Representing uncertainty: randomness vs. incomplete information

Didier Dubois

IRIT-CNRS, University of Toulouse, France.

Dedicated to the memory of Janos Fodor

Introduction

Science aims at being precise and objective but...

- We seldom access reality as such, due to limited perception capabilities.
- To-date, we are able to collect and to receive more and more information, but this information
  - can be of poor quality,
  - can be conflicting
- In many areas, what we take for knowledge is but reasonable belief.
- The increasing role of computers in daily lives seems to have pushed reality away, while claiming to make it closer.

Computers have modifyed paradigms of scientific investigation

- Computations that were impossible to run some time ago become feasible
- More and more data, including from human origin (like testimonies)
- A variety of data types, including images, natural language, etc.
- More and more sources of various origins : data bases, sensors, humans....

Importance of modeling the imperfection of information

The fantastic computation power available is counterbalanced by the possible lack of good quality of the data to be processed, and the necessity of merging information prior to using it.

Computers have modifyed paradigms of scientific investigation

- Computations that were impossible to run some time ago become feasible
- More and more data, including from human origin (like testimonies)
- A variety of data types, including images, natural language, etc.
- More and more sources of various origins : data bases, sensors, humans....

Importance of modeling the imperfection of information

The fantastic computation power available is counterbalanced by the possible lack of good quality of the data to be processed, and the necessity of merging information prior to using it.

Uncertainty

In consequence, *uncertainty* still pervades many of the conclusions we can draw from information we receive and we need to model it

Uncertainty

The lack of capability for an agent to answer questions of interest positively or negatively .

As we finally cannot access to as much information as we would like to :

- The more informative a statement, the more uncertain it may be.
- Useful statements : a balance between precision and certainty

Sources of uncertainty

- Randomness : observed instability of repeatable phenomena
- Lack of information : just missing data or lack of precision (incompleteness)
- Excess of information : many conflicting items from various sources

These aspects cannot be accounted for by a unique approach, like probability.

What to use when

- Randomness : additive probability theory
- Incompleteness : sets instead of points (logic, intervals, fuzzy sets)
- Inconsistency : set-theoretic connectives, aggregation functions, argumentation

Claim

Need to reconcile probability and logic

Aim : Constructing and quantifying beliefs

Historical aspects of uncertainty

In the XVIIth century, scholars distinguished between

- Chances : uncertainty resulting from games (flipping coins, dice, deck of cards)
- Probabilities : trust in potentially unreliable testimonies at courts of law

- Chances are objective, probabilities are subjective.
- Until the end of XVIIIth century, the problem of merging unreliable testimonies was important
- Some proposals of the time cannot be understood using standard probability theory

The divorce between probability and logic

- *Until the end of XIXth century* logic and probability go along together (Boole, Venn, De Morgan...),e.g. probabilistic syllogisms
- *First half of XXth century* : logic as the foundations of mathematics, probability as the foundation of statistics.
- *End of XXth century on* : artificial intelligence. Logic and probability to model human articulated reasoning
  - logical databases, epistemic logic, non-monotonic reasoning
  - Bayesian networks, possibilistic logic, Markov logic, theory of evidence, imprecise probabilities
  - multi-agent reasoning

Handling incomplete information

The basic approach relies on classical logic

- A collection of beliefs is modeled by a set of logical assertions.
- A statement p is **certainly true** if deducible from the belief base : $N(p) = 1$ (and $0 =$ not certain)
- A statement is **plausible** if it is consistent with the belief base : $\Pi(p) = 1$ (and $0 =$ impossible)

Duality property

A statement is certainly true if its negation is impossible :

$$N(p) = 1 - \Pi(not\ p)$$

$N(p\ and\ q) = \min(N(p), N(q))$; $\Pi(p\ or\ q) = \max(\Pi(p), \Pi(q))$

The basic approach relies on classical logic

- A collection of beliefs is modeled by a set of logical assertions.
- A statement p is **certainly true** if deducible from the belief base : $N(p) = 1$ (and $0 =$ not certain)
- A statement is **plausible** if it is consistent with the belief base : $\Pi(p) = 1$ (and $0 =$ impossible)

Duality property

A statement is certainly true if its negation is impossible :

$$N(p) = 1 - \Pi(not\ p)$$

$N(p\ and\ q) = \min(N(p), N(q))$; $\Pi(p\ or\ q) = \max(\Pi(p), \Pi(q))$

Human originated information : incomplete and non-Boolean

It is in natural language hence gradual (fuzzy) : truth becomes a matter of degree(Zadeh)

- Linguistic terms referring to measurable scales (no meaningful threshold between yes and no)
- Typicality relations underlying linguistic terms (no flat extension to concepts)

The non-Boolean truth scale make fuzzy concepts commensurate

Interesting issues

- How to extend logical connectives (conjunction, disjunction implication)
- Other aggregation functions : means, uninorms, nullnorms (J. Fodor)
- Can we build syntactic logical systems like in the Boolean case ?

Human originated information : incomplete and non-Boolean

It is in natural language hence gradual (fuzzy) : truth becomes a matter of degree(Zadeh)

- Linguistic terms referring to measurable scales (no meaningful threshold between yes and no)
- Typicality relations underlying linguistic terms (no flat extension to concepts)

The non-Boolean truth scale make fuzzy concepts commensurate

Interesting issues

- How to extend logical connectives (conjunction, disjunction implication)
- Other aggregation functions : means, uninorms, nullnorms (J. Fodor)
- Can we build syntactic logical systems like in the Boolean case ?

A uniform probability cannot model ignorance

- Confusion between randomness and lack of information using subjective probability
- Uniform distributions are not scale invariant
- In the face of partial ignorance, people do not always make decisions based on expectations (Ellsberg Paradox)

Three epistemic values

Certainty that yes, certainty that no, uncertainty (ignorance)

Need two set functions.

Possibility theory for incomplete information

Possibility distributions (Zadeh) : represent states of information as sets of more or less plausible states of facts.

- *Degree of certainty* $N(p)$: to what extent p is true in all the most plausible situations
- *Degree of possibility* $\Pi(p)$: to what extent p is true in at least one plausible situation

$$N(p) = 1 - \Pi(not\ p);$$

$$N(p\ and\ q) = \min(N(p), N(q)); \Pi(p\ or\ q) = \max(\Pi(p), \Pi(q))$$

Remarks

- Degrees of possibility need not be numerical : order is enough
- Ignorance : everything is possible
- Precise information : only one state of facts is possible

Possibility theory and logic

Given a possibility distribution over a set of possible situations, the family of propositions with a certainty degree higher than a threshold is deductively closed

Possibilistic logic

- Handles propositional formulas with attached certainty degrees
- The validity of a reasoning path is the validity of the weakest link.

At each level of certainty, reasoning in possibilistic logic is like reasoning in classical logic

Possibility vs. probability

A clear contrast

- Probability : precise and scattered pieces of information (sensors)
- Possibility : imprecise but coherent pieces of information (human expert)

A fuzzy set (understood as a possibility distribution) can also model

- A likelihood function in the sense of the maximum likelihood principle
- A nested family of confidence intervals
- A kind of cumulative distribution function obtained via probabilistic inequalities (Chebyshev)

Possibility vs. probability

A clear contrast

- Probability : precise and scattered pieces of information (sensors)
- Possibility : imprecise but coherent pieces of information (human expert)

A fuzzy set (understood as a possibility distribution) can also model

- A likelihood function in the sense of the maximum likelihood principle
- A nested family of confidence intervals
- A kind of cumulative distribution function obtained via probabilistic inequalities (Chebyshev)

Blending probability and sets

Modeling epistemic states by

- *Sets of probability measures* (credal sets) : ill-known probabilistic models or subjective probabilities with non-echangeable bets : lower previsions (P. Walley)
- *Random sets (evidence theory)* : statistics with incomplete observations (A. Dempster), or unreliable imprecise testimonies (G. Shafer)

Generalisations of both probability and possibility theories that use distinct set functions for certainty and plausibility : upper and lower probabilities.

Conjoint generalisations of set-theoretic and probabilistic notions

Logical connectives, conditioning, expectation.

Main ideas

- Choose the formal representation of data *in agreement with its level of precision* : probability, intervals, fuzzy intervals, etc.
- Propagate probabilistic and incomplete information *conjointly* through a mathematical model.
- quantify the amount of belief in the risky event, and *also the amount of ignorance* about it

Main contribution

Choose between collecting more information and taking action to protect against variability.

A reference : G. Bárdossy, J. Fodor, Evaluation of Uncertainties and Risks in Geology, Springer, 2003.

Information fusion

Incomplete information coming from several sources : expert opinions, databases, sensors, etc.

Aim : exploit conflict between sources to extract plausible information

- merging techniques use extensions of conjunction, disjunction and averaging operations
- rely on consistent subsets of sources
- lay bare and ranking alternative remaining possibilities after fusion

Merging techniques

- weighted average vs. Bayesian methods in probability
- Using fuzzy set connectives in possibility theory
- Dempster's rule of combination and variants in evidence theory

Decision under incomplete information

- Incompleteness limits our capability to choose between options
- Either accept *undecisiveness* or introduce *decision-maker attitude* (pessimism, optimism) to solve incomparabilities
- ordinal representations face impossibility theorems, while numerical approaches may be scale-dependent

Main decision rules beyond expected utilities

- maxmin (optimistic) and minmax (pessimistic) rules in possibility theory
- Lower expectations (pessimistic) based on Choquet integrals
- Lower expectation of difference between gambles (Walley)

Conclusion : the nature of epistemic models

Epistemic modeling accounts for incompleteness of information, e.g. set-valued functions

- It is not usual : In classical approaches, a model is an approximate but precise representation, *a substitute* of reality.
- An imprecise model
  - is of higher order, not objective (it is observer-dependent)
  - it represents *incomplete knowledge about reality*
  - we may wish to check it encloses reality between its bounds
  - we may strive to make it more precise by acquiring more information.

One should reconsider system modeling theories under the epistemic point of view

References

- D.D., H. Prade, Possibility Theory. Plenum Press, New York
- D. D., J.C. Fodor, H. Prade, M. Roubens : Aggregation of decomposable measures with application to utility theory. Theory and Decision, 41, 1996, 59-95.
- D. D., D. Guyonnet. Risk-informed decision-making in the presence of epistemic uncertainty. In : International Journal of General Systems, Vol. 40 N. 2, p. 145-167, 2011.
- D. D., H. Prade (2014) Possibilistic Logic Ñ An Overview In JŽrg H. Siekmann, Ed. : Handbook of the History of Logic, Volume 9, 283-342
- I. Couso, D. Dubois, L. Sanchez. Random Sets and Random Fuzzy Sets as Ill-Perceived Random Variables, Springer, 2014.
- D.D., W. Liu, J. Ma, H. Prade. The basic principles of uncertain information fusion. Information Fusion, Vol. 32, p. 12-39, 2016.