Managing Imperfect Information Using Imperfect Rules

Approximate Guidelines

Davide Sottara, Ph.D.
Research Scientist
University of Bologna

Outline

1. Imperfection : Why and What
2. Imperfection : How
   - On Premise Evaluation
   - On Conclusion Entailment
3. Conclusions

Outline

Basic Inference

$$\frac{\langle P(x), P(X) \rightarrow C(Y) \rangle}{C(y)}$$

- Classic Modus Ponens
  - Premise and Implication entail Consequence

Basic Inference

$$\frac{\langle P(x), P(X) \rightarrow C(Y) \rangle}{C(y)}$$

- Classic Modus Ponens
  - Premise and Implication entail Consequence

Is this model really applicable in practice?

Once upon a time...

A well known story... [Bezdek] :

$$thirsty(X) \wedge has(X, Bottle) \wedge \neg lethal(Bottle) \Rightarrow drink(Bottle)$$

Once upon a time...

A well known story... [Bezdek] :

$$thirsty(X) \wedge has(X, Bottle) \wedge \neg lethal(Bottle) \Rightarrow drink(Bottle)$$

- Lack of information
- Ill-defined information
- Erroneous information
- ...

What is Imperfection?

Imperfection

Imperfection, be it Imprecision or Uncertainty, pervades . . . systems that attempt to provide an accurate model of the real world

P.Smets, 1999

What is Imperfection?

Imperfection

Imperfection, be it Imprecision or Uncertainty, pervades . . . systems that attempt to provide an accurate model of the real world

P.Smets, 1999

Uncertainty

Uncertainty is a condition where Boolean truth values are unknown, unknowable, or inapplicable . . .
W3C Incubator Group on Uncertainty Reasoning for the Web, 2005

What is Imperfection?

Imperfection - a negative definition

Uncertainty/Imperfection is the opposite of preciseness and certainty, i.e. of what Boolean logic models

Using Imperfection

Rules should handle imperfection, not ignore it

Benefits

- Conciseness
- Robustness

Drawbacks

- Complexity
- Correctness and Coherence

More remarks

My three cents on predictive modelling technologies...

More remarks

My three cents on predictive modelling technologies...

- Knowledge
- Business Rules
- Predictive Models

More remarks

Knowledge

- Business Rules
- Predictive Models

More remarks

Knowledge

- “Hard” AI
- “Soft” AI

More remarks

Knowledge

- Perfect AI
- Imperfect AI

More remarks

Knowledge

- Perfect AI
- Imperfect AI

We have a (serious) problem...

More remarks

Knowledge

- Perfect AI
- Imperfect AI

We have a (serious) problem...

Business Rules $\subset$ Rules

More remarks

Knowledge

- Perfect AI
- Imperfect AI

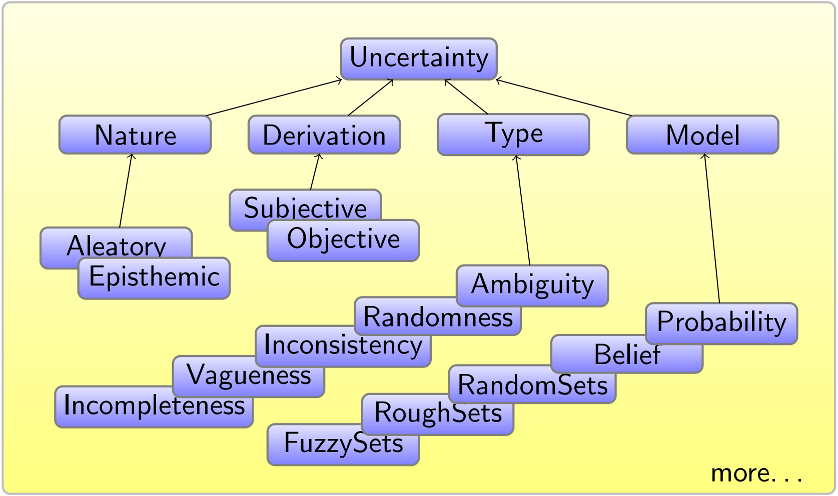
We have a (serious) problem...

Business Rules $\subset$ Rule-Based Programming

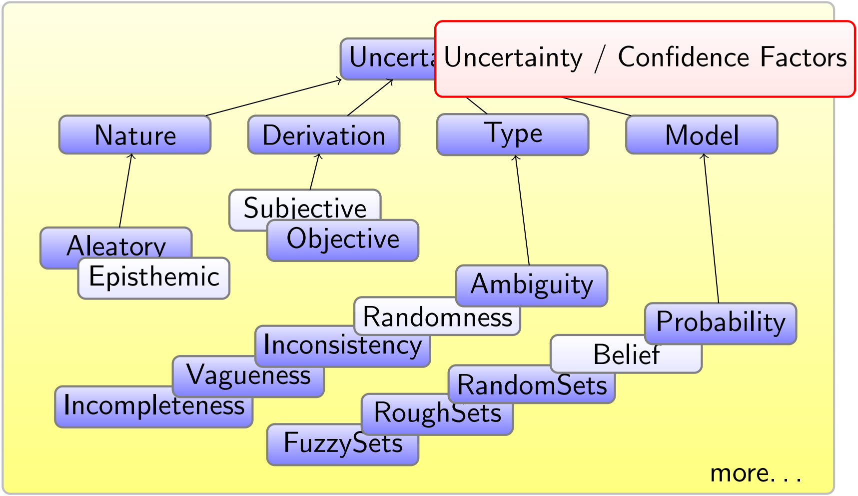
Outline

1. Imperfection : Why and What
2. Imperfection : How
   - On Premise Evaluation
   - On Conclusion Entailment
3. Conclusions

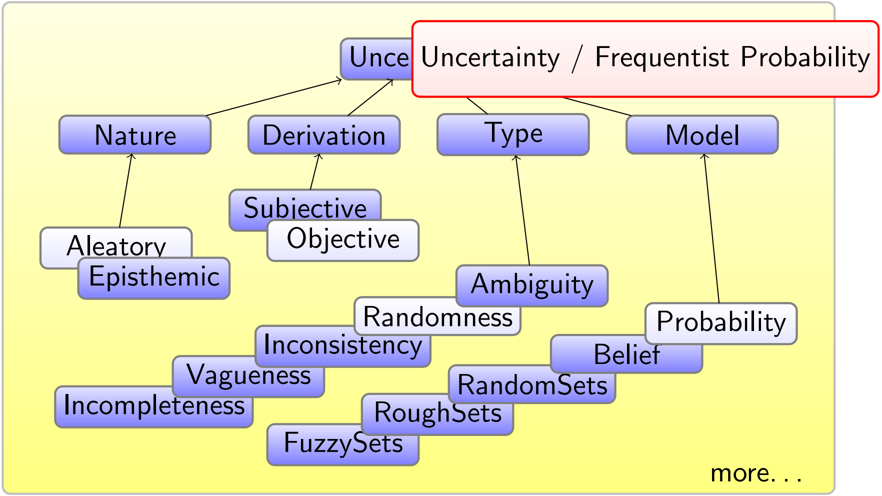
An Ontology for Imperfection



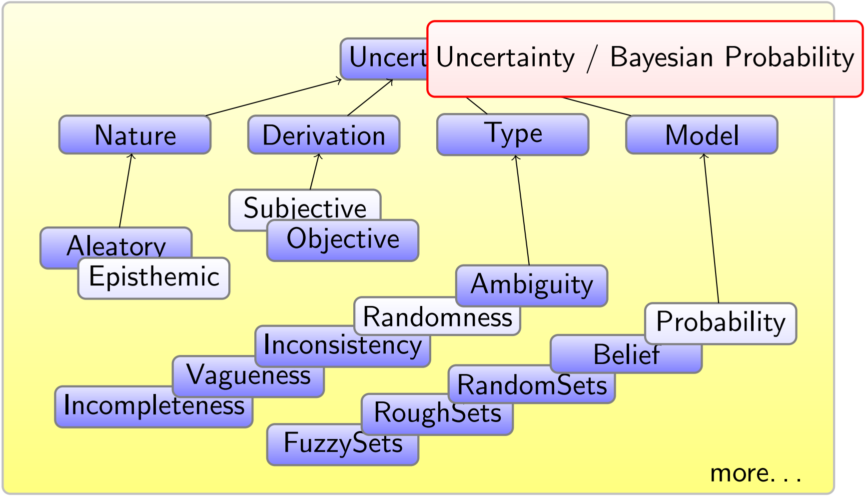
An Ontology for Imperfection



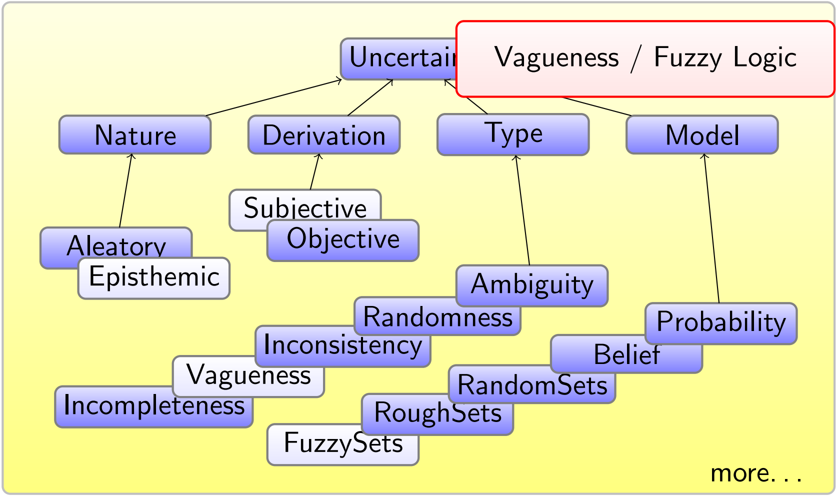
An Ontology for Imperfection



An Ontology for Imperfection



An Ontology for Imperfection



Generalized Inference

$$\frac{\langle P(x), P(X) \rightarrow C(Y) \rangle}{C(y)}$$

- Classic Modus Ponens
  - Premise and Implication entail Consequence

Generalized Inference

$$\frac{\langle \Phi(\ldots, A_j(x)/\varepsilon_j, \ldots), P(X) \rightarrow C(Y) \rangle}{C(y)}$$

- Premise
  - Atomic constraints are evaluated
  - General, pluggable `Evaluators`
  - A `Degree` is returned

Generalized Inference

$$\frac{\langle \Phi(\ldots, A_j(x)/\varepsilon_j, \ldots)/\varepsilon_P, P(X) \rightarrow C(Y) \rangle}{C(y)}$$

- Premise
  - Atomic constraints are evaluated
  - General, pluggable `Evaluators`
  - A `Degree` is returned

- Premise
  - Atoms are aggregated in formulas
  - using generalized logic Connectives
  - evaluated by `Operators`

Generalized Inference

$$\frac{\langle P(x)/\varepsilon_P \;,\; \rightarrow_{(X,Y)}/\varepsilon_{\rightarrow}\rangle}{C(y)}$$

- Implication
  - Implication has a Degree
  - often given *a priori*

Generalized Inference

$$\frac{\langle P(x)/\varepsilon_P \;,\; \rightarrow(X,Y)/\varepsilon_\rightarrow \rangle}{C(y)/\varepsilon_C}$$

- Implication
  - Implication has a Degree
  - often given *a priori*

- Modus Ponens
  - MP computes the Degree of the Consequence

Generalized Inference

$$\frac{\frac{\langle P_1, \rightarrow_1 \rangle}{C_1 /_{\varepsilon} C_1}, \ldots, \frac{\langle P_n, \rightarrow_n \rangle}{C_n /_{\varepsilon} C_n}}{C(y) /_{\varepsilon} C}$$

- Merging multiple sources
  - Multiple premises for the same conclusion
  - Solve conflicts
  - Handle missing values

Outline

Generalized Inference - Imperfect Facts

$$\frac{\langle \Phi(\ldots, A_j(x)/\varepsilon_j, \ldots), P(X) \rightarrow C(Y) \rangle}{C(y)}$$

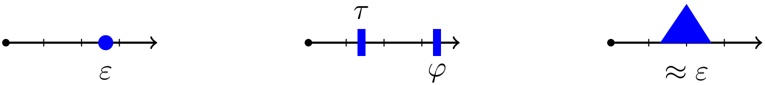
- Premise : (Complex) Formulas
  - Atomic constraints are evaluated
  - A `Degree` is returned

Generalized Degrees

Degrees generalize the boolean `true/false`

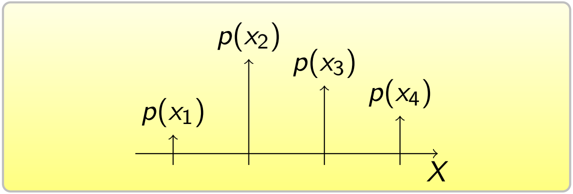
- **truth**: compatibility with a prototype
- **probability**: ratio of relevant events over total
- **belief**: opinion in assuming a property to be true.
- **possibility**: disposition towards accepting a situation to be true.
- **confidence**: strength of an agent's belief in a statement.

Different models, including:



Simple Interval Type-II degrees

Uncertainty : Frequentist Probability



Objective Probabilities

- Repeated trials

Random Variables

- 1 out of N
- Expected value

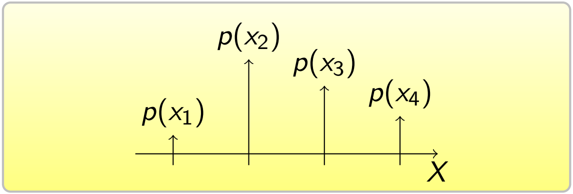
Applications

- Predict unobservable events
- What-if scenarios

Drawbacks

- High amount of experimental data

Uncertainty : Frequentist Probability



Objective Probabilities

- Repeated trials

Random Variables

- 1 out of N
- Expected value

Applications

- Predict unobservable events
- What-if scenarios

Drawbacks

- High amount of experimental data

Evaluation:

$$\text{somePredicate}(\ X/y\ ) \rightarrow p(X = y)$$

Uncertainty : Bayesian Probability

Subjective Probabilities

- Non-repeatable events
- Prior, Conditional and Posterior

Parametric Models

- Gaussian
- Poisson
- Beta, Gamma
- ...

Applications

- Express **belief** on the outcome of an event

Drawbacks

- Model selection
- Parameter estimation

Uncertainty : Bayesian Probability

Subjective Probabilities

- Non-repeatable events
- Prior, Conditional and Posterior

Parametric Models

- Gaussian
- Poisson
- Beta, Gamma
- ...

Applications

- Express **belief** on the outcome of an event

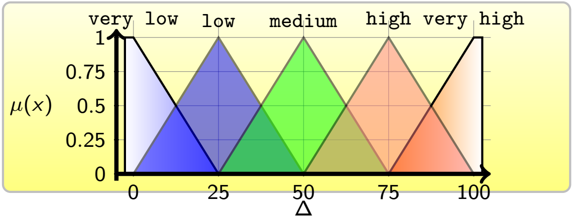
Drawbacks

- Model selection
- Parameter estimation

How to build that belief? **Evidence**:

$$p(a|b) \leftarrow p(a) \cdot \frac{p(b|a)}{p(b)}$$

Vagueness : Many-Valued and Fuzzy Logic



Partial membership

- Tertium datur
- Fuzzy sets define predicates

Domains

- Quantitative domain
- Qualitative property

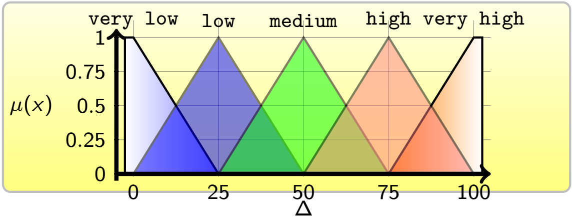
Applications

- Gradual Properties

Drawbacks

- Don't overlook definitions

Vagueness : Many-Valued and Fuzzy Logic



Partial membership

- Tertium datur
- Fuzzy sets define predicates

Domains

- Quantitative domain
- Qualitative property

Applications

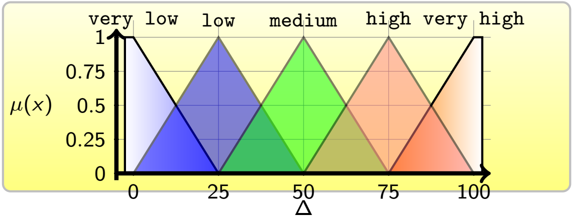
- Gradual Properties

Drawbacks

- Don't overlook definitions

Approximate relations/functions

Vague Uncertainty : Possibility



Possibility

- Not probability!

Linguistic Variables

- Vague statements
- Fuzzy sets as values

Applications

- Manage vague facts

Drawbacks

- "Defuzzification"

Examples I

- ?- age(davide,X)
  $\{X/20 : 30\%, X/30 : 60\%\}$

Examples I

- ?- age(davide,X)
  $\{X/20 : 30\%, X/30 : 60\%\}$
- ?- age(davide,X)
  $p(X) \propto \mathcal{N}(30, 2)$

Examples I

- ?- age(davide,X)
  $\{X/20 : 30\%, X/30 : 60\%\}$
- ?- age(davide,X)
  $p(X) \propto \mathcal{N}(30, 2)$
- age(davide,30)
  $\Rightarrow \{old(davide)/0.25, young(davide)/0.75\}$

Examples I

- ?- age(davide,X)
  $\{X/20 : 30\%, X/30 : 60\%\}$
- ?- age(davide,X)
  $p(X) \propto \mathcal{N}(30, 2)$
- age(davide,30)
  $\Rightarrow \{old(davide)/0.25, young(davide)/0.75\}$
- young(davide).

Examples II

- ?- hasDisease(davide,X)
  $\{X/cold : 30\%, X/allergy : 10\%, X/itchynose : 60\%\}$

Examples II

- ?- hasDisease(davide,X)
  $\{X/cold : 30\%, X/allergy : 10\%, X/itchynose : 60\%\}$
- hasDisease(davide,cold)
  $\{true/60\%, false/20\%\}$
- hasDisease(davide,allergy)
  $\{true/30\%, false/70\%\}$
- hasDisease(davide,itchy nose)
  $\{true/40\%\}$

Generalized Inference

$$\frac{\langle \Phi(\ldots,A_j(x)/\varepsilon_j,\ldots)/\varepsilon_P, P(X)\rightarrow C(Y)\rangle}{C(y)}$$

- Premise
  - Atomic constraints are evaluated
  - General, pluggable `Evaluators`
  - A `Degree` is returned

- Premise
  - Atoms are aggregated in formulas
  - using generalized logic Connectives
  - evaluated by `Operators`

Custom Operators

$$\{(Args), Degree\}^n \mapsto Degree$$

- Define connectives $(\wedge, \vee, \neg, \ldots)$
- Return a *Degree*

Custom Operators

$\{(Args), Degree\}^n \mapsto Degree$

- Define connectives $(\wedge, \vee, \neg, \ldots)$
- Return a *Degree*

Truth-functionality ?

Custom Operators

$\{(Args), Degree\}^n \mapsto Degree$

- Define connectives $(\wedge, \vee, \neg, \ldots)$
- Return a *Degree*

Truth-functionality ?

- Vagueness
  - Truth-functional
  - Three main "families"
  - (...and many others)

- Uncertainty
  - Not Truth-functional...
  - (... unless assumptions are made)

Examples III

- ?- hasDisease(davide,X)
  $\{X/cold : 30\%, X/allergy : 10\%, X/itchynose : 60\%\}$

$$hasDisease(davide, cold) \vee hasDisease(davide, allergy)$$

Examples III

- ?- hasDisease(davide,X)
  $\{X/cold : 30\%, X/allergy : 10\%, X/itchynose : 60\%\}$

$$hasDisease(davide, cold) \vee hasDisease(davide, allergy)$$

Mutual exclusion : $p(cold)+p(allergy) = 40\%$

Examples IV

- hasDisease(davide,cold) $\{true/60\%, false/20\%\}$
- hasDisease(davide,allergy) $\{true/30\%, false/70\%\}$

$$hasDisease(davide, cold) \vee hasDisease(davide, allergy)$$

Examples IV

- hasDisease(davide,cold) $\{true/60\%, false/20\%\}$
- hasDisease(davide,allergy) $\{true/30\%, false/70\%\}$

$$hasDisease(davide, cold) \vee hasDisease(davide, allergy)$$

Independence ? : $p(cold) \oplus p(allergy) = (60 + 30 - 18)\%$

Examples IV

- hasDisease(davide,cold) $\{true/60\%, false/20\%\}$
- hasDisease(davide,allergy) $\{true/30\%, false/70\%\}$

$$hasDisease(davide, cold) \vee hasDisease(davide, allergy)$$

Independence ? : $p(cold) \oplus p(allergy) = (60 + 30 - 18)\%$

Non-interactivity ? : $p(cold) \vee p(allergy) = \max\{60, 30\}\%$

Examples IV

- hasDisease(davide,cold)
  $\{true/60\%, false/20\%\}$
- hasDisease(davide,allergy)
  $\{true/30\%, false/70\%\}$

$$hasDisease(davide, cold) \vee hasDisease(davide, allergy)$$

Independence ? : $p(cold) \oplus p(allergy) = (60 + 30 - 18)\%$
Non-interactivity ? : $p(cold) \vee p(allergy) = \max\{60, 30\}\%$
Non-Independence ? : we need e.g. $p(cold|allergy)$

Outline

1. Imperfection : Why and What
2. Imperfection : How
   - On Premise Evaluation
   - On Conclusion Entailment
3. Conclusions

Generalized Inference

$$\frac{\langle P(x)/\varepsilon_P \;,\; \rightarrow(X,Y)/\varepsilon_\rightarrow \rangle}{C(y)/\varepsilon_C}$$

- Implication
  - Implication has a Degree
  - often given *a priori*

- Modus Ponens
  - MP computes the Degree of the Consequence

Implication and Deduction

- Two operators
- $A \to B$ vs $A \Rightarrow B$

Implication and Deduction

- Two operators
- $A \rightarrow B$ vs $A \Rightarrow B$

- Vagueness
  - Gradual implications
  - Gradual rules
  - ...

- Uncertainty
  - Bayesian inference
  - Distribution Semantics
  - ...

Examples V

$$hasDisease(X, cold) \Rightarrow sneeze(X)$$

- Vague: "the more serious the cold, the stronger the sneeze"
  $\mu(hasDisease(X, cold)) \otimes \mu(hasDisease(X, cold) \to sneeze(X))$
- Statistical : "given that you have cold, you'll probably sneeze"
  $p(hasDisease(X, cold)) \times p(sneeze(X) | hasDisease(X, cold))/p(sneeze(X))$
- Epistemical : "is it true that cold causes sneeze?"
  $p(hasDisease(X, cold)) \times p(sneeze(X))$

Generalized Inference

$$\frac{\dfrac{\langle P_1, \rightarrow_1 \rangle}{C_1 /_{\varepsilon} C_1}, \ldots, \dfrac{\langle P_n, \rightarrow_n \rangle}{C_n /_{\varepsilon} C_n}}{C(y) /_{\varepsilon} C}$$

- Merging multiple sources
  - Multiple premises for the same conclusion
  - Solve conflicts
  - Handle missing values

Generalized Inference

$$\frac{\dfrac{\langle P_1, \rightarrow_1 \rangle}{C_1 /_{\varepsilon} C_1}, \ldots, \dfrac{\langle P_n, \rightarrow_n \rangle}{C_n /_{\varepsilon} C_n}}{C(y) /_{\varepsilon} C}$$

- Merging multiple sources
  - Multiple premises for the same conclusion
  - Solve conflicts
  - Handle missing values
- Usually or-like

Uncertainty : Dirichlet model

$$Dirichlet_N(\mathbf{p}|\beta) = \frac{\Gamma\left(\sum_{j=0}^{N-1} \beta_j + N\right)}{\prod_{j=0}^{N-1} \Gamma\left(\beta_j + 1\right)} \prod_{j=0}^{N-1} p_j^{\beta_j}$$

Subjective Discrete Probabilities

- Belief from experience

Confidence

- Adds variance to probabilities

Applications

- Incremental, adaptive systems

Drawbacks

- Bootstrap

Uncertainty : Dirichlet model

$$Dirichlet_N(\mathbf{p}|\beta) = \frac{\Gamma\left(\sum_{j=0}^{N-1} \beta_j + N\right)}{\prod_{j=0}^{N-1} \Gamma(\beta_j + 1)} \prod_{j=0}^{N-1} p_j^{\beta_j}$$

Subjective Discrete Probabilities

- Belief from experience

Confidence

- Adds variance to probabilities

Applications

- Incremental, adaptive systems

Drawbacks

- Bootstrap

Consider source reliability when merging

Dempster-Shafer Theory and TBM

$$m(A) = \frac{1}{1-K} \sum_{B \cap C = A \neq \emptyset} m_0(B) \star m_1(C)$$

Uncertain Probabilities

- Belief vs Plausibility

Set-valued events

- Ignorance
- Inconsistency

Applications

- Incremental, evidence-driven systems

Drawbacks

- Complexity

Outline

Conclusions

Rules should handle imperfection, not ignore it

Conclusions

Rules should handle imperfection, not ignore it

- Expressiveness
- Declarativeness
- Integration

BUT

Conclusions

Rules should handle imperfection, not ignore it

- Expressiveness
- Declarativeness
- Integration

BUT

Imperfection must be modelled correctly