Uncertainty in Software Models

Lola Burgueño, Paula Muñoz, Antonio Vallecillo
Universidad de Málaga, Spain

Málaga, XX Noviembre 2021

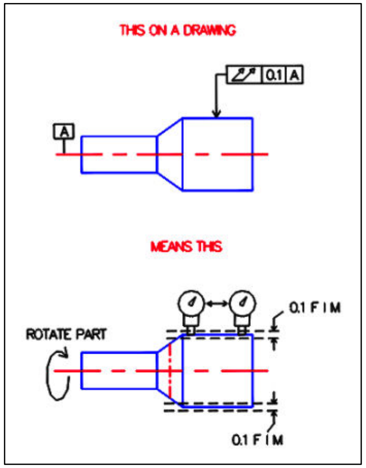
Uncertainty

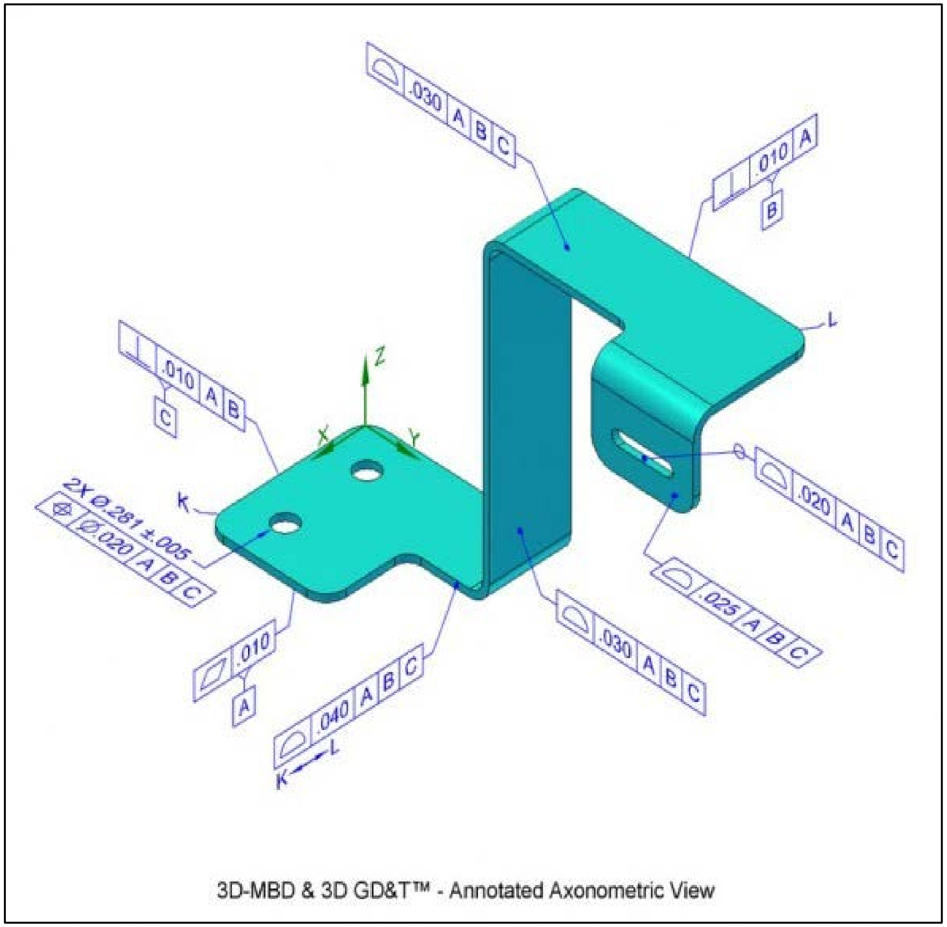
Uncertainty: Quality or state that involves imperfect and/or unknown information

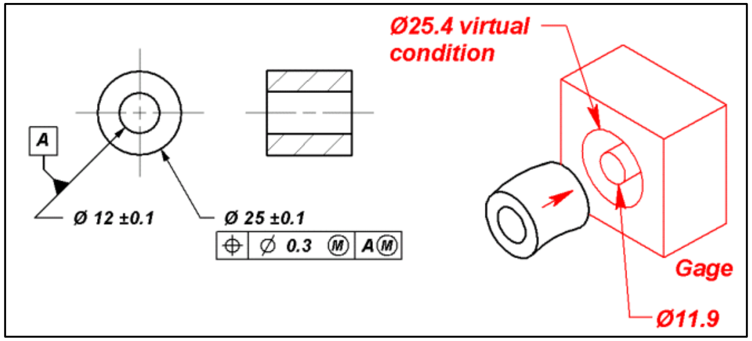
- It applies to: predictions of future events,
  estimations,
  physical measurements, or
  properties of a system, its elements or its environment
- due to:
  - Underspecification of the problem or solution domains
  - Lack of knowledge of the system, its environment, or its underlying physics
  - Lack of precision in measurements
  - Imperfect, incorrect, or missing information
  - Numerical approximations
  - Values and parameters indeterminacy
  - Different interpretations of the same evidences by separate parties

"There is nothing certain, but the uncertain" (*proverb*)

Uncertainty in mechanical systems

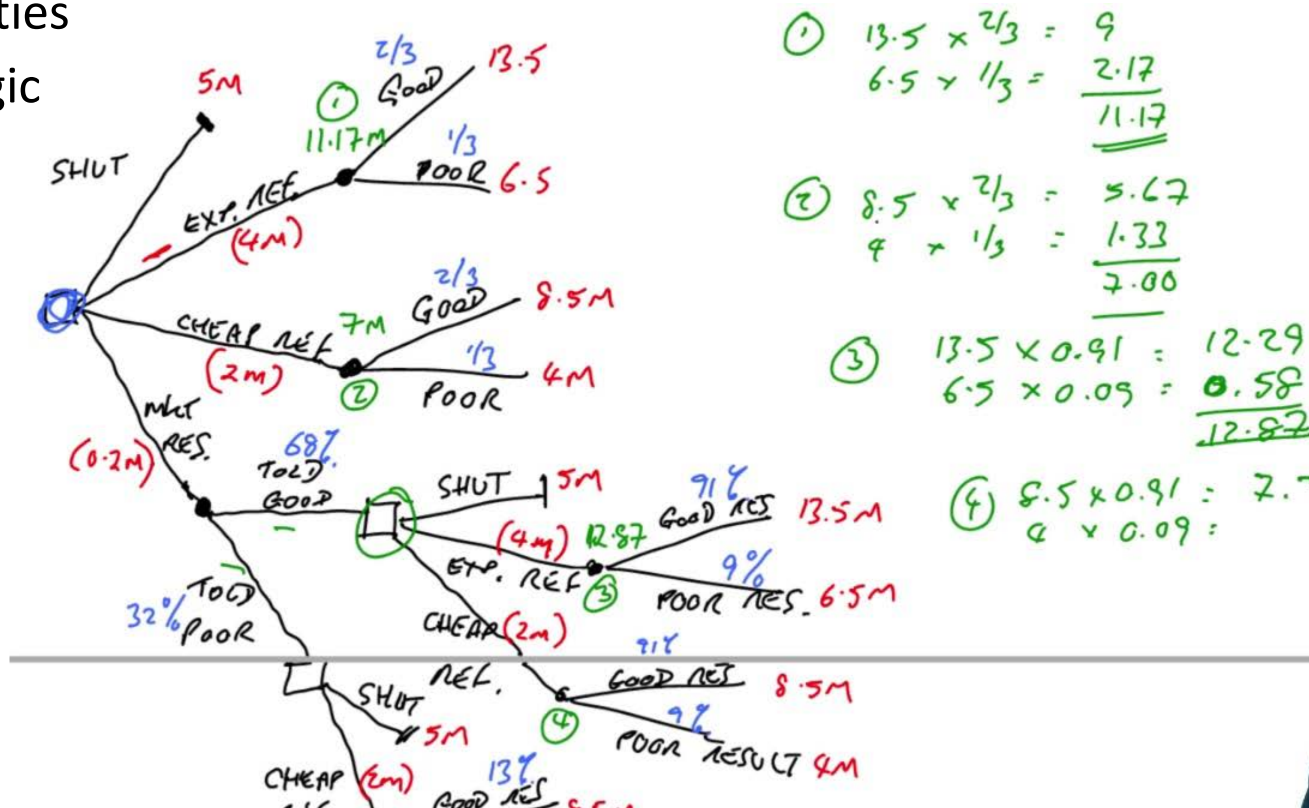






Many different formalisms and theories to quantify uncertainty

- Bayesian Belief Networks (BBN)
- Monte Carlo simulations
- Decision theory/trees
- Probabilities
- Fuzzy Logic
- …



A classification of uncertainty (according to its nature)

- **Aleatory** Uncertainty – A kind of uncertainty that refers to the *inherent* uncertainty due to the *probabilistic variability or randomness* of a phenomenon
  - Examples: measuring the speed of a car, or the duration of a software development process
  - This type of uncertainty is **irreducible**, in that there will always be variability in the underlying variables.

- **Epistemic** Uncertainty – A kind of uncertainty that refers to *the lack of knowledge* we may have about the system (modeled or real).
  - Examples: Ambiguous or imprecise requirements about the expected system functionality, its envisioned operating environment, etc.
  - This type of uncertainty is **reducible**, in that additional information or knowledge may reduce it.

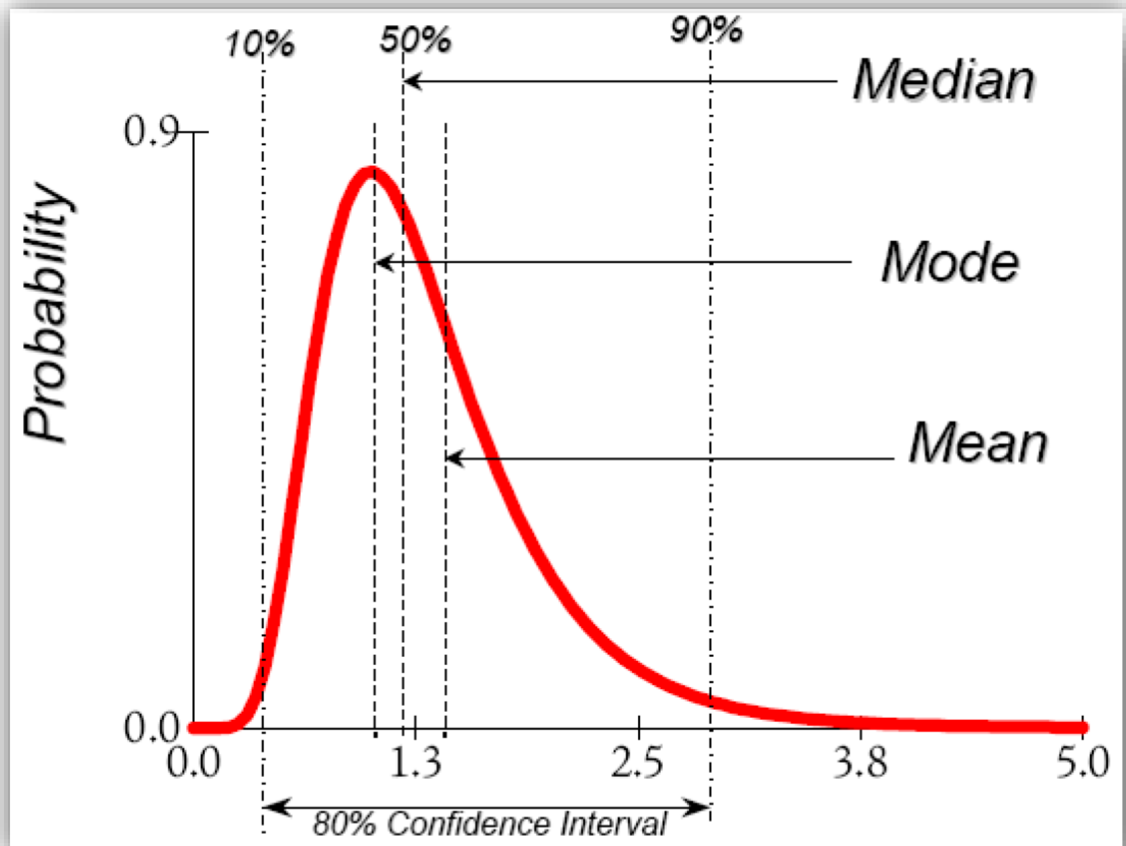
A. Der Kiureghian and O. Ditlevsen: "Aleatory or epistemic? Does it matter?" Structural Safety 31(2):105-112, 2009

Types of uncertainty (according to their sources)

- **Measurement** uncertainty: A kind of *aleatory* uncertainty that refers to a set of possible states or outcomes of a measurement, where probabilities are assigned to each possible state or outcome
- **Occurrence** uncertainty: a kind of *epistemic* uncertainty that refers to the degree of belief that we have on the actual existence of an entity, i.e., the real entity that a model element represents
- **Belief** uncertainty: A kind of *epistemic* uncertainty in which a *belief agent* is uncertain about any of the *statements* made about the system or its environment.
- **Design** uncertainty: A kind of *epistemic* uncertainty that refers to a set of possible design decisions or options, where probabilities are assigned to each decision or option
- **Environment** uncertainty: lack of certainty about the surroundings, boundaries and usages of a system and of its elements
- **Location** uncertainty: lack of certainty about the geographical or physical location of a system, its elements or its environment
- **Time** uncertainty: lack of certainty about the time properties expressed in a statement about the system or its environment

Based on M. Zhang, B. Selic, S. Ali, T. Yue, O. Okariz, and R. Norgren, "Understanding Uncertainty in Cyber-Physical Systems: A Conceptual Model" In Proc. of ECMFA 2016, LNCS vol. 9764, pp. 247-264. Springer, 2016.

Measurement Uncertainty



Measurement uncertainty

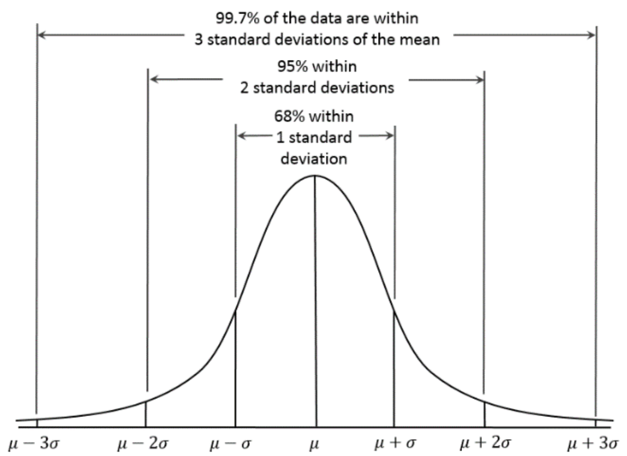
- **Measurement** uncertainty: A kind of *aleatory* uncertainty that refers to a set of possible states or outcomes of a measurement

- Normally expressed by a parameter, associated with the result of a measurement x, that characterizes the dispersion of the values that could reasonably be attributed to the measurand: the standard deviation u of the possible variation of the values of x

- Representation: $\boldsymbol{x} \pm \boldsymbol{u}$ or (x, u)

- *Examples:*
  - Normal distribution: (x, σ) with mean x, and and standard deviation σ
  - Interval $[a, b]$: Uniform distribution is assumed

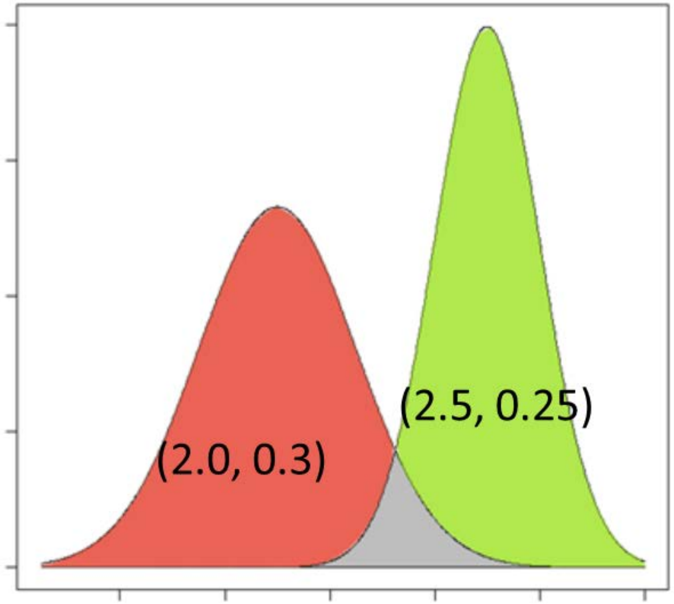
$$(x, u) \text{ with } x = \frac{a+b}{2}, u = \frac{(b-a)}{2\sqrt{3}}$$



JCGM 100:2008. Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM). http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

Some problems with Measurement Uncertainty

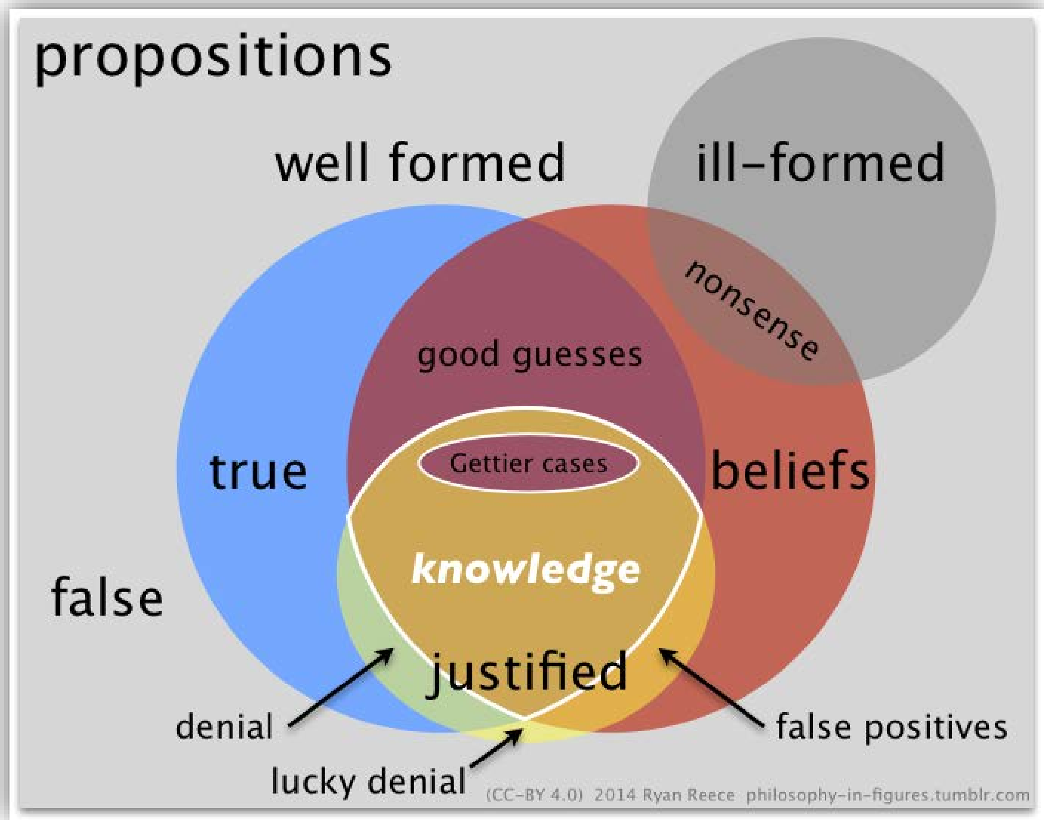
- Computations with uncertain values must respect the *propagation of uncertainty* (uncertainty analysis)
  - In general, this is a complex problem, which cannot be manually managed
- *Comparison of uncertain values* is no longer a Boolean property!
  - How to compare 17.7 ± 0.2 with 17.8 ± 0.2?
- Other primitive datatypes are also affected by uncertainty
  - Strings (OCR)
  - Enumerations
  - Collections



UBoolean

- UBooleans are pairs (b, c)
  - where b:Boolean and c:Real, c ϵ [0, 1]
  - c represents the confidence that the actual value of the value is indeed b
  - Canonical form: (true, c)
  - Equivalence relation: (b, c) = (not b, 1 - c)
- Constants
  - UBoolean(true, 0.999), UBoolean(false, 0.001)
- Operations
  - Redefined basic operations: and, or, not
  - Redefined secondary operations: implies, equivalent, xor
  - Conversion operations: toBoolean() and toBooleanC(c:Real)

Belief Uncertainty



Belief uncertainty

- **Belief** uncertainty: A kind of *epistemic* uncertainty in which the modeler, or any other *belief agent*, is uncertain about any of the *statements* made about the system or its environment.
  - By nature, it is always *subjective*
  - *It may not always be possible to determine whether or not a belief statement is valid*.
    - A belief statement may not necessarily correspond to objective reality.
    - This means that it could be completely false, or only partially true, or completely true.
  - The validity of a statement may only be meaningfully defined within a given context or purpose.
    - Thus, the statement that “the Earth can be represented as a perfect sphere” may be perfectly valid for some purposes but invalid or only partly valid for others.

- **Belief agent:** An entity (human, institution, even a machine) that holds one or more beliefs

- **Belief statement:** *Statement* qualified by a *degree of belief*

- **Degree of belief:** Confidence assigned to a statement by a belief agent. Normally expressed by quantitative or qualitative methods (e.g., a grade or a probability “*credence*”)

Loli Burgueño, Robert Clarisó, Jordi Cabot, Sébastien Gérard, Antonio Vallecillo. “Belief uncertainty in software models.” Proc. of MiSE 2019@ICSE, pp. 19-26. ACM, 2019. https://dl.acm.org/citation.cfm?id=3340709

Classical Boolean Logic

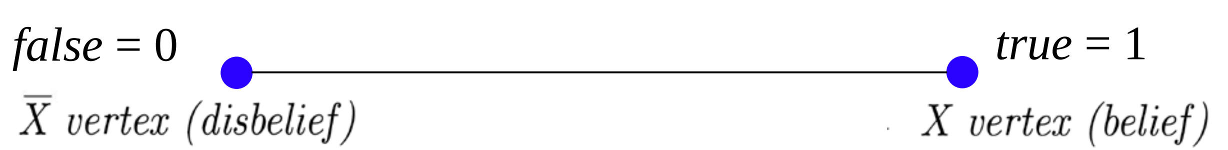
false = 0

$\overline{X}$ *vertex (disbelief)*

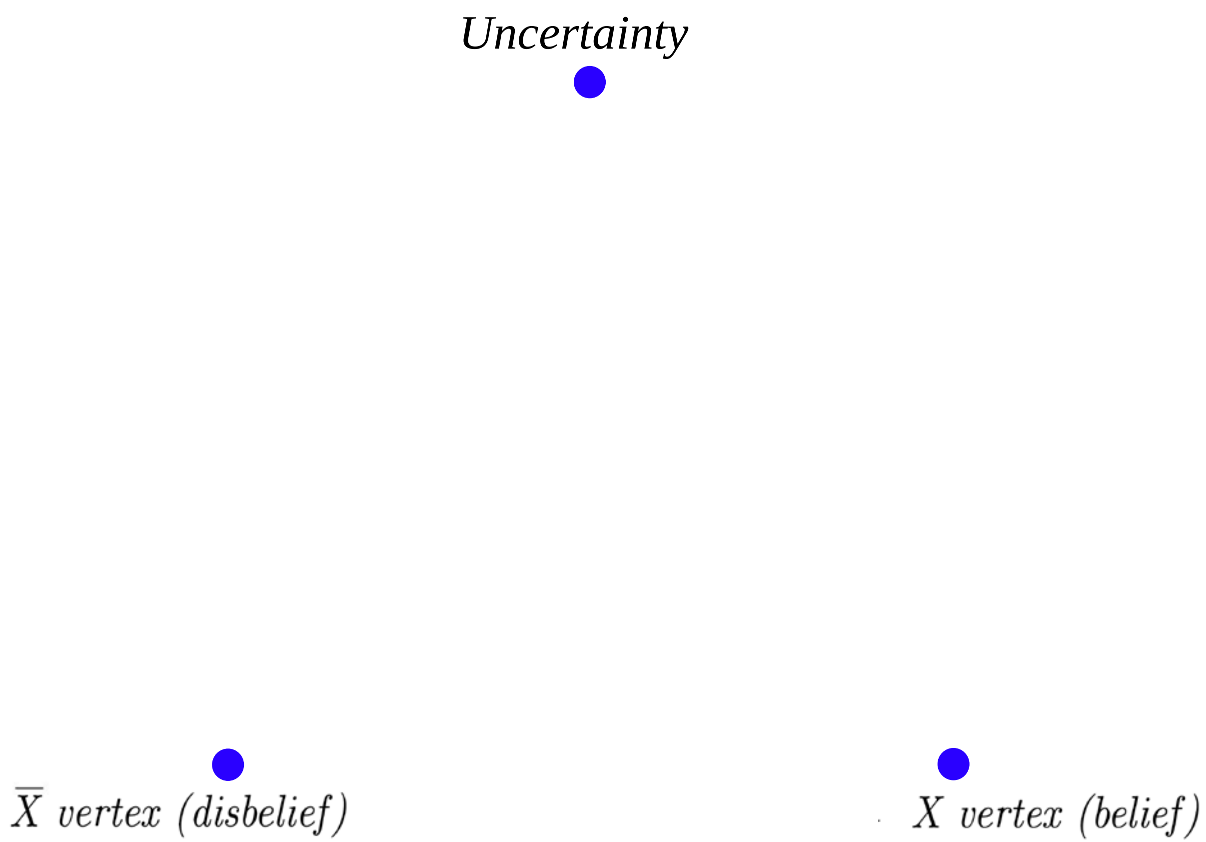
true = 1

X *vertex (belief)*

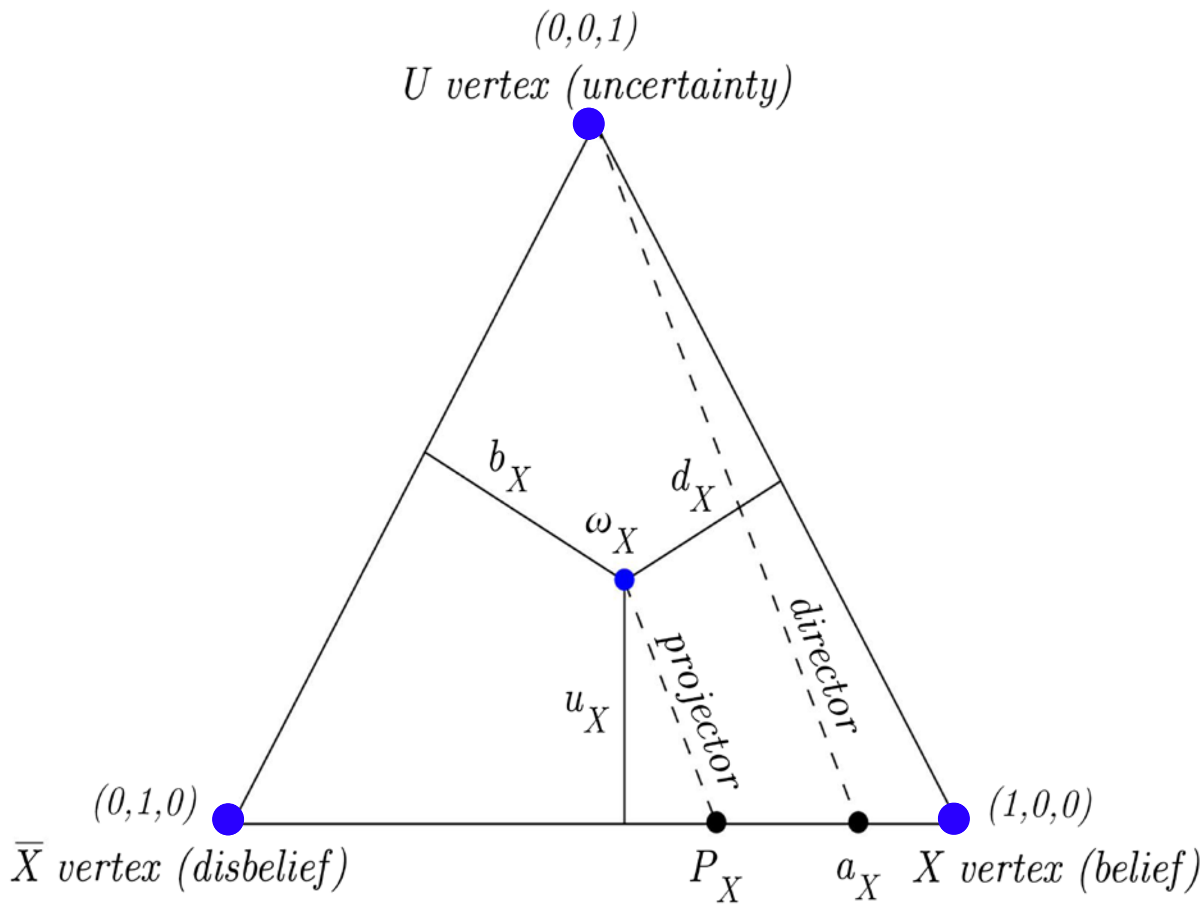
Probabilistic Logic



Kleene (three-valued) Logic



Subjective Logic



Subjective logic in UML/OCL

- `SBoolean(`*b*`,`*d*`,`*u*`,`*a*`)`
  - *b* represents the degree of ***belief*** that the agent has about the statement
  - *d* represents the degree of ***disbelief***
  - *u* represents the ***uncertainty*** that the agent expressing the opinion has about the statement, i.e., the degree of trust
  - *a* is the (objective) ***prior probability*** assigned to the statement (also called "***base rate***").

- $b + d + u = 1$

- Boolean values lifted to SBoolean:
  - `true  ≡ SBoolean(1, 0, 0, 1)`
  - `false ≡ SBoolean(0, 1, 0, 0)`
- UBoolean values lifted to SBoolean:
  - `Uboolean(true,c) = SBoolean(c,1-c,0,c)`
- Projection from SBoolean to UBoolean:
  - `SBoolean(`*b*`,`*d*`,`*u*`,`*a*`).projection() = UBoolean(true, `*b + u*a*`)`