The weakest link

- Resolution or modus ponens are exact
  - there is no possibility of mistake if the rules are followed exactly.
- These methods of inference (also known as deductive methods) require that information be complete, precise, and consistent.
- By contrast, the real world requires common sense reasoning in the face of incomplete, inexact, and potentially inconsistent information.

Incomplete Facts

- A logic is monotonic if the truth of a sentence does not change when more facts are added. FOL is monotonic.
- A logic is non-monotonic if the truth of a proposition may change when new information (facts) is added or old information is deleted.

> "It rained last night if the grass is wet and the sprinkler was not on last evening. I am looking right now and see that the grass is wet."
> Did it rain last night?

`rained :- grass_is_wet,`
`           \+ sprinkler_was_on.`
`grass_is_wet.`
	`?- rained.`
	`Yes`
	`?- assert(sprinkler_was_on).`
	`Yes`
	`?- rained.`
	`No`
	`?- retract(sprinkler_was_on).`
	`Yes`
	`?- rained.`
	`Yes`

CIRCUMSCRIPTION

- Similar to the closed world assumption but more precise
- We specify particular predicates that are "as false as possible"
  - Meaning that they are false for all the objects except for those for which we know them to be true

    $$Bird(x) \land \neg Abnormal_1(x) \Rightarrow Flies(x)$$

    provided that $Abnormal_1$ is circumscribed
  - We draw the conclusion that $Flies(Tweety)$ out of $Bird(Tweety)$ provided that we do not know that $Abnormal_1(Tweety)$ holds
- Implemented in Prolog by the `not` predicate (more or less)

- Default logic adds a new inference rule: if α is true and β is not known to be false then γ:

$$\frac{\alpha \quad : \quad \beta}{\gamma}$$

e.g.,

$$\frac{\texttt{grass\_is\_wet} \quad : \quad \neg\texttt{sprinkler\_was\_on}}{\texttt{rained}}$$

- Nonmonotonic logic adds a new operator $\mathbb{M}$:

$$\alpha \wedge \mathbb{M}\beta \Rightarrow \gamma$$

stands for “if α is true and β is not known to be false then γ.” e.g.,

$$\texttt{grass\_is\_wet} \wedge \mathbb{M}\neg\texttt{sprinkler\_was\_on} \Rightarrow \texttt{rained}$$

$$\begin{array}{c} american(X) \wedge adult(X) \wedge \\ \mathbb{M}(\exists A \ \ (car(A) \wedge owns(X, A))) \end{array} \Rightarrow (\exists A \ \ (car(A) \wedge owns(X, A)))$$

Truth Maintenance Systems

- Problem: If we assert $\neg P$ we will have to retract P (if present)
  - Simple enough, but what if we inferred things starting from P? They will all need to be retracted.

- Efficient solution: Justification Truth Maintenance Systems (JTMS)
  - We annotate every sentence in the knowledge base with a justification = set of sentences from which it was inferred
  - If we have $P \Rightarrow Q$ and we assert P then we can add Q with the justification $\{P, P \Rightarrow Q\}$
  - A sentence can have any number of justifications
  - If we retract P the JTMS will also retract the sentences for which P is a member of every justification.

$$\begin{array}{lcl} \{P, P \Rightarrow Q\} & \longrightarrow & Q \text{ retracted} \\ \{P, P \vee R \Rightarrow Q\} & \longrightarrow & Q \text{ retracted} \\ \{R, P \vee R \Rightarrow Q\} & \longrightarrow & Q \text{ not retracted} \end{array}$$

TRUTH MAINTENANCE SYSTEMS (CONT'D)

- A JTMS will actually mark sentences as “out” instead of retracting them
  - The assumption being that a sentence that is retracted might become pertinent again in the future
  - A JTMS will thus thus retain the whole inference chain should a justification become valid again
  - JTMS also provide a mechanism for generating explanations

INEXACT FACTS

- Action A_t = leave for airport t minutes before flight.
  - Will A_t get me there on time?
  - Problems:
    1. partial observability (road state, other drivers' plans. . .)
    2. noisy sensors (traffic reports over the radio)
    3. uncertainty in action outcomes (flat tire. . .)
    4. intractable complexity of modelling and predicting traffic

- A logical approach:
  - risks falsehood: "A_{120} will get me there on time"
  - leads to conclusions that are too weak for decision making: "A_{120} will get me there on time if there's no jam on Pont Champlain and it doesn't rain and my tires remain intact. . . "
  - Note: I might reasonably expect that A_{1440} will get me there on time, but such a logical approach will make me spend a night in the airport.

- Nonmonotonic/default logic: I assume won't get a flat tire, that there is no traffic jam on Champlain...

$$drive(sherbrooke, dorval, 120) \land \mathbb{M}\neg flat\_tire \Rightarrow A_{120}$$

$$\frac{drive(sherbrooke, dorval, 120) \quad : \quad \neg jammed(champlain)}{A_{120}}$$

i.e., assume that A_{120} works unless contradicted by evidence.

- Nonmonotonic/default logic: I assume won't get a flat tire, that there is no traffic jam on Champlain...

$$drive(sherbrooke, dorval, 120) \wedge \mathbb{M}\neg flat\_tire \Rightarrow A_{120}$$

$$\frac{drive(sherbrooke, dorval, 120) \quad : \quad \neg jammed(champlain)}{A_{120}}$$

  i.e., assume that A_{120} works unless contradicted by evidence.
  But what assumptions are reasonable?

- Rules with fudge factors:
  $sprinkler \Rightarrow_{0.99} wet\_grass$
  $wet\_grass \Rightarrow_{0.7} rained$

HANDLING UNCERTAINTY

- Nonmonotonic/default logic: I assume won't get a flat tire, that there is no traffic jam on Champlain...

$$drive(sherbrooke, dorval, 120) \wedge \mathbb{M}\neg flat\_tire \Rightarrow A_{120}$$

$$\frac{drive(sherbrooke, dorval, 120) \quad : \quad \neg jammed(champlain)}{A_{120}}$$

i.e., assume that A_{120} works unless contradicted by evidence.
But what assumptions are reasonable?

- Rules with fudge factors:
$sprinkler \Rightarrow_{0.99} wet\_grass$
$wet\_grass \Rightarrow_{0.7} rained$

$sprinkler$ causes $rain$?? (problems with combinations).

- Nonmonotonic/default logic: I assume won't get a flat tire, that there is no traffic jam on Champlain...

$$drive(sherbrooke, dorval, 120) \wedge \mathbb{M}\neg flat\_tire \Rightarrow A_{120}$$

$$\frac{drive(sherbrooke, dorval, 120) \quad : \quad \neg jammed(champlain)}{A_{120}}$$

  i.e., assume that A_{120} works unless contradicted by evidence.
  But what assumptions are reasonable?

- Rules with fudge factors:
  $sprinkler \Rightarrow_{0.99} wet\_grass$
  $wet\_grass \Rightarrow_{0.7} rained$

  $sprinkler$ causes $rain$?? (problems with combinations).

- Probability: given the available evidence, A_{120} will get me to the airport in time with probability 0.03.

Probability

- Probability summarizes
  - laziness to enumerate all the exceptions, facts, ...
  - ignorance, i.e., lack of relevant facts, initial conditions, ...

- Bayesian (or subjective) probability relates probability to one's own state of knowledge.

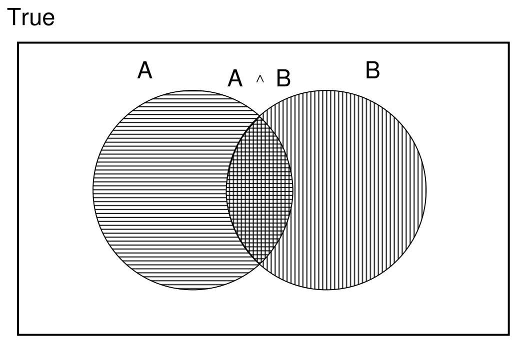
$$P(A_{120}|intact\_tires) = 0.06$$

  - Probabilities change with new evidence.

$$P(A_{120}|intact\_tires \wedge 3am) = 0.75$$

  - Analogous to logical entailment ($KB \models \alpha$), not truth.
  - Axioms of probability:

    1. $0 \leq A \leq 1$
    2. $P(True) = 1$; $P(False) = 0$
    3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$



SYNTAX

- possible worlds defined by assignment of values to random variables.
- Propositional (Boolean) random varianles: $Cavity$ (do I have a cavity?)
  - including propositional logic expressions: $\neg Burglary \vee Earthquake$
- Multivalued random variables: $Weather$ is one of $\langle sunny, rain, cloudy, snow\rangle$. Values must be exhaustive and mutually exclusive.
- Propositions constructed by assignment of a value: $Weather = sunny$.
- Unconditional (prior) probabilities of propositions: $P(Weather = sunny) = 0.72$
- Conditional (posterior) probabilities: $P(Cavity|Toothache) = 0.8$ (i.e., probability given that $Toothache$ is all I know).

Syntax (cont'd)

- Probability distribution gives values for all possible assignments: $\mathbb{P}(Weather) = \langle 0.72, 0.1, 0.08, 0.1 \rangle$ (normalized)

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

$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$				
$Cavity = false$				

SEMANTICS

- If we know more, e.g., $Cavity$ is also given, then we have $P(Cavity|Toothache, Cavity) = 1$.

- New evidence may be irrelevant, allowing simplification: $P(Cavity|Toothache, Midterm) = P(Cavity|Toothache) = 0.8$.

- Conditional probability:

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} \text{ if } P(B) \neq 0$$

  alternatively

$$P(A \wedge B) = P(A|B)P(B) = P(B|A)P(A)$$

- Bayes' rule:

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

  Why is this useful?

Bayes' rule

$$P(Meningitis|StiffNeck) = \frac{P(StiffNeck|Meningitis)P(Meningitis)}{P(StiffNeck)}$$
$$= \frac{0.8 \times 0.0001}{0.1} = 0.0008$$

BAYES' RULE

$$P(Meningitis|StiffNeck) = \frac{P(StiffNeck|Meningitis)P(Meningitis)}{P(StiffNeck)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$$

- i.e., Bayes' rule is useful for assessing diagnostic probability from causal probability:

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

- Chain rule: successive application of the product rule (on joint probability distributions)

$$\begin{aligned} \mathbb{P}(X_1, \ldots, X_n) &= \mathbb{P}(X_1, \ldots, X_{n-1})\mathbb{P}(X_n|X_1, \ldots, X_{n-1}) \\ &= \mathbb{P}(X_1, \ldots, X_{n-2})\mathbb{P}(X_{n_1}|X_1, \ldots, X_{n-2})\mathbb{P}(X_n|X_1, \ldots, X_{n-1}) \\ &= \ldots \\ &= \prod_{i=1}^{n} \mathbb{P}(X_i|X_1, \ldots, X_{i-1}) \end{aligned}$$

NORMALIZATION

- We want to compute a posterior distribution over A given $B = b$, and suppose A has possible values $\langle a_1, \ldots, a_m \rangle$.

$$\begin{aligned} P(A = a_1 | B = b) &= P(B = b | A = a_1) P(A = a_1) / P(B = b) \\ &\ldots \\ P(A = a_m | B = b) &= P(B = b | A = a_m) P(A = a_m) / P(B = b) \end{aligned}$$

$$\begin{aligned} \sum_i P(A = a_i | B = b) &= \left(\sum_i P(B = b | A = a_i) P(A = a_i) \right) / P(B = b) \\ 1 &= \left(\sum_i P(B = b | A = a_i) P(A = a_i) \right) / P(B = b) \\ 1 / P(B = b) &= 1 / \sum_i P(B = b | A = a_i) P(A = a_i) \end{aligned}$$

$\rightarrow$ normalization factor α

- $\mathbb{P}(A | B = b) = \alpha \mathbb{P}(B = b | A) \mathbb{P}(A)$
e.g., let $\mathbb{P}(B = b | A) \mathbb{P}(A) = \langle 0.4, 0.2, 0.2 \rangle$;
then $\mathbb{P}(A | B = b) = \alpha \langle 0.4, 0.2, 0.2 \rangle = \frac{\langle 0.4, 0.2, 0.2 \rangle}{0.4 + 0.2 + 0.2} = \langle 0.5, 0.25, 0.25 \rangle$

COMBINING EVIDENCE

- Often easier to analyze each specific circumstance instead of the whole situation:

$$
\begin{aligned}
&P(RunOver|Cross)\\
&\quad = P(RunOver|Cross, Light = green)P(Light = green|Cross)\\
&\quad + P(RunOver|Cross, Light = yellow)P(Light = yellow|Cross)\\
&\quad + P(RunOver|Cross, Light = red)P(Light = red|Cross)
\end{aligned}
$$

- I.e., we can introduce a variable as an extra condition:

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

- When Y is absent, we have summing out or marginalization:

$$P(X) = \sum_z P(X|Z = z)P(Z = z) = \sum_z P(X, Z = z)$$

- Given a joint distribution over a set of variables, the distribution over any subset can be calculated by summing out the other variables.

FULL JOINT DISTRIBUTION

- 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 w_i.
  - Possible worlds are exclusive and exhaustive, hence the sum of the probabilities in the matrix is always 1: $\sum_i P(w_i) = 1$.

	$Toothache = true$	$Toothache = false$
$Cavity = true$	0.04	0.06
$Cavity = false$	0.01	0.89

- For any proposition ϕ defined on the random variables: $\phi(w_i)$ is true or false ϕ is equivalent to the disjunction of w_is 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.

- We are interested in the posterior joint distribution of the query variables $\mathbf{Y}$ given specific values $\mathbf{e}$ for the evidence variables $\mathbf{E}$.

- We may have *hidden variables* $\mathbf{H} = \mathbf{X} \setminus \mathbf{Y} \setminus \mathbf{E}$.

- Then the required summation of joint entries is done by summing out the hidden variables:

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

- The terms in the summation are joint entries because $\mathbf{Y}$, $\mathbf{E}$, and $\mathbf{H}$ together exhaust the set of random variables.

- Problem: Huge time and space complexity.