REASONING WITH UNCERTAINTIES

Sources of uncertainties

Dealing with vagueness

Dealing with probabilities

KINDS OF UNCERTAINTIES

Knowledge

- **Modeling of agents' limited/uncertain knowledge**
- **Separate contexts for each agent**
- **Reasoning about rational and cooperative behavior**

Vagueness

- **Degree of uncertainty about continuous properties**
- **Modeled by Fuzzy logic**
- **Reasoning with aspects of lexical semantics**

Probabilities

- **Several possible outcomes, unknown at present**
- **Modeled by statistic methods**
- **Reasoning about combinations and actions**

Underspecification

- **Uncertainty due to partial, local knowledge**
- **Modeled by interpretation-neutral representations**
- **Reasoning in staged process**

PROBABILITIES

Positions towards probabilities

- ***Frequencies*** **- numbers based on experiments**
- ***Objectivist*** **- real aspects of the universe**
- ***Subjectivist*** **- according to beliefs of agents**

Example **- "The sun will still exist tomorrow"**

- ***Undefined*****, due to lack of experiments**
- ***1*****, all experiments in the past succeeded**
- ***1-ε*****, where *ε* is the proportion of stars going supernova per day**
- ***(d+1)/(d+2)*****, where *d* is the number of days the sun has existed so far (Laplace)**
- ***Probability*** **can be derived from the type, age, size and temperature of the sun (similar to other stars)**

The first three methods are frequentist

The last two subjectivist

Choosing among the reference class for frequentist views is subjective

FUZZY LOGIC

Applicability

- **How well an object satisfies a vague description (e.g, being "tall", and "smart")**
- **Similarity to a prototype - "sort of", …**

Idea

- ***TallPerson* is a *fuzzy* predicate**
- **Truth of value of *TallPerson*(Nate) is p, $0 \leq p \leq 1$**
- ***Fuzzy set* interprets a predicate as a set of its members without sharp boundaries**

Combining uncertainties

- $T(A \wedge B) = \min(T(A),T(B))$
- $T(A \vee B) = \max(T(A),T(B))$
- $T(\neg A) = 1 - T(A)$

but $T(A \vee \neg A) \neq T(True)$

***Commercial applications* control systems (e.g., shavers)**

- ***Small* rule bases with no (little) chaining**
- ***Tunable* and *adjustable* parameters (learning)**

BAYES THEOREM

Conditional probability

$P(A|B)$ **Probability of A given B**

more practical than *joint probability* (complete assignment of values to random variables)

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} \qquad P(B) > 0$$

$$P(A \wedge B) = P(A|B)P(B),\ P(A \wedge B) = P(B|A)P(A)$$ **Product rule**

Derivation of the theorem (law of Bayes)

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \qquad P(A) > 0$$

***Example* – How many patients with with stiff neck have meningitis?**

M, S **Patient has meningitis (M), stiff neck (S)**

$P(S|M) = 0{,}5$ **Meningitis causes a stiff neck in 50%**

$P(M) = 1/50000$ **Probability a patient has meningitis**

$P(S) = 1/20$ **Probability a patient has a stiff neck**

$$P(S|M) = \frac{P(S|M)P(M)}{P(S)} = \frac{0{,}5x1/50000}{1/20} = 0{,}0002$$

BAYES THEOREM APPLICATION

A generalization

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\Sigma^n_{k=1} P(E|H_k)P(H_k)}$$

- **Probability of an evidence depends on all possible hypotheses**
- **The set of all hypotheses must be mutually exclusive and exhaustive**

Problems

- **Knowledge acquisition is hard**
- **Too many probabilities needed**
- **Computation time is too large**
- **Updating new information is difficult and time consuming**
- **Exceptions like “None of the above” cannot be represented**
- **Humans are not very good probability estimators**

Simple Bayes rule-based systems are impractical

CERTAINTY FACTOR

Occurrence

- **Associated with rules in MYCIN**
- **Measures of belief and disbelief of an hypothesis**

Computation

$$B(H_i|E) = \frac{\max[P(H_i|E)P(H_i)]-P(H_i)}{(1-P(H_i))P(H_i|E)} \quad \text{unless } P(H_i) = 1$$

$$D(H_i|E) = \frac{P(H_i)-\min[P(H_i|E)P(H_i)]}{P(H_i)P(H_i|E)} \quad \text{unless } P(H_i) = 1$$

$$C(H_i|E) = B(H_i|E) - D(H_i|E)$$

Combination

$$B(H_i|E_1,E_2) = B(H_i|E_1) + B(H_i|E_2)\,(1-B(H_i|E_1))$$

Disbelief calculated analoguously

Assessment

- **Much simpler than Bayes theorem**
- **Semantics and combination increasingly unclear**

DEMPSTER SHAFER MODELS

Motivation

- **Addresses distinction between**
  Uncertainty **and** ***ignorance***
  Probability axioms insist: $P(A) + P(\neg A) = 1$
  This may not meaningfully be appplicable under conditions of incomplete knowledge (i.e., presuppositions for knowing about P **and** $\neg P$**)**

Basic Idea

- **Probability that evidence supports a proposition Belief function** $Bel(X)$
- **No knowledge –** $Bel(X) = 0$, $Bel(\neg X) = 0$ ***sceptical*** **position**
  Good knowledge – "competence" 0.9
  $Bel(X) = 0.9 \times 0.5 = 0.45$, $Bel(\neg X) = 0.9 \times 0.5 = 0.45$

Interpretation

- **Utility for actions unclear – semantics for it unclear**
- **Defines probability interval**

BELIEF NETWORKS

Purpose

- **Expresses dependence between variables**
- **Specifications of joint probability distributions**

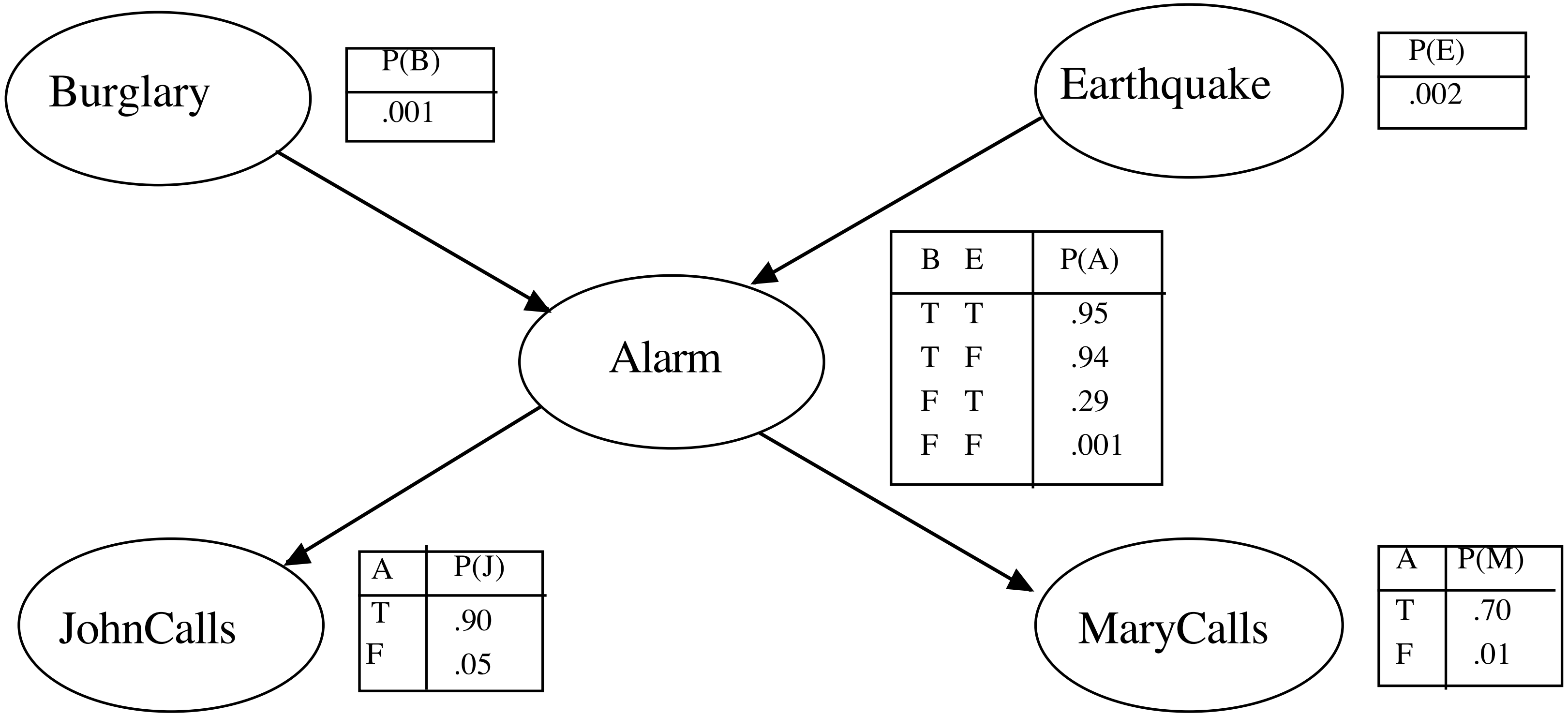
Components

- **Set of random variables are nodes of the network**
- **Directed links between nodes –** ***direct*** **influence**
- **Conditional probability table for each nodes:**
  quantifies effects that parents have on nodes
- **No directed cycles in the graph –** ***direct*** **influence**

Usage

- **Decide about** ***direct*** **influence to determine topology**
- **Defines** ***conditional*** **probabilities for variables in direct influence**

BELIEF NETWORKS – EXAMPLE



P(B)
.001

P(E)
.002

B	E	P(A)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J)
T	.90
F	.05

A	P(M)
T	.70
F	.01

BELIEF NETWORKS – SEMANTICS

Joint probability distribution

Conjunction of a particular assignment to variables

$$P(X_1=x_1 \wedge \ldots \wedge X_n=x_n)=P(x_1,\ldots,x_n)=\Pi^n_{i=1} P(x_i|\mathrm{Parents}(X_i))$$

Example

- **Alarm has sounded, neither a burglary nor earthquake has occurred, both John and Mary call**

  $P(J \wedge M \wedge A \wedge \neg B \wedge \neg E)=$

  $P(J|A)P(M|A)P(A|\neg B \wedge \neg E)P(\neg B)P(\neg E)=$

  0.9 x 0.7 x 0.001 x 0.999 x 0.998 = 0.00062

CONSTRUCTING BELIEF NETWORKS

Sketch of a procedure

1. **Choose the set of relevant variables X_i**
2. **Choose an ordering on the variables**
3. **While there are still variables left**

   a) Pick a variable X_i and add a node to the network

   b) Set *parents(X_i)* to some minimal set of nodes in the net such that conditional independence is satisfied (independent of other nodes)

   a) Define the conditional probability table for X_i

Properties

- **Network is acyclic**
- **No redundant probability values**
- **"Impossible" to violate probability axioms**

Techniques

- **"Direct influencers" first – start with "root causes"**
- **Conditional probabilities for – deterministic (logical OR)**
  – noisy-OR (generalization of logical OR)

BELIEF NETWORK INFERENCES

Categories

- ***Diagnostic* inferences (from effects to causes)**
  Given JohnCalls, *P(Burglary|JohnCalls)* = 0.016
- ***Causal* inferences (from causes to effects)**
  Given Burglary, *P(JohnCalls|Burglary)* = 0.86 and *P(MaryCalls|Burglary)* = 0.67
- ***Intercausal* inferences (between causes of an effect)**
  Given Alarm, *P(Burglary|Alarm)* = 0.376
  If also Earthquake, *P(Burglary|Alarm ^ Earthquake)* = 0.003
- ***Mixed* inferences (between two or more of these)**
  Given JohnCalls and ¬Earthquake, *P(Alarm|JohnCalls ^ ¬Earthquake)* = 0.03
  Given JohnCalls and ¬Earthquake, *P(Burglary|JohnCalls ^ ¬Earthquake)* = 0.017

Uses

- ***Decisions* about actions**
- ***Decisions* about observations for gaining evidence**
- ***Sensitivy* analysis – degree of impact on result**
- ***Explaining* results of probabilistic inference**

CASE STUDY: THE PATHFINDER SYSTEM

System scope

- **Diagnostic expert system for lymph-node diseases**
- **Over 60 diseases and over 100 disease findings**

History

- **PATHFINDER I: rule-based system, no uncertainty**
- **PATHFINDER II: experimental, including certainty factors and Dempster-Shafer. Simplified Bayesian model outperformed other methods**
- **PATHFINDER III: with simplified Bayesian model, paying attention to low probability events**
- **PATHFINDER IV: Belief network for handling dependencies (simplified Bayesian model does not)**

Evaluation of correctness in diagnoses

- **PATHFINDER III: (7,9/10)**
- **PATHFINDER IV: (8,9/10)**

Amounts to saving one more life every 1000 cases

Most recent results: system outperform experts creators

REPRESENTING ARGUMENTS IN BELIEF NETWORKS

- **Represent discrete variables and dependencies in terms of conditional probabilities**
- **Enriching the semantics to represent all elements of Toulmin's and Walton's models**

Extended node types

Evidence nodes

Domain facts, prior probability may be assigned to them

Roots of the network, cannot be justified

Truth nodes

Domain facts, may be assigned by some argumentation step, parents may be one or more warrant nodes and premises according to the warrants' structure

Warrant nodes

Relationship between premises and conclusions, according to argumentation scheme

Degrees of belief associated with warrants

See later for *rebuttal nodes* and *proof nodes*

EXAMPLE – EXPERT OPINION

Warrant

(C) Statements of FDA dealing with healthful living, in which they are expert, are true

Premises

(D) FDA says that eating vegetables is a form of healthy eating

(A) FDA is and expert in healthful eating

(B) The statement "eating vegetables is a form of healty eating" deals with healthful living

Conclusion

Eating vegetables is a form of healthful living

Probabilities associated with each component may change the degree of belief in the conclusion (see the conditional probability table following)

EXAMPLE – CONDITIONAL PROBABILITY TABLE

Conclusion: **"Eating vegetables is a form of healthful living" is true/false,**

depending on truth/falsity of the premises (A, B, D) and

degree of likelihood of the warrant (C)

A	**false**															
B	**false**								**true**							
C	**alm. certain**		**very likely**		**likely**		**not applic.**		**alm. certain**		**very likely**		**likely**		**not applic.**	
D	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**
False	**.8**	**.75**	**.82**	**.8**	**.85**	**.8**	**.01**	**.01**	**.75**	**.7**	**.8**	**.75**	**.8**	**.78**	**.99**	**.99**
True	**.2**	**.25**	**.18**	**.2**	**.15**	**.2**	**.99**	**.99**	**.25**	**.3**	**.2**	**.25**	**.2**	**.22**	**.01**	**.01**

A	**true**															
B	**false**								**true**							
C	**alm. certain**		**very likely**		**likely**		**not applic.**		**alm. certain**		**very likely**		**likely**		**not applic.**	
D	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**	**false**	**true**
False	**.75**	**.7**	**.8**	**.75**	**.8**	**.75**	**.99**	**.99**	**.65**	**.01**	**.7**	**.2**	**.75**	**.3**	**.99**	**.99**
True	**.25**	**.3**	**.2**	**.25**	**.2**	**.22**	**.01**	**.01**	**.35**	**.99**	**.3**	**.8**	**.25**	**.7**	**.01**	**.01**

CHAINING OF ARGUMENTS

Implementation in belief networks

Attaching subnetworks corresponding to argumentation schemata together

Addressing the premise attacked by a critical question

Replacing that premise by the conclusion of the related network

Example

Concluding expertise of FDA trustfulness and credibility

Extending argumentation scheme *From Expert Opinion* by
the argumentation scheme *From Verbal Classification*

From answering "Is the FDA a credible information source in the domain?" and
"Is the FDA a trustful source in the domain?"

Concluding that FDA is an expert in the domain

REBUTTAL NODES – HANDLES EXCEPTIONS IN RULES

Example

Premises

"If someone desires to be in good health and an action contributes to maintaining good health, the he or she should perform that action"

"Eating vegetables contributes to maintaining good health"

"Person U desires to be in good health"

Conclusion

"Person U should eat vegetables"

Exception

"Unless person U suffers from colitis"

Belief in the conclusion is *high* if premises are highly believed and the exception not

Belief in the conclusion is *low* if warrant not applicable or belief in the exception is high

PROOF NODES – CONVERGENT/LINKED ARGUMENTS

Example

Argumemts

1. **"You should eat more vegatables because eating vegetables contributes to maintaining good health" (degree of belief .7)**

2. **"You should eat more vegatables because eating vegetables contributes to maintaining good health and also to having a good appearance" (degree of belief .8)**

Conditional probability tables

"Person U should eat vegetables"

Exception

"Unless person U suffers from colitis"

Belief in the conclusion is *high* if premises are highly believed and the exception not

Belief in the conclusion is *low* if warrant not applicable or belief in the exception is high