Agents and Uncertainty
CS7032: AI for IET

Liliana Mamani Sanchez
lmamanis@tcd.ie

October 27, 2015
Agent architectures and uncertainty

- How does uncertainty arise in $\text{Arch}_s = \langle S, A, action, env \rangle$?
- Incompleteness or incorrectness of an agent’s understanding of the properties of $S$
- The rationality of $\text{action}(.)$ depends on
  - the relative importance of the agent’s goals, and
  - the likelihood that those goals will be achieved (under various configurations of $S$)
Let action $A_t =$ leave for airport $t$ minutes before flight. Will $A_t$ get me there on time?

Recall the issues the agent has to deal with:

1. **Partial observability** (road state, other drivers’ plans, etc.)
2. **Noisy sensors** (traffic reports)
3. **Uncertainty** in action outcomes (flat tire, etc.)
4. **Immensely complex** of modelling and predicting traffic
Example (ctd.)

- Hence a purely logical approach either
  1. risks falsehood: “$A_{25}$ will get me there on time”
     or
  2. leads to conclusions that are too weak for decision making:
     “$A_{25}$ will get me there on time if there’s no accident on the
     bridge and it doesn’t rain and my tires remain intact etc etc.”

- ($A_{1440}$ might reasonably be said to get me there on time but
  I’d have to stay overnight in the airport . . .)
Encoding uncertainty

Knowledge representation in standard logic:

\[ \forall p \ Symptom(p, \ ToothAche) \Rightarrow \ Disease(p, \ Cavity) \]
Encoding uncertainty

- Knowledge representation in standard logic:
  \[
  \forall p \text{ Symptom}(p, \text{ToothAche}) \Rightarrow \text{Disease}(p, \text{Cavity})
  \]

- ...but the consequent needs to be exhaustively specified:
  \[
  \forall p \text{ Symptom}(p, \text{ToothAche}) \Rightarrow \text{Disease}(p, \text{Cavity})
  \]

- We could try working with causal rules:
  \[
  \forall p \text{ Disease}(p, \text{Cavity}) \Rightarrow \text{Symptom}(p, \text{ToothAche})
  \]

- ... but not all cavities cause pain
Knowledge representation in standard logic:

\[ \forall p \ Symptom(p, \ ToothAche) \Rightarrow Disease(p, \ Cavity) \]

...but the consequent needs to be exhaustively specified:

\[ \forall p \ Symptom(p, \ ToothAche) \Rightarrow Disease(p, \ Cavity) \lor Disease(p, \ GumDisease) \lor Disease(p, \ ImpactedWisdom) \lor \ldots \]
Encoding uncertainty

- Knowledge representation in standard logic:
  \[ \forall p \, \text{Symptom}(p, \text{ToothAche}) \Rightarrow \text{Disease}(p, \text{Cavity}) \]
- …but the consequent needs to be exhaustively specified:
  \[ \forall p \, \text{Symptom}(p, \text{ToothAche}) \Rightarrow \text{Disease}(p, \text{Cavity}) \lor \text{Disease}(p, \text{GumDisease}) \lor \text{Disease}(p, \text{ImpactedWisdom}) \lor \ldots \]
Encoding uncertainty

► Knowledge representation in standard logic:
\[ \forall p \ Symptom(p, \text{ToothAche}) \Rightarrow Disease(p, \text{Cavity}) \]

► ...but the consequent needs to be exhaustively specified:
\[ \forall p \ Symptom(p, \text{ToothAche}) \Rightarrow Disease(p, \text{Cavity}) \lor Disease(p, \text{GumDisease}) \lor Disease(p, \text{ImpactedWisdom}) \lor \ldots \]

► We could try working with causal rules:
\[ \forall p \ Disease(p, \text{Cavity}) \Rightarrow Symptom(p, \text{ToothAche}) \]
Encoding uncertainty

- Knowledge representation in standard logic:
  \[ \forall p \ Symptom(p, \ ToothAche) \Rightarrow \ Disease(p, \ Cavity) \]

- ...but the consequent needs to be exhaustively specified:
  \[ \forall p \ Symptom(p, \ ToothAche) \Rightarrow \ Disease(p, \ Cavity) \lor \ Disease(p, \ GumDisease) \lor \ Disease(p, \ ImpactedWisdom) \lor ... \]

- We could try working with causal rules:
  \[ \forall p \ Disease(p, \ Cavity) \Rightarrow \ Symptom(p, \ ToothAche) \]

- ... but not all cavities cause pain
Why standard logic rules fail

According to [?], they fail due to:

- **Laziness**: it may be too hard to specify (and verify) all relevant rules
- **Theoretical ignorance**: a complete theory for the domain may not exist
- **Practical ignorance**: it may not be practical to test all possible relevant sensorial input (in order to verify the antecedents of the rules above, for instance)
D. Rumsfeld’s contribution

Donald Rumsfeld made the distinction between known unknowns — the things we know we do not know — and unknown unknowns — the things we do not know we do not know. Mr Rumsfeld’s musings won awards for gobbledygook. Yet this remark was one of the wisest things the former US defence secretary said, although the competition is not intense. The distinction between known and unknown unknowns aptly expresses the distinction between risk and uncertainty [...]

(John Kay, Financial Times, 16 Oct 2007)
Methods for handling uncertainty (and risk)

- **Probability theory**
  Given the available evidence, 
  $A_{25}$ will get me there on time with probability 0.04

- **Default or nonmonotonic logic:**
  Assume my car does not have a flat tire
  Assume $A_{25}$ works unless contradicted by evidence
  Issues: What assumptions are reasonable? How to handle contradiction?

- **Rules with fudge factors:**
  $A_{25} \mapsto_{0.3}$ get there on time
  $Sprinkler \mapsto_{0.99} WetGrass$, $WetGrass \mapsto_{0.7} Rain$
  Issues: Problems with combining evidence, e.g., $Sprinkler$ causes $Rain$?

- **Fuzzy logic:** encode degrees of truth e.g., $WetGrass$ is true to degree 0.2
Methods for handling uncertainty (and risk)

- **Probability theory**
  Given the available evidence,
  \( A_{25} \) will get me there on time with probability 0.04

- **Default or nonmonotonic logic:**
  Assume my car does not have a flat tire
  Assume \( A_{25} \) works unless contradicted by evidence
  Issues: What assumptions are reasonable? How to handle contradiction?

- **Rules with fudge factors:**
  \( A_{25} \mapsto_{0.3} \) get there on time
  \( Sprinkler \mapsto_{0.99} WetGrass, \; WetGrass \mapsto_{0.7} Rain \)
  Issues: Problems with combining evidence, e.g., \( Sprinkler \) causes \( Rain \)??

- **Fuzzy logic:** encode degrees of truth e.g., \( WetGrass \) is true to degree 0.2
Probabilities

- Similar to propositional logic: possible worlds defined by assignment of values to random variables.
- **Propositional** or **Boolean** random variables e.g., \textit{Cavity} (do I have a cavity?)
  - Include propositional logic expressions e.g., \(\neg \text{Burglary} \lor \text{Earthquake}\)
- **Multivalued** random variables e.g., \textit{Weather} is one of \(\langle \text{sunny}, \text{rain}, \text{cloudy}, \text{snow}\rangle\)
  - Values must be exhaustive and mutually exclusive
- Proposition constructed by assignment of a value: e.g., \(\text{Weather} = \text{sunny}\); also \(\text{Cavity} = \text{true}\) for clarity
Prior or unconditional probabilities of propositions e.g., $P(Cavity) = 0.1$ and $P(Weather = sunny) = 0.72$ correspond to belief prior to arrival of any (new) evidence.

- Probability distribution gives values for all possible assignments: $P(Weather) = \langle 0.72, 0.1, 0.08, 0.1 \rangle$ (normalised, i.e., sums to 1)

- Joint probability distribution for a set of variables gives values for each possible assignment to all the variables $P(Weather, Cavity) = a 4 \times 2$ matrix of values:

\[
\begin{array}{c|cccc}
\text{Weather} = & \text{sunny} & \text{rain} & \text{cloudy} & \text{snow} \\
\hline
\text{Cavity} = \text{true} & & & & \\
\text{Cavity} = \text{false} & & & & \\
\end{array}
\]
Conditionals and computational shortcuts

- Conditional or posterior probabilities e.g.,
  \[ P(\text{Cavity} | \text{Toothache}) = 0.8 \]
  i.e., given that Toothache is all I know.

- Notation for conditional distributions:
  \[ P(\text{Weather} | \text{Earthquake}) = 2\text{-element vector of 4-element vectors} \]

- If we know more, e.g., Cavity is also given, then we have
  \[ P(\text{Cavity} | \text{Toothache}, \text{Cavity}) = 1 \]

- New evidence may be irrelevant, allowing simplification, e.g.,
  \[ P(\text{Cavity} | \text{Toothache}, \text{Rain}) = P(\text{Cavity} | \text{Toothache}) = 0.8 \]

This kind of inference, sanctioned by domain knowledge, is crucial.
Axioms of probability

- **Notational variants:**

<table>
<thead>
<tr>
<th>Logic</th>
<th>Set theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(A ∨ B)$</td>
<td>$P(A ∪ B)$</td>
</tr>
<tr>
<td>$P(A ∧ B)$ or</td>
<td>$P(A ∩ B)$</td>
</tr>
<tr>
<td>$P(A, B)$</td>
<td></td>
</tr>
<tr>
<td>$P(\text{false}),$</td>
<td>$P(\emptyset),$</td>
</tr>
<tr>
<td>$P(\text{true})$</td>
<td>$P(\Omega)$</td>
</tr>
</tbody>
</table>

- **Axioms:** For any propositions $A, B$
  1. $0 \leq P(A) \leq 1$
  2. $P(\text{true}) = 1$ and $P(\text{false}) = 0$
  3. $P(A ∨ B) = P(A) + P(B) - P(A ∧ B)$
Conditional probabilities

\[ P(A|B) = \frac{P(A \land B)}{P(B)} \text{ if } P(B) \neq 0 \]

Product rule (PR) gives an alternative formulation:
\[ P(A \land B) = P(A|B)P(B) = P(B|A)P(A) \]

A general version holds for whole distributions, e.g.,
\[ P(Weather, Cavity) = P(Weather|Cavity)P(Cavity) \]
(View as a 4 × 2 set of equations, not matrix mult.)

Chain rule is derived by successive application of PR:
\[ P(X_1, \ldots, X_n) = P(X_1, \ldots, X_{n-1}) P(X_n|X_1, \ldots, X_{n-1}) \]
\[ = P(X_1, \ldots, X_{n-2}) P(X_{n-1}|X_1, \ldots, X_{n-2}) P(X_n|X_1, \ldots, X_{n-1}) \]
\[ = \ldots \]
\[ = \prod_{i=1}^{n} P(X_i|X_1, \ldots, X_{i-1}) \]
Bayes’ rule

- Product rule $P(A \land B) = P(A|B)P(B) = P(B|A)P(A)$

  $\implies$ Bayes’ rule $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

- Why is this useful?

  For assessing diagnostic probability from causal probability:

  $$P(\text{Cause} | \text{Effect}) = P(\text{Effect} | \text{Cause})P(\text{Cause})P(\text{Effect})$$

  E.g., let $M$ be meningitis, $S$ be stiff neck:

  $$P(M | S) = \frac{P(S | M)P(M)}{P(S)} = 0.0001 \times 0.0008 = 0.0000008$$
Bayes’ rule

- Product rule \( P(A \land B) = P(A|B)P(B) = P(B|A)P(A) \)

\[ \implies \text{Bayes’ rule } P(A|B) = \frac{P(B|A)P(A)}{P(B)} \]

- Why is this useful? For assessing diagnostic probability from causal probability:

\[ P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)} \]
Bayes’ rule

- Product rule $P(A \land B) = P(A|B)P(B) = P(B|A)P(A)$

$$\implies \text{Bayes’ rule } P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- Why is this useful? For assessing diagnostic probability from causal probability:

$$P(\text{Cause}|\text{Effect}) = \frac{P(\text{Effect}|\text{Cause})P(\text{Cause})}{P(\text{Effect})}$$

E.g., let $M$ be meningitis, $S$ be stiff neck:

$$P(M|S) = \frac{P(S|M)P(M)}{P(S)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$$
Normalisation

- Suppose we wish to compute a posterior distribution over $A$ given $B = b$, and suppose $A$ has possible values $a_1 \ldots a_m$.
- We can apply Bayes’ rule for each value of $A$:

$$
P(A = a_1 | B = b) = \frac{P(B = b | A = a_1) P(A = a_1)}{P(B = b)}
$$

$$
\ldots
$$

$$
P(A = a_m | B = b) = \frac{P(B = b | A = a_m) P(A = a_m)}{P(B = b)}
$$

- Summing these, and noting that $\sum_i P(A = a_i | B = b) = 1$, we get the normalisation factor:

$$
1/P(B = b) = 1/\sum_i P(B = b | A = a_i) P(A = a_i) \quad (1)
$$
Normalisation (ctd)

The normalisation factor is constant w.r.t. $i$. If we denote it by $\alpha$, we can rewrite Bayes’ rule for the general, multi-valued case as follows:

$$P(A|B = b) = \alpha P(B = b|A)P(A)$$

Typically compute an unnormalized distribution, normalize at the end. E.g.:

$$P(B = b|A)P(A) = \langle 0.4, 0.2, 0.2 \rangle$$

then

$$P(A|B = b) = \alpha \langle 0.4, 0.2, 0.2 \rangle$$

$$= \frac{\langle 0.4, 0.2, 0.2 \rangle}{0.4 + 0.2 + 0.2}$$

$$= \langle 0.5, 0.25, 0.25 \rangle$$
Conditioning

Introducing a variable as an extra condition:

\[ P(X|Y) = \sum_z P(X|Y, Z = z)P(Z = z|Y) \]

Intuition: often easier to assess each specific circumstance, e.g.:

\[ P(\text{RunOver}|\text{Cross}) = P(\text{RunOver}|\text{Cross}, \text{Light} = \text{green})P(\text{Light} = \text{green}|\text{Cross}) \]
\[ + P(\text{RunOver}|\text{Cross}, \text{Light} = \text{yellow})P(\text{Light} = \text{yellow}|\text{Cross}) \]
\[ + P(\text{RunOver}|\text{Cross}, \text{Light} = \text{red})P(\text{Light} = \text{red}|\text{Cross}) \]

When \( Y \) is absent, we have summing out or marginalisation:

\[ P(X) = \sum_z P(X|Z = z)P(Z = z) = \sum_z P(X, Z = z) \]
Full joint distributions

A complete probability model specifies every entry in the joint distribution for all the variables $\mathbf{X} = X_1, \ldots, X_n$

I.e., a probability for each possible world $X_1 = x_1, \ldots, X_n = x_n$

E.g., suppose *Toothache* and *Cavity* are the random variables:

<table>
<thead>
<tr>
<th></th>
<th>Toothache = true</th>
<th>Toothache = false</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavity = true</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Cavity = false</td>
<td>0.01</td>
<td>0.89</td>
</tr>
</tbody>
</table>

▶ Possible worlds are mutually exclusive $\implies P(w_1 \land w_2) = 0$

▶ Possible worlds are exhaustive $\implies w_1 \lor \cdots \lor w_n$ is True

hence $\sum_i P(w_i) = 1$
For any proposition $\phi$ defined on the random variables $\phi(w_i)$ is true or false

$\phi$ is equivalent to the disjunction of $w_i$s where $\phi(w_i)$ is true

Hence $P(\phi) = \sum_{\{w_i: \phi(w_i)\}} P(w_i)$

I.e., the unconditional probability of any proposition is computable as the sum of entries from the full joint distribution
Conditional probabilities can be computed in the same way as a ratio:

\[
P(\phi | \xi) = \frac{P(\phi \land \xi)}{P(\xi)}
\]

For example:

\[
P(\text{Cavity} | \text{Toothache}) = \frac{P(\text{Cavity} \land \text{Toothache})}{P(\text{Toothache})}
\]

\[= \frac{0.04}{0.04 + 0.01}
\]

\[= 0.8
\]
Inference from joint distributions

- Typically, we are interested in the posterior joint distribution of the
  - query variables $\mathbf{X}$ given
  - specific values $\mathbf{e}$ for the evidence variables $\mathbf{E}$
Inference from joint distributions

- Typically, we are interested in the posterior joint distribution of the
  - query variables $X$ given
  - specific values $e$ for the evidence variables $E$
- Let the hidden variables be $H = V \setminus (X \cup E)$
Inference from joint distributions

- Typically, we are interested in the posterior joint distribution of the
  - query variables $X$ given
  - specific values $e$ for the evidence variables $E$
- Let the hidden variables be $H = V \setminus (X \cup E)$
- Then the required summation of joint entries is done by summing out the hidden variables:

$$P(X|E = e) = \alpha P(X, E = e) = \alpha \sum_{h} P(X, E = e, H = h)$$

The terms in the summation are joint entries because $X$, $E$, and $H$ together exhaust the set of random variables $V$
Inference by enumeration

$\text{Enumeration}(X, e, P)$ returns a distribution over $X$

1. inputs: $X$, the query variable
2. $e$, evidence specified as an event
3. $P$, a joint distrib on $\{X\} \cup E \cup H$ /* $H =$ hidden vars */
4. $Q(X) \leftarrow$ a distribution over $X$, initially empty
5. for each value $x_i$ of $X$ do
6. \hspace{1em} $Q(x_i) \leftarrow \text{EnumerateJoint}(x_i, e, H, [], P)$
7. return $\text{Normalize}(Q(X))$

$\text{EnumerateJoint}(x, e, vars, values, P)$ returns a real number

1. if $\text{Empty?}(vars)$ then return $P(x, e, values)$
2. $Y \leftarrow \text{First}(vars)$
3. return $\sum_h \text{EnumerateJoint}(x, e, \text{Rest}(vars), [h|values], P)$
Limitations of inference by enumeration

- Worst-case time complexity $O(d^n)$ where $d$ is the largest arity
- Space complexity $O(d^n)$ to store the joint distribution
- How to find the numbers for $O(d^n)$ entries?