{x_{t-1}} P(x_t | x_{t-1}) m_{t-1}[x_{t-1}] \\
    = P(\epsilon_t | x_t) \max_{x_{t-1}} [m_{t-1}[x_{t-1}]] \\
    \]
  - Also store backtraces which record the argmaxes

Example

\[ <s> \quad y \quad s \quad n \quad o \quad <s> \]
\[ e_0 \quad \gamma' \quad \gamma'' \quad \gamma'' \quad \gamma'' \quad \gamma'' \]

Utilities

- So far: talked about beliefs

- Important difference between:
  - Belief about some variables
  - Rational action involving those variables
  - Remember the midterm question?

- Next: utilities

Preferences

- An agent chooses among:
  - Prizes: A, B, etc.
  - Lotteries: situations with uncertain prizes
  \[ L = [p, A; (1 - p), B] \]

- Notation:
  \[ A \succ B \] A preferred over B
  \[ A \sim B \] indifference between A and B
  \[ A \succeq B \] B not preferred over A

Rational Preferences

- We want some constraints on preferences before we call them rational

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

- Preferences of a rational agent must obey constraints.
- These constraints (plus one more) are the axioms of rationality

  \begin{align*}
  \text{Orderability} & \quad (A \succ B) \Rightarrow (B \succ A) \Rightarrow (A \sim B) \\
  \text{Transitivity} & \quad (A \succ B) \Rightarrow (B \succ C) \Rightarrow (A \succ C) \\
  \text{Continuity} & \quad A > B \Rightarrow [p, A; 1 - p, C] \sim B \\
  \text{Substitutability} & \quad A \sim B \Rightarrow [p, A; 1 - p, C] \sim [p, B; 1 - p, C] \\
  \text{Monotonicity} & \quad A > B \Rightarrow (p \geq q \Rightarrow [p, A; 1 - p, B] \succeq [q, A; 1 - q, B]) \\
  \end{align*}

- 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:
    
    $$ U(A) \geq U(B) \iff A \succeq B $$
    
    $$ U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i) $$

- Maximum expected likelihood (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

### Human Utilities

- Utilities map states to real numbers. Which numbers?
- Standard approach to assessment of human utilities:
  - Compare a state $A$ to a standard lottery $L_p$ between
    - “best possible prize” $u_+$ with probability $p$
    - “worst possible catastrophe” $u_-$ with probability $1-p$
  - Adjust lottery probability $p$ until $A \sim L_p$
  - Resulting $p$ is a utility in $[0,1]$

### Utility Scales

- Normalized utilities: $u_+ = 1.0, u_- = 0.0$
- 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
- Note: behavior is invariant under positive linear transformation
  
  $$ U'(x) = k_1 U(x) + k_2 \quad \text{where } k_1 > 0 $$

- With deterministic prizes only (no lottery choices), only ordinal utility can be determined, i.e., total order on prizes

### Money

- Money does not behave as a utility function
- Given a lottery $L$:
  - Define expected monetary value $EMV(L)$
  - Usually $U(L) < U(EMV(L))$
  - I.e., people are risk-averse
- Utility curve: for what probability $p$ am I indifferent between:
  - A prize $x$
  - A lottery $[p, \$M; (1-p), \$0]$ for large $M$?
- Typical empirical data, extrapolated with risk-prone behavior:

### 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-prone, no insurance needed!

### 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, $\$3; 0.75, $\$0$]
  - Most people prefer $B > A, C > D$
  - But if $U(\$0) = 0$, then
    - $B > A \Rightarrow U(\$3k) > 0.8 U(\$4k)$
    - $C > D \Rightarrow 0.8 U(\$4k) > U(\$3k)$
Decision Networks

- Extended BNs
  - Chance nodes (circles, like in BNs)
  - Decision nodes (rectangles)
  - Utility nodes (diamonds)
- Can query to find action with max expected utility
- Online applets if you want to play with these

Value of Information

- Idea: compute value of acquiring each possible piece of evidence
  - Can be done directly from decision network
- Example: buying oil drilling rights
  - Two blocks A and B, exactly one has oil, worth k
  - Prior probabilities 0.5 each, mutually exclusive
  - Current price of each block is k/2
  - "Consultant" offers accurate survey of A. Fair price?
- Solution: compute expected value of information
  = expected value of best action given the information minus expected value of best action without information
- Survey may say "oil in A" or "no oil in A", prob 0.5 each (given!)
  = [0.5 * value of "buy A" given "oil in A"] + [0.5 * value of "buy B" given "no oil in A"]
  = 0
  = [0.5 * k/2] + [0.5 * k/2] - 0 = k/2

General Formula

- Current evidence $E$, current best action $a$
- Possible action outcomes $S_j$, potential new evidence $E_j$

$$ EU(a|E) = \max_{a} \sum_j U(S_j) P(S_j|E, a) $$

- Suppose we knew $E_j = e_j$, then we would choose $a(e_j)$:

$$ EU(a_{e_j}|E, E_j = e_j) = \max_{a} \sum_j U(S_j) P(S_j|E, a, E_j = e_j) $$

- BUT $E_j$ is a random variable whose value is currently unknown, so:
  - Must compute expected gain over all possible values

$$ VPI_E(E_j) = \left( \sum_k P(E_j = e_j|E) EU(a_{e_j}|E, E_j = e_j) \right) - EU(a|E) $$

(VPI = value of perfect information)

VPI Properties

- Nonnegative in expectation
  $$ \forall j, E : VPI_E(E_j) \geq 0 $$

- Nonadditive -- consider, e.g., obtaining $E_j$ twice
  $$ VPI_E(E_j, E_k) \neq VPI_E(E_j) + VPI_E(E_k) $$

- $VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_E, E_j(E_k)$

Next Class

- Start on reinforcement learning!
  - Central idea of modern AI
  - How to learn complex behaviors from simple feedback
  - Basic technique for robotic control
  - Last large technical unit of the course