Remedial Probability

CS267A - Fall 2018
Guy Van den Broeck
World View

• Propositional
  Global properties that are true or false

• Probabilistic
  – Belief is still a set of possible worlds
  – But now they have a degree of belief Pr(.)
  – Knowledge Base KB ≈ Pr

• “Uncertainty is epistemological – pertaining to an agent’s beliefs of the world – rather than ontological – how the world is.” [Poole et al.]
  – We can have different beliefs about the same world
  – What’s the probability that the world ends tomorrow?
### Possible Worlds

<table>
<thead>
<tr>
<th><em>world</em></th>
<th>Earthquake</th>
<th>Burglary</th>
<th>Alarm</th>
<th>Pr(.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_1$</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>.0190</td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>true</td>
<td>true</td>
<td>false</td>
<td>.0010</td>
</tr>
<tr>
<td>$\omega_3$</td>
<td>true</td>
<td>false</td>
<td>true</td>
<td>.0560</td>
</tr>
<tr>
<td>$\omega_4$</td>
<td>true</td>
<td>false</td>
<td>false</td>
<td>.0240</td>
</tr>
<tr>
<td>$\omega_5$</td>
<td>false</td>
<td>true</td>
<td>true</td>
<td>.1620</td>
</tr>
<tr>
<td>$\omega_6$</td>
<td>false</td>
<td>true</td>
<td>false</td>
<td>.0180</td>
</tr>
<tr>
<td>$\omega_7$</td>
<td>false</td>
<td>false</td>
<td>true</td>
<td>.0072</td>
</tr>
<tr>
<td>$\omega_8$</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>.7128</td>
</tr>
</tbody>
</table>
Propositions are only Boolean?

• Categorical variables
  – Weather=sunny, Weather=rainy, Weather=snowy
  – 3 Boolean variables that are mutually exclusive
    • Sometimes called “indicator variables”
    • Can all be encoded in sentences...

• Continuous variables
  – Temperature=73.514, Temperature=78.785, ...
  – Infinitely many Boolean variables (and worlds).
    • In logic, see SAT Modulo Theories (SMT)
    • Special accommodations for continuous variables in statistics; we will mostly stick to the discrete world.
Sentences or “Events”

• Knowledge is a probability for every world: $\Pr(\omega)$
• What is the probability of a sentence $\alpha$? (also called an “event” $\alpha$ in probability)
• Need to **axiomatize** probability [Kolmogorov]:
  1. Probabilities are non-negative: $0 \leq \Pr(\alpha)$
  2. The probability of a true event is 1: $\Pr(\text{true}) = 1$
  3. If $\alpha$ and $\beta$ are mutually exclusive, then $\Pr(\alpha \lor \beta) = \Pr(\alpha) + \Pr(\beta)$. 
Sentences or “Events”

• Knowledge is a probability for every world: \( \Pr(\omega) \)
• What is the probability of a sentence \( \alpha \)?
  (also called an “event” \( \alpha \) in probability)
• A sentence \( \alpha \) is equivalent to the disjunction of its models: \( \alpha \equiv \omega_1 \lor \omega_8 \lor \omega_{11} \lor \omega_{17} \lor \cdots \)

\[
\Pr(\alpha) = \sum_{\omega \models \alpha} \Pr(\omega) = \sum_{\omega \in \text{Mods}(\alpha)} \Pr(\omega)
\]
Properties of Probability

• Complement events
  \(- \Pr(\alpha) + \Pr(\neg \alpha) = 1\)
  \(- \text{Why?}\)

• Inclusion-exclusion
  \(- \Pr(\alpha \lor \beta) = \Pr(\alpha) + \Pr(\beta) - \Pr(\alpha \land \beta)\)
  \(- \text{Why?}\)

\[
\begin{align*}
\Pr(\text{Earthquake}) &= \Pr(\omega_1) + \Pr(\omega_2) + \Pr(\omega_3) + \Pr(\omega_4) = .1 \\
\Pr(\text{Burglary}) &= \Pr(\omega_1) + \Pr(\omega_2) + \Pr(\omega_5) + \Pr(\omega_6) = .2 \\
\Pr(\text{Earthquake} \land \text{Burglary}) &= \Pr(\omega_1) + \Pr(\omega_2) = .02 \\
\Pr(\text{Earthquake} \lor \text{Burglary}) &= .1 + .2 - .02 = .28
\end{align*}
\]
Conditional Probability

• What if I observe new information in the form of a sentence $\beta$?
• Belief changes from $\Pr(\alpha)$ to $\Pr(\alpha|\beta)$
• Can also be axiomatized...
• But briefly

\[
\Pr(\alpha|\beta) = \frac{\Pr(\alpha \land \beta)}{\Pr(\beta)}
\]

Pr(Burglary) = .2
Pr(Burglary|Earthquake) = .2
Pr(Alarm) = .2442
Pr(Alarm|Earthquake) $\approx$ .75

[Darwiche ‘08]
Monotonicity of Belief?

- Recall: monotonicity of logic
- Is it possible to observe something new and undo prior beliefs?

\[
\begin{align*}
\Pr(\text{Alarm}) &= .2442 \\
\Pr(\text{Alarm} | \neg \text{Earthquake}) &\approx .75 \\
\end{align*}
\]

Example:
- Alarm and not Earthquake: \(.1620 + .0072 = 0.1692\)
- Not Earthquake: \(.9\)
- Alarm given not Earthquake: \(.188\)

[Darwiche ’08]
Product Rule
Computing Probabilities: Example

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th></th>
<th>¬toothache</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>catch</td>
<td>¬catch</td>
<td>catch</td>
<td>¬catch</td>
</tr>
<tr>
<td>cavity</td>
<td>0.108</td>
<td>0.012</td>
<td>0.072</td>
<td>0.008</td>
</tr>
<tr>
<td>¬cavity</td>
<td>0.016</td>
<td>0.064</td>
<td>0.144</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Motivation: Maximum Expected Utility
Betting Semantics
Inconsistent Beliefs

<table>
<thead>
<tr>
<th>Agent 1 Proposition</th>
<th>Agent 1 Belief</th>
<th>Agent 2 Bet</th>
<th>Agent 2 Stakes</th>
<th>Outcomes and payoffs to Agent 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>0.4</td>
<td>( a )</td>
<td>4 to 6</td>
<td>(-6) (-6) (4) (4)</td>
</tr>
<tr>
<td>( b )</td>
<td>0.3</td>
<td>( b )</td>
<td>3 to 7</td>
<td>(-7) (3) (-7) (3)</td>
</tr>
<tr>
<td>( a \lor b )</td>
<td>0.8</td>
<td>( \neg(a \lor b) )</td>
<td>2 to 8</td>
<td>2 (2) (2) (-8)</td>
</tr>
</tbody>
</table>

\(-11\) \(-1\) \(-1\) \(-1\)
Independence

Cavity
Toothache  Catch
Weather

decomposes into

Cavity
Toothache  Catch

Weather
Independence

$\text{Coin}_1 \hspace{1cm} \ldots \ldots \hspace{1cm} \text{Coin}_n$

decomposes into

$\text{Coin}_1 \hspace{1cm} \ldots \ldots \hspace{1cm} \text{Coin}_n$
Notation: Events or Variables
Conditional Independence
Bayes Rule
Naïve Bayes Assumption

\[ P(Cause, Effect_1, \ldots, Effect_n) = P(Cause) \prod_{i} P(Effect_i | Cause) \]

This is how spam filters work!
Bayesian Networks

Weather

Cavity

Toothache

Catch
Background

• https://artint.info/2e/html/ArtInt2e.Ch8.html