Recap: Some Simple Cases

Hidden Markov Models

Battleship HMM

Passage of Time
Example: Passage of Time

As time passes, uncertainty “accumulates”

\[ B'(X) = \sum_{x'} P(x' | x) B(x) \]

Transition model: ships usually go clockwise

Example: Observation

As we get observations, beliefs get reweighted, uncertainty “decreases”

\[ B(X) \propto P(e|X) B'(X) \]

Example HMM

The Forward Algorithm

- We are given evidence at each time and want to know
  \[ B_t(X) = P(X_t | e_{1:t}) \]
- We can derive the following updates
  \[ P(x_t | e_{1:t}) \propto P(x_t, e_{1:t}) \]
  \[ = \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t}) \]
  \[ = \sum_{x_{t-1}} P(x_{t-1}, e_{t-1:1}) P(x_t | x_{t-1}) P(e_t | x_t) \]
  \[ = P(e_t | x_t) \sum_{x_{t-1}} P(x_t | x_{t-1}) P(x_{t-1}, e_{1:t-1}) \]
Belief Updates

- Every time step, we start with current $P(X \mid \text{evidence})$
- We update for time:
  $$P(x_t | e_{1:t-1}) = \sum_{x_{t-1}} P(x_t | e_{t-1})P(x_{t-1} | e_{1:t-1})$$
- We update for evidence:
  $$P(x_t | e_{1:t}) \propto P(x_t | e_t) P(x_t | e_{1:t-1})$$
- The forward algorithm does both at once (and doesn’t normalize)
- Problem: space is $|X|$ and time is $|X|^2$ per time step

Particle Filtering

- Sometimes $|X|$ is too big to use exact inference
  - $|X|$ may be too big to even store $B(X)$
  - E.g. $X$ is continuous
  - $|X|^2$ may be too big to do updates
- Solution: approximate inference
  - Track samples of $X$, not all values
  - Time per step is linear in the number of samples
  - But: number needed may be large
- This is how robot localization works in practice

Particle Filtering: Time

- Each particle is moved by sampling its next position from the transition model
  $$x' = \text{sample} \left( P(X' | x) \right)$$
- This is like prior sampling – samples’ frequencies reflect the transition probs
- Here, most samples move clockwise, but some move in another direction or stay in place
- This captures the passage of time
  - If we have enough samples, close to the exact values before and after (consistent)

Particle Filtering: Observation

- Slightly trickier:
  - We don’t sample the observation, we fix it
  - This is similar to likelihood weighting, so we downweight our samples based on the evidence
    $$w(z) = P(e | z)$$
    $$B(X) \propto P(e | X) B'(X)$$
- Note that, as before, the probabilities don’t sum to one, since most have been downweighted (in fact they sum to an approximation of $P(e)$)

Particle Filtering: Resampling

- Rather than tracking weighted samples, we resample
- $N$ times, we choose from our weighted sample distribution (i.e. draw with replacement)
- This is equivalent to renormalizing the distribution
- Now the update is complete for this time step, continue with the next one

Robot Localization

- In robot localization:
  - We know the map, but not the robot’s position
  - Observations may be vectors of range finder readings
  - State space and readings are typically continuous (works basically like a very fine grid) and so we cannot store $B(X)$
  - Particle filtering is a main technique
- [DEMOS]
SLAM

- **SLAM = Simultaneous Localization And Mapping**
  - We do not know the map or our location
  - Our belief state is over maps and positions!
  - Main techniques: Kalman filtering (Gaussian HMMs) and particle methods

[DEMOS]

DP-SLAM, Ron Parr