Reasoning with uncertainty III

\[
P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i | \text{Parents}(X_i))
\]

\[
m_{ji}(x_i) = \max_{x_j} f_{ij}(x_i, x_j) \prod_{\{N(j)\setminus i\}} m_{kj}(x_j)
\]

\[
m_{ji} = \sum_{x_j} f_{ij}(x_i, x_j) \prod_{\{N(j)\setminus i\}} m_{kj}(x_j)
\]

Sum- and max-product, belief propagation
Exact on trees
Junction trees
Complexity w.r.t. treewidth
Temporal Models

- $X_t =$ state of the world at time $t$
- $Y_t =$ Observation (evidence, measurement,..) at time $t$
- Examples:
  - $X =$ types of actions, $Y =$ observations from video
  - $X =$ location/orientation in the world, $Y =$ observations

Assumptions

- Stationarity
  - Technically, infinite set of variables
  - Assume that process does not change over time: the conditional probabilities that define the model do not change over time
- Markov
  - State and measurements depend only on variables within a bounded time window
  - Conditionally independent from all the other variables conditioned on that time window
  
  \begin{align*}
  & P(X_t | X_{[1:t-1]}) = P(X_t | X_{[t-d:t-1]}) \\
  & P(Y_t | X_{[1:t]}, Y_{[0:t-1]}) = P(Y_t | X_t)
  \end{align*}
Example

- First-order:
- From d-separation, $X_{t+1}$ is conditionally independent of $X_{t-1}$ given $X_t$
- No independence of measurements $Y_t$

Alternate view

- Polytree in directed representation
- Tree in factored graph
Operations

• Filtering: $P(\text{current state given all previous observations})$

  $P(X_{t+1}|y_{1:t+1}) \propto P(y_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t) P(x_t|y_{1:t})$

• Prediction: $P(\text{future states given all previous observations})$

  $P(X_{t+1}|y_{1:t}) \propto \sum_{x_t} P(X_{t+1}|x_t) P(x_t|y_{1:t})$

• Smoothing: $P(\text{past state given all observations})$

  $P(X_k|y_{1:k}) \propto P(X_k|y_{1:k}) P(y_{k+1:t}|X_k)$

  $P(y_{k+1:t}|X_k) = \sum_{x_{k+1}} P(y_{k+1}|x_{k+1}) P(x_{k+1}|X_k) P(y_{k+2:t}|x_{k+1})$
\[ P(X_{t+L+1}|y_{1:t}) \alpha \sum_{x_t} P(X_{t+L+1}|x_{t+L})P(x_{t+L}|y_{1:t}) \]

- Further prediction?
- No added evidence
- Stationary distribution

\[ P(X_{t+1}|y_{1:t+1}) \alpha P(y_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)f_t \]

\[ f_t = P(x_t|y_{1:t}) \]

- Estimate from past: Propagate forward in time
- Estimate from future: Propagate backward in time

\[ b_{k+1:t} = P(y_{k+1:t} | X_k) \]

\[ b_{k+1:t} = \sum_{x_{k+1}} P(y_{k+1} | x_{k+1}) P(x_{k+1} | X_k) b_{k+2:t} \]

- Estimate from both: Combine forward and backward terms

\[ P(X_k | y_{1:t}) \alpha f_{k} b_{k+1:t} \]

• Inference over 1:t linear in number of states
• (Polytree case in which inference is efficient; sum-product)
Inferring the most likely set of states

Same as max-product vs. sum-product

\[
\max_{x_1,\ldots,x_t} P(x_1,\ldots,x_t,X_{t+1}|y_{1:t+1}) \alpha \\
= P(y_{t+1}|X_{t+1}) \max_{x_t} (P(X_{t+1}|x_t) \max_{x_1,\ldots,x_{t-1}} P(x_1,\ldots,x_{t-1},x_t|y_{1:t})) \\
= m_{1:t+1} \alpha \\
= P(y_{t+1}|X_{t+1}) \max_{x_t} (P(X_{t+1}|x_t)m_{1:t})
\]

Special case: HMM

- \( X \) discrete with \( K \) states
- \( K^2 \) transitions: \( P(X_{t+1}|X_t) \)
- Emission probabilities \( P(Y_t|X_t) \)
Special case: HMM

\[ m_{1:t}(x_t) = \max_{x_{1:t-1}} P(x_1, \ldots, x_t | y_{1:t}) \]

\[ m_{1:t+1}(x_{t+1}) = P(y_{t+1} | x_{t+1}) \max_{x_t} (P(x_{t+1} | x_t)m_{1:t}) \]
Special case: HMM

\[ ( \quad | \quad ) \quad ( \quad | \quad ) \rightarrow \text{DP} \]

- Special case of max-product (same as tree example the last time)
- Time in general (excepting special form of \( ( \quad | \quad ) \))

Example

- Observations: (Noisy) estimates of link locations
- States: Actions executed at each time step

Nazli Ikizler and David Forsyth, “Searching video for complex activities with finite state models” IEEE Conference on Computer Vision and Pattern Recognition, 2007
• Different action model for each part (learned from motion capture data)

Nazli Ikizler and David Forsyth, “Searching video for complex activities with finite state models” IEEE Conference on Computer Vision and Pattern Recognition, 2007
Nazli Ikizler and David Forsyth, “Searching video for complex activities with finite state models” IEEE Conference on Computer Vision and Pattern Recognition, 2007
• Longer-range connections (e.g., tracking, ..)
• No problem (in principle): Update representation of $X^i$ and $X^j$ are not separated
• Need to represent all states
• Grouping states does not solve the problem: Standard HMM back/forward is $Nk^{2D}$
Linear dynamical models

- Canonical case: \( x = [u, v, \dot{u}, \dot{v}] \)
- What parametric representations?
- Desirable property (closure):
  \[
P(X_{t+1}|y_{1:t+1}) \propto P(y_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|y_{1:t})
\]
  and
  \[
P(X_t|y_{1:t})
\]
  must be of the same form (only parameters change)
- Need to use exponential family
- Gaussian model \( N(\mu, \Sigma) \) satisfies closure
- Other models grow in complexity without bounds

- Linear transform of variable \( A \sim N(\mu, \Sigma) \) to \( MA + U \) \( U \sim N(0, \Gamma) \) yields \( N(M\mu, M\Sigma M^T + \Gamma) \)
Linear dynamical models

\[ P(X_{t+1}|X_t) = N(AX_t, \Sigma) \]
\[ P(Y_t|X_t) = N(BX_t, \Gamma) \]

- Equivalent to:
  \[ X_{t+1} = AX_t + w \quad w \sim N(0, \Sigma) \]
  \[ Y_t = BX_t + v \quad v \sim N(0, \Gamma) \]

E.g., constant velocity:

- \( x = [u, v, \dot{u}, \dot{v}] \)
- \( A = [I \Delta t; 0 I] \)

Linear dynamic systems

Gaussian model propagates through the chain:

\[ P(x_{t+1}|y_{1:t+1}) \propto \sum_{x_t} P(x_{t+1}|x_t)P(x_t|y_{1:t}) \]
\[ P(x_{t+1}|y_{1:t+1}) \propto \int P(x_{t+1}|x_t)P(x_t|y_{1:t})dx_t \]
\[ P(x_t|y_{1:t}) \sim N(\mu_t, U_t) \]

\[ \mu_{t+1} = A\mu_t + K(y_{t+1} - BA\mu_t) \]
\[ V_t = AU_tA^T + \Sigma \]
\[ K = V_tB^T(BV_tB^T + \Gamma)^{-1} \]
\[ U_{t+1} = (I - KB)V_t \]
Linear dynamic systems

Gaussian model propagates through the chain:

\[ P(x_{t+1}|y_{1:t+1}) \propto P(y_{t+1}|x_{t+1}) \sum_{x_t} P(x_{t+1}|x_t) P(x_t|y_{1:t}) \]

\[ P(x_{t+1}|y_{1:t+1}) \propto P(y_{t+1}|x_{t+1}) \int P(x_{t+1}|x_t) P(x_t|y_{1:t}) dx_t \]

\[ P(x_t|y_{1:t}) \sim N(\mu_t, U_t) \]

**Innovation**

Gain = \( B^{-1} \) if trust new observation
0 if trust \( X_t \)

\[ \mu_{t+1} = A\mu_t + K(y_{t+1} - BA\mu_t) \]

\[ V_t = AU_tA^T + \Sigma \]

\[ K = V_t B^T (BV_t B^T + \Gamma)^{-1} \]

\[ U_{t+1} = (I - KB)V_t \]

**Non linear case:**

\[ x_{t+1} = f(X_t) + w \quad w \sim N(0, \Sigma) \]

**First-order approximation**

\[ f(X) \sim f(\mu) + J(X - \mu) \]

**Local approximation**

\[ -JX_t \]

**Local linear approximation can be very bad (e.g, \( X = [x \ y \ \theta] \))**

**Other possibility:** Sample \( U_t \), transform the samples, estimate \( V_t \) from the samples
More general case

- Arbitrary connections between state and observation variables at any time $t$
  1. Replicate over time (unroll) $\rightarrow$ General graph, can’t do exact inference directly (in general)
  2. Collapse state variables wrt observed $\rightarrow$ $K^2$ state tables in general
- In the discrete case, DBN $\iff$ HMM but note the complexity issue
Example

- Observations: IDs of (60) objects manipulated (RFID tags)
- State: Activity performed (11 fine-grained activities requiring extensive observations)
- Hypothesis: “invisible human hypothesis”


- Baselines:
  A. Each activity has its own HMM (11 HMMs) → take the best
  B. A single HMM for all the activities (11-valued states)
  C. A single HMM with state = activities x objects (660-valued states)
More complicated relations:
- The number of objects is an indication of the type of activities (setting the table vs. eating breakfast)
- E node (Exit) indicates end of previous activity
- AD node (Aggregate Distribution)
• Does not scale well (660^2 tables) for C
• Better representation of relations in D
• Hierarchical representation of object list to address robustness issues?

<table>
<thead>
<tr>
<th></th>
<th>Time-Slice</th>
<th>Edit Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mu (Mean)</td>
<td>68%</td>
<td>12%</td>
</tr>
<tr>
<td>Sigma (Standard Deviation)</td>
<td>(5.9)</td>
<td>(2.9)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>88%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>(4.2)</td>
<td>(6.2)</td>
</tr>
<tr>
<td></td>
<td>87%</td>
<td>14%</td>
</tr>
<tr>
<td></td>
<td>(9.3)</td>
<td>(10.4)</td>
</tr>
<tr>
<td></td>
<td>88%</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>(3.1)</td>
<td>(2.2)</td>
</tr>
</tbody>
</table>