Uncertainty In Artificial Intelligence
Theory

Christian Rossi

Academic Year 2023-2024

Abstract

The topics of the course are:

- Uncertainty sources that affect models: typology, issues, and modeling approaches.
- Measure-based uncertainty modeling.
- Logic-based uncertainty modeling.
- Fuzzy models: fuzzy sets, fuzzy logic, fuzzy rules, motivations for fuzzy modeling, tools for fuzzy systems, design of fuzzy systems, applications.
- Bayesian networks: basics, design, learning, evaluation, applications.
- Hidden Markov Models: basics, design, learning, evaluation, applications.
- Applications: motivations, choices, models, case studies.

Contents

Chapter 1

Introduction

1.1 Definition

Definition (*Uncertainty*)

> *Uncertainty* refers to epistemic situations involving imperfect or unknown information. It applies to predictions of future events, to physical measurements that are already made, or to the unknown.

Uncertainty arises in partially observable or stochastic environments, as well as due to ignorance, indolence, or both. It arises in any number of fields, including insurance, philosophy, physics, statistics, economics, finance, medicine, psychology, sociology, engineering, metrology, meteorology, ecology and information science.

The lack of certainty, a state of limited knowledge where it is impossible to exactly describe the existing state, a future outcome, or more than one possible outcome. This puts in evidence that uncertainty is related to the need of describing a piece of reality.

1.2 Modelling

Modelling is at the base of our life: the way we interact with the world is through models that interpret data coming from sensors and generate knowledge and actions. Modelling is also the way we may represent entities in a computer and possibly making it reasoning on them.

Definition (*Model*)

> A *model* is a representation of some entity, defined for a specific purpose. A model captures only the aspects of the entity modelled that are relevant for the purpose. A model is necessarily different from the

modelled entity. So, intrinsic to modelling are all sort of uncertainties.

All Artificial Intelligence applications are based on models, either defined by somebody or learned. These models are represented In different ways, but share uncertainty issue mainly on inputs.

1.3 Uncertainty classification

The uncertainty can be of two main types:

- Epistemic uncertainty: it is due to things one could in principle know but does not in practice. This may because a measurement is not accurate, because the model neglects certain effects, or because particular data have been deliberately hidden. It is also known as systematic uncertainty and can in principle be reduced by enriching the model.
- Aleatoric uncertainty: it is representative of unknown unknowns that differ each time we run the same experiment. Aleatoric uncertainty is also known as statistical uncertainty, since only statistical information can describe it. This may also depend on the way we get and elaborate data. In general, it is present when the model is missing some aspects.

The sources of uncertainty can be:

- Parameter: it comes from the model parameters, whose exact values are unknown to experimentalists and cannot be controlled in experiments, or whose values cannot be inferred by statistical methods.
- Parametric variability: it comes from the variability of input variables of the model.
- Structural: also known as model inadequacy, model bias, or model discrepancy, this comes from the lack of knowledge of the problem.
- Algorithmic: also known as numerical uncertainty, or discrete uncertainty. This type comes from numerical errors and numerical approximations in the implementation of the computer model.
- Experimental: also known as observation error, this comes from the variability of experimental measurements.
- Interpolation: this comes from a lack of variable data collected from computer model simulations and/or experimental measurements.

1.4 Uncertainty modelling

The type of uncertainty model depends on the type of uncertainty, its sources and the information we have in uncertainty and mostly has to do with qualification and quantification of uncertainty. The possible models for uncertainty are: statistical, logical and cognitive.

Artificial Intelligence and Machine Learning technologies are based on models that include uncertainty models of these sorts, essential not only for the implementation of effective models, but also to define learning models able to cope with complex situations, and to evaluate the quality of learned/developed models. There are two major types of problems in uncertainty quantification:

- Forward propagation of uncertainty: the various sources of uncertainty are propagated through the model to predict the overall uncertainty in the system response:
  - To evaluate low-order moments of the outputs (mean and variance).
  - To evaluate the reliability of the outputs.
  - To assess the complete probability distribution of the outputs.

  This is what is done also in Bayesian networks and graphical models.

- Inverse assessment of model uncertainty and parameter uncertainty, where the model parameters are calibrated simultaneously using test data: given some experimental measurements of a system and some results from its mathematical model, inverse uncertainty quantification estimates the discrepancy between the experiment and the mathematical model (bias correction) and estimates the values of unknown parameters in the model if there are any (parameter calibration).

The models used in Artificial Intelligence can be classified in three main types:

- Symbolic models: elements of the models are expressed as terms related to entities to be modelled. The state of the world is represented by facts expressed in formal languages close to natural languages.
- Sub-symbolic models: elements of the models are expressed by code.
- Black-box models: the model can be computed and possibly investigated, but it is only regarded as a computational way to map inputs to outputs.

For symbolic models a fact is true in a model if it is possible to collect enough evidence to support it. The only really true facts are the ones true by definition. All the others may be supported by evidence.

1.5 Ignorance management

There are many potential sources of ignorance when reasoning in the real world:

- Insufficient data.
- Biased data: data are collected by sensors affected by errors.
- Variable data: data are collected by imprecise sensors.
- Reliability of data.
- Fuzzyness.
- Reliability of the model: depends on the model design, implementation and parametrization.
- Incompleteness of the model.

Example : Let's consider the sentence "The elephant weighs 2 tons". This can be interpreted in various ways, each slightly different:

- The elephant weighs exactly 2 tons.
- The elephant weighs 2 tons $\pm$ 10 kg, given the resolution of the weight scales of the instrument.
- The elephant weighs approximately 2 tons, but we cannot say anything more precise.
- We are not sure about any previous sentence because we do not have enough evidence.

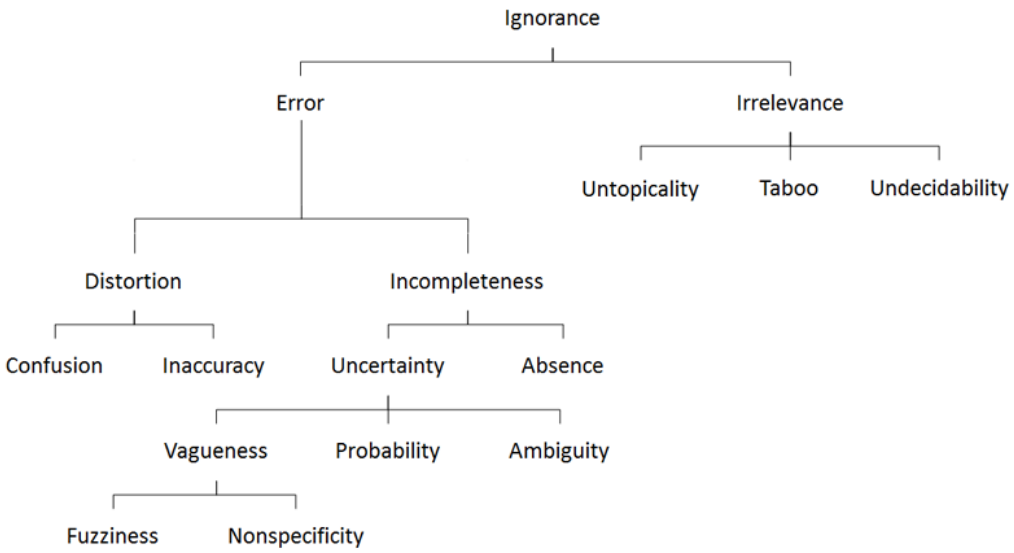

Figure 1.1: Smithson's taxonomy of ignorance and uncertainty

To model ignorance most often it is decided to associate measures of some aspects. Let's distinguish between two aspects:

- The type of representation: numbers, labels, intervals, ...
- The represented ignorance that we would like to model: i.e, probability, reliability, subjective evaluation, ...

The probability is represented with numbers between zero and one, and a well-established set of rules and properties are associated to its management, among which, given a set of alternative hypothesis:

- The sum of their probabilities should be one.
- The probability a posteriori of a hypothesis h_i given some evidence e is given by the Bayes theorem:

$$P(h_i \mid e) = \frac{P(e \mid h_i)P(h_i)}{P(e)}$$

Probability was used, for example, in the MYCIN that was one of the first expert systems, aimed at diagnosing blood illness. They modeled certainty by considering two numerical factors:

- Measure of increased Belief: $MB = \frac{P(\frac{h}{e})-P(h)}{1-P(h)}$.

- Measure of decreased Disbelief: $MD = \frac{P(h)-P(\frac{h}{e})}{P(h)}$.

The measure of a statement is given by the certain factor:

$$CF = MB - MD \in [-1; 1]$$

The main hypothesis for this solution is that the number given as MB and MD are not statistical probabilities, but subjective probabilities, provided by different experts and combined by rules (this may be ambiguous).

Compared to probabilities, linguistic terms are less ambiguous than numbers. Using a limited set of labels it is possible to associate to statements subjective evaluation, on which it is relatively easy to make subjective judgements converge. Then, a computational mechanism is needed to define how to combine labels. This is done by using fuzzy systems, that are a representation of truth of a statement in linguistic terms, as evaluation of its fuzzyness.