Probabilistic Reasoning in First-Order Logic

- Set of possible worlds
  (be they represented as a full joint distribution or belief network)

- Each world ω has a probability $P(\omega)$

- Taking any sentence ϕ we can compute its probability:

$$P(\phi) = \sum_{\omega : \phi \; is \; true \; in \; \omega} P(\omega)$$

- Problem?

Probabilistic Reasoning in First-Order Logic

- Set of possible worlds
  (be they represented as a full joint distribution or belief network)

- Each world ω has a probability $P(\omega)$

- Taking any sentence ϕ we can compute its probability:

$$P(\phi) = \sum_{\omega : \phi \text{ is true in } \omega} P(\omega)$$

- Problem?
  - FOL instoduces an infinite set of possible worlds!
  - Possible solution:
    unique names assumption + domain closure = database semantics

Relational Probability Models (RPM)

- Database semantics (ensures the finitness of possible worlds)
  - Except that the closed world assumption is eliminated
  - Probabilistically it does not make sense that all unknown fact are false!
- RPM have constants, functions, and predicates (considered Boolean functions)
- Each function has a type signature

$$\begin{aligned} Honest &: Customer \longrightarrow \{true, false\} \\ Kindness &: Customer \longrightarrow \{1,2,3,4,5\} \\ Quality &: Book \longrightarrow \{1,2,3,4,5\} \\ Recommendation &: Customer \times Book \longrightarrow \{1,2,3,4,5\} \end{aligned}$$

- Random variables are obtained by the instantiating each function with each possible argument
  - Each type has finitely many instances $\Rightarrow$ number of random variables is finite

RPM (cont'd)

- Dependencies between random variables are stated as one dependency statement for each function

$$\begin{aligned} Honest(c) &\sim \langle 0.99, 0.01\rangle \\ Kindness(c) &\sim \langle 0.1, 0.1, 0.2, 0.3, 0.3\rangle \\ Quality(b) &\sim \langle 0.05, 0.2, 0.4, 0.2, 0.15 \\ Recommendations(c,b) &\sim RecCPT(Honets(c), Kindness(c), Quality(b)) \end{aligned}$$

  - $RecCPT$ is a conditional distribution with $2 \times 5 \times 5$ rows
  - Conditional expressions are possible:

$$\begin{aligned} Recommendations(c,b) \sim\ & \textbf{if } Honest(c) \textbf{ then} \\ & \quad HonestRecCPT(Kindness(c), Quality(b)) \\ & \textbf{else } \langle 0.4, 0.1, 0.0, 0.1, 0.4\rangle \end{aligned}$$

    just a more compact way of representing the conditional distribution $RecCPT$

- Instantiate these dependencies $\Rightarrow$ a belief network = the semantics of the RPM

RPM (CONT'D)

- Eliminating the closed world assumption
  - Real-life problem: relational uncertainty
  - How can we ascertain that $Fan(C1, Author(B1))$ if the author of $B1$ is unknown?
    - We reason about all the possible authors!
    - Suppose there are n possible authors $A1, \ldots, An$
    - Then $Author(B1)$ is a random variable with possible values $A1, \ldots, An$
- Other conditions:
  - No dependency must be cyclic (since a belief network cannto have cycles)
  - Recursive dependencies are not supported (since this will generate infinite paths in the belief network)

Inference in RPMs

- Unrolling:
  - Collect constants and evidence
  - Construct the dependencies
  - Build the associated belief network
  - Apply inference in the belief network
- The resulting networks are very large
  - The usual solution is to construct the network on the fly, not at the beginning
  - Many of the factors constructed during variable elimination will be identical
    - Efficient caching of previous results improves the algorithm dramatically

Other Approaches to Reasoning with Uncertain Data

- Major apparent discrepancy between out mind (qualitative) and the probability theory (quantitative)
  - However, no better solution is known
- Dealing with ignorance: interval-valued degrees of belief (the Dempster-Shafer theory)
- Fuzzy logic allows vagueness (a sentence can be “sort of” true)
  - Vagueness and uncertainty are however orthogonal issues

Fuzzy Logic

- Is Jim tall? It depends. If he is around 180cm tall, then many people will hesitate
- Instead we can recognize that there are degrees of tallness:
  - the truth value of $Tall(Jim)$ is a number between 0 and 1 instead of just true or false
  - Generally to every fact A we assign a degree of truth $T(A)$ (between 0 and 1)– T is the fuzzy truth function
- Once the truth value of facts is known the truth value of complex sentences can be established inductively:

$$\begin{aligned} T(A \wedge B) &= \min(T(A), T(B)) \\ T(A \vee B) &= \max(T(A), T(B)) \\ T(\neg A) &= 1 - T(A) \end{aligned}$$

- All the inference methods work well, but there are problems with relating fuzzy truth with reality:

$$T(Tall(Jim) \wedge \neg Tall(Jim)) = 0.4 \text{ ???}$$