Quantitative Measures of Uncertainty

Jonathan Lawry

Department of Engineering Mathematics
University of Bristol

Reasoning under Certainty

- Classical logic provides a formalism for reasoning given a set of propostions which are assumed to be (certainly) true.
- Given a number of inference rules we are then able to deduce other propositions as being (certainly) true.
- **Example:** Suppose it is known that the following hold:
- Patient X *has* symptom S.
- *All* patients with symptom S have disease D.
- Therefore, patient X *has* disease D.
- This an example of a syllogism: X is a Y, All Y are Z, Therefore, X is a Z.

Reasoning under Uncertainty

- Almost all practical reasoning involves uncertainty.
- In otherwords, our knowledge consists of propositions in which we have some level of belief but about which we are not certain.
- This type of reasoning is beyond the scope of classical logic.
- **Example:** Suppose it is known that the following hold:
- It is *highly likely* that patient X has symptom S.
- Disease D is *very common* in patients with symptom S.
- Does patient X have disease D?
- Classical logic cannot answer this question.

Intelligent Agents

- We aim to investigate frameworks according to which a rational intelligent agent could reason under uncertainty.
- Such an agent could be an artificial intelligence program or a person.
- The underlying assumption is that their reasoning is governed by some internally consistent strategy or algorithm.
- In order to develop reasoning frameworks for intelligent systems we must consider computational aspects of reasoning under uncertainty as well as just rationality requirements.

Quantifying Belief

- Most approaches to reasoning under uncertainty assume that an agent's level of belief can be quantified on a scale between 0 and 1.
- So for a proposition A a belief level of 0 means that the agent is certain that A is false, and a belief value of 1 means that they are certain that A is true.
- A belief value of 0.5 means that the agent has an intermediate level of belief in A.
- In this course we will consider a number of different ways of quantifying belief on a 0 to 1 scale.
- But what are the basic elements to which we should allocate belief values?

Possible Worlds

- Possible worlds corresponds to complete descriptions of a systems, of the state of knowledge, or to the outcome of an experiment.
- Possible worlds are exclusive in that only one world can correspond to the true world. i.e. there is exactly one possible world, denote w^*, which is the true world.
- **Example:** Suppose that only 4 people have registered to attend this unit; Bill, Mary, Ethel and Fred.
- In this case, possible worlds corresponding to the possible attendance at any given lecture are tuples

$$\langle BILL, MARY, ETHEL, FRED \rangle$$

  where $BILL, \ldots, FRED$ are binary valued variables such that, for example, $BILL = 1$ means that Bill is present and $BILL = 0$ means that Bill is absent.

Possible Worlds: More Examples

- Consider an experiment in which a dice is thrown twice and the score on each dice recorded.
- In this case, a possible world is simply a pair of numbers identifying the score on each dice e.g $\langle 5, 3 \rangle$ records a 5 on throw 1 and a 3 on throw 2.
- In an experiment in which a coin is repeatedly tossed 3 times the possible worlds correspond to sequences of 3 Hs or Ts.
- In an experiment where a single card is drawn from a pack of cards, the possible worlds are all the different cards in the pack.

Possible Worlds as State Descriptions

- In his work on Inductive Logic Rudolf Carnap proposed the following formulation:
- Suppose there are n unary predicates $Q_1, \ldots, Q_n$ which can be applied to m objects $a_1, \ldots, a_m$.
- Any given object must satisfy exactly one of the formula of the following form:

$$\alpha_i(x) = \pm Q_1(x) \land \pm Q_2(x) \land \ldots \land \pm Q_n(x)$$

  where $\pm Q_i(x)$ denotes either $Q_i(x)$ or $\neg Q_i(x)$.
- A state description then indicates which of the α_i formula applies to each of the objects $a_1, \ldots, a_m$, and have the form:

$$\alpha_{i_1}(a_1) \land \ldots \land \alpha_{i_m}(a_m)$$

Possible Worlds as State Descriptions:2

- Suppose that we have the predicated *Happy* and *Rich* to describe the students registered for this unit i.e. Bill, Mary, Ethel and Fred
- We have $\alpha_1(x) = Happy(x) \wedge Rich(x)$, $\alpha_2(x) = Happy(x) \wedge \neg Rich(x)$, $\alpha_3 = \neg Happy(x) \wedge Rich(x)$, and $\alpha_4 = \neg Happy(x) \wedge \neg Rich(x)$
- Then the following is an example of a state description: $\alpha_1(Bill) \wedge \alpha_1(Mary) \wedge \alpha_4(Ethel) \wedge \alpha_2(Fred)$.
- As such it is a complete description of the state of knowledge concerning the happiness and wealth of those students enroled on this unit.
- i.e. Bill and Mary are both, both happy and rich, Ethel is neither happy nor rich and Fred is happy but not rich.

Propositions as Sets of Possible Worlds

- The propositions with which an intelligent agent would wish to associate belief values can be represented by subsets of possible worlds.
- For example, the proposition 'All the female students registered on the unit are present' is represented by the set of possible worlds $\{\langle 0,1,1,0\rangle, \langle 1,1,1,0\rangle, \langle 0,1,1,1\rangle, \langle 1,1,1,1\rangle\}$ corresponding to all those worlds in which both Mary and Ethel are present.
- Similarly, the proposition 'the sum of the scores from both throws of the dice is 4' is represented by the set $\{\langle 1,3\rangle, \langle 3,1\rangle, \langle 2,2\rangle\}$.
- A proposition represented by set A if true, if $w^* \in A$

Uncertainty Measures

- We now define a generic quantitative measure of belief as follows:
- Let W denote the set of all possible worlds (which is assumed to be finite)
- An uncertainty measure $\mu : 2^W \to [0, 1]$ is a function mapping subsets of W into the interval $[0, 1]$, such that:
- $\mu(W) = 1,\ \mu(\emptyset) = 0.$
- For $S \subseteq W$ $\mu(S) = 1$ means that the proposition represented by S is certainly true.
- For $S \subseteq W$ $\mu(S) = 0$ means that the proposition represented by S is certainly false.

Properties of Uncertainty Measures

- There are many types of uncertainty measures each satisfying different additional *properties* beyond those on the previous slide.
- Such *properties* characterise different assumption about how a *rational* agent should quantify their beliefs.
- Some *properties* are more general and some more specific resulting in more or less *expressive* measures.
- We shall now introduce a number of these *properties* and throughout this unit consider different uncertainty measures which satisfy them.
- The most fundamental property is *monotonicity* as follows:

Monotonic Measures

- Consider the proposition 'At least one of the female students is present'. It would seem natural to assume that an agent's belief in this proposition should be greater than or equal to their belief in the proposition 'Both female students are present'.
- Since the former is a more general statement than the latter.
- For uncertainty measures this requires that: If $A \subseteq B$ then $\mu(A) \leq \mu(B)$
- e.g. $\{\langle 0,1,1,0\rangle, \langle 1,1,1,0\rangle, \langle 0,1,1,1\rangle, \langle 1,1,1,1\rangle\} \subseteq \{\langle 0,1,1,0\rangle, \langle 0,0,1,0\rangle, \langle 0,1,0,0\rangle, \langle 1,1,1,0\rangle, \langle 1,0,1,0\rangle, \langle 1,1,0,0\rangle, \langle 0,1,1,1\rangle, \langle 0,0,1,1\rangle, \langle 0,1,0,1\rangle, \langle 1,1,1,1\rangle, \langle 1,0,1,1\rangle, \langle 1,1,0,1\rangle\}$

Additive Measures

- Consider the disjunction 'Either all the male students were present and none of the female OR all the female students were present and none of the male'
- Since the disjunctions are exclusive, it might seem reasonable that an agent's belief in the disjunction should be the sum of their beliefs in each of the disjuncts.
- For uncertainty measures: If $A \cap B = \emptyset$ then $\mu(A \cup B) = \mu(A) + \mu(B)$.
- e.g. $\mu(\{\langle 1, 0, 0, 1 \rangle, \langle 0, 1, 1, 0 \rangle\}) = \mu(\{\langle 1, 0, 0, 1 \rangle\}) + \mu(\{\langle 0, 1, 1, 0 \rangle\})$
- General form: For any A and B, $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

Sub and Super Additive Measures

- Sub-Additive Measures: If $A \cap B = \emptyset$ then $\mu(A \cup B) \leq \mu(A) + \mu(B)$
- General form: For any A and B $\mu(A \cup B) \leq \mu(A) + \mu(B) - \mu(A \cap B)$
- Super-Additive Measures: If $A \cap B = \emptyset$ then $\mu(A \cup B) \geq \mu(A) + \mu(B)$
- General form: For any A and B $\mu(A \cup B) \geq \mu(A) + \mu(B) - \mu(A \cap B)$
- Maxitive Measures: For any A and B $\mu(A \cup B) = \max(\mu(A), \mu(B))$: Special case of sub-additive measure.
- Minimal Measures: For any A and B $\mu(A \cap B) = \min(\mu(A), \mu(B))$: Special case of super-additive measure.

Random Variables

- *Random Variables are like the Holy Roman Empire...*
- It is often the case that states of the world W have different *measurements* associated with them.
- e.g. the magnitude of the force acting in a certain direction, temperature of a component, colour of a playing card, the sum of the scores on two throws of a dice, the number of heads in a sequence of 100 coin tosses, the number of male students present.
- A belief measure on W then gives us information about the uncertainty regarding such a measurement.

Random Variables: 2

- A random variable $X : W \to \Omega$ is a function from W into some measurement domain Ω (often $\Omega = \mathbb{R}$ the real numbers).
- An uncertainty measure on W induces an uncertainty measure on X as follows:
- $\mu(X = x) = \mu(\{w : X(w) = x\})$
- For $S \subseteq \Omega$ $\mu(X \in S) = \mu(\{w : X(w) \in S\})$
- e.g. $X =$ number of male students present, then $\mu(X = 2) = \mu(\{\langle 1, 0, 0, 1\rangle, \langle 1, 0, 1, 1\rangle, \langle 1, 1, 0, 1\rangle, \langle 1, 1, 1, 1\rangle\})$

Joint Uncertainty Measures

- Often there are a number of different measurements associated with a state of the world, represented by different random variables $X_1, \ldots, X_n$.
- A joint uncertainty measure on these variables is then given by: $\mu(X_1 = x_1, \ldots, X_n = x_n) = \mu(\{w : X_1(w) = x_1, \ldots, X_n(w) = x_n\})$
- $\mu(X_1 \in S_1, \ldots, X_n \in S_n) = \mu(\{w : X_1(w) \in S_1, \ldots, X_n(w) \in S_n\})$
- e.g. $X_1 =$ number of male students present, $X_2 =$ number of female students present. Then $\mu(X_1 = 1, X_2 = 1) = \mu(\{\langle 1,0,1,0\rangle, \langle 1,1,0,0\rangle, \langle 0,0,1,1\rangle, \langle 0,1,0,1\rangle\})$

Random Sets

- We now consider measurements which are imprecise so that $R(w) = S$ where $S \subseteq \Omega$ (the measurement domain).
- In other words $R : W \to 2^{\Omega}$.
- As with random variables an uncertainty measure on W defines an uncertainty measure on R as follows: $\mu(R = S) = \mu(\{w : R(w) = S\})$
- Let R denote the lecturer's estimate of the number of people present. He or she is unable to count accurately so for example $R(\langle 1, 0, 1, 1 \rangle) = \{3, 4\}$