# The role of coherent probabilities... nowadays

Andrea Capotorti

Dipartimento di Matematica e Informatica
Università degli Studi di Perugia
Italy

# Some "seeds" in the ground of AI

It's from the '90s that coherent (conditional or not) probability assessments have been proposed as good models for Expert Systems and Artificial intelligence, e.g.:

**1990**:

- G Coletti, A Gilio, R Scozzafava: **Coherent** qualitative probability and uncertainty in **Artificial Intelligence**, Proc. of *8th Int. Congr. of Cybernetics and Systems*
- G Coletti, A Gilio, R Scozzafava: **Conditional events** with vague information in **expert systems**, Proc. of *Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems*

# Some "seeds" in the ground of AI

It's from the '90s that coherent (conditional or not) probability assessments have been proposed as good models for Expert Systems and Artificial intelligence, e.g.:

**1990**:


- G Coletti, A Gilio, R Scozzafava: **Coherent** qualitative probability and uncertainty in **Artificial Intelligence**, Proc. of *8th Int. Congr. of Cybernetics and Systems*
- G Coletti, A Gilio, R Scozzafava: **Conditional events** with vague information in **expert systems**, Proc. of *Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems*

# Some "seeds" in the ground of AI

It's from the '90s that coherent (conditional or not) probability assessments have been proposed as good models for Expert Systems and Artificial intelligence, e.g.:

**1990**:

- G Coletti, A Gilio, R Scozzafava: **Coherent** qualitative probability and uncertainty in **Artificial Intelligence**, Proc. of *8th Int. Congr. of Cybernetics and Systems*
- G Coletti, A Gilio, R Scozzafava: **Conditional events** with vague information in **expert systems**, Proc. of *Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems*

# Some "seeds" in the ground of AI

**1991**:

- G Coletti, A Gilio, R Scozzafava: **Assessment** of qualitative judgements for **conditional events** in **expert systems**, Proc of *European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*

**1992**:

- G Coletti, G Regoli: How can an **expert system** help in choosing the optimal decision? *Theory and decision*

- A. Gilio, F. Spezzaferri: Knowledge integration for **conditional probability assessments**, In: "*Uncertainty in* **Artificial Intelligence**" (Eds. D. Dubois, M. P. Wellman, B. D'Ambrosio, P. Smets)

# Some "seeds" in the ground of AI

**1991**:

- G Coletti, A Gilio, R Scozzafava: **Assessment** of qualitative judgements for **conditional events** in **expert systems**, Proc of *European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*

**1992**:

- G Coletti, G Regoli: How can an **expert system** help in choosing the optimal decision? *Theory and decision*
- A. Gilio, F. Spezzaferri: Knowledge integration for **conditional probability assessments**, In: "*Uncertainty in* **Artificial Intelligence**" (Eds. D. Dubois, M. P. Wellman, B. D'Ambrosio, P. Smets)

# Some "seeds" in the ground of AI

**1991**:

- G Coletti, A Gilio, R Scozzafava: **Assessment** of qualitative judgements for **conditional events** in **expert systems**, Proc of *European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*

**1992**:

- G Coletti, G Regoli: How can an **expert system** help in choosing the optimal decision? *Theory and decision*
- A. Gilio, F. Spezzaferri: Knowledge integration for **conditional probability assessments**, In: "*Uncertainty in* ***Artificial Intelligence***" (Eds. D. Dubois, M. P. Wellman, B. D'Ambrosio, P. Smets)

# Some "seeds" in the ground of AI

**1993**:

- G. Coletti and R. Scozzafava: A **coherent** qualitative Bayes' theorem and its application in **artificial intelligence**, Proc. of *Int. Symposium on Uncertainty Modeling and Analysis*
- G. Coletti: Numerical and qualitative judgements in probabilistic **expert systems**, Proc. of *International Workshop on Probabilistic Methods in* **Expert Systems**
- G. Coletti: Comparative probabilities ruled by **coherence** conditions and its use in **expert systems**, *International Journal Of General System*
- R.Scozzafava: **Subjective probability** versus Belief Functions in **Artificial Intelligence**, *International Journal of General Systems*

# Some "seeds" in the ground of AI

**1993**:

- G. Coletti and R. Scozzafava: A **coherent** qualitative Bayes' theorem and its application in **artificial intelligence**, Proc. of *Int. Symposium on Uncertainty Modeling and Analysis*
- G. Coletti: Numerical and qualitative judgements in probabilistic **expert systems**, Proc. of *International Workshop on Probabilistic Methods in* **Expert Systems**
- G. Coletti: Comparative probabilities ruled by **coherence** conditions and its use in **expert systems**, *International Journal Of General System*
- R.Scozzafava: **Subjective probability** versus Belief Functions in **Artificial Intelligence**, *International Journal of General Systems*

# Some "seeds" in the ground of AI

**1995**:

- in *Mathematical Models for Handling Partial Knowledge in* **Artificial Intelligence** Ed: G. Coletti, D. Dubois, R. Scozzafava:
- P. Vicig: **Conditional and Comparative Probabilities** in **Artificial Intelligence**
- A. Gilio: **Algorithms** for Precise and Imprecise **Conditional Probability Assessments**.

# Some "seeds" in the ground of AI

**1995**:

- in *Mathematical Models for Handling Partial Knowledge in* **Artificial Intelligence** Ed: G. Coletti, D. Dubois, R. Scozzafava:
- P. Vicig: **Conditional and Comparative Probabilities** in **Artificial Intelligence**
- A. Gilio: **Algorithms** for Precise and Imprecise **Conditional Probability Assessments**.

# Some "seeds" in the ground of AI

**2001**:

- G. Coletti, R. Scozzafava: From **conditional events** to conditional measures: a new axiomatic approach, *Annals of Mathematics and* **Artificial Intelligence**
- A. Gilio, V. Biazzo, G. Sanfilippo: **Algorithms** for **coherence** checking and propagation of **conditional probability bounds**", Proc. of *KI-2001 Workshop on Uncertainty in* **Artificial Intelligence**
- B. Vantaggi: Conditional independence in a finite **coherent setting**, *Annals of Mathematics and* **Artificial Intelligence**

# Some "seeds" in the ground of AI

**2001**:

- G. Coletti, R. Scozzafava: From **conditional events** to conditional measures: a new axiomatic approach, *Annals of Mathematics and* **Artificial Intelligence**
- A. Gilio, V. Biazzo, G. Sanfilippo: **Algorithms** for **coherence** checking and propagation of **conditional probability bounds**", Proc. of *KI-2001 Workshop on Uncertainty in* **Artificial Intelligence**
- B. Vantaggi: Conditional independence in a finite **coherent setting**, *Annals of Mathematics and* **Artificial Intelligence**

# Some "seeds" in the ground of AI

**2002**: in a special issue of *Annals of Mathematics and* **Artificial Intelligence**:

- M. Baioletti, A. Capotorti, S. Tulipani, B. Vantaggi: Simplification Rules for the **Coherent Probability Assessment** Problem
- A. Capotorti, B. Vantaggi: Locally strong **coherence** in inferential processes
- G. Coletti, R. Scozzafava: Stochastic independence in a **coherent setting**
- A. Gilio: Probabilistic Reasoning Under **Coherence** in System-P
- A. Gilio, V. Biazzo: On the linear structure of betting criterion and the checking of **coherence**

# Some "seeds" in the ground of AI

**2002**: in a special issue of *Annals of Mathematics and* **Artificial Intelligence**:

- M. Baioletti, A. Capotorti, S. Tulipani, B. Vantaggi: Simplification Rules for the **Coherent Probability Assessment** Problem
- A. Capotorti, B. Vantaggi: Locally strong **coherence** in inferential processes
- G. Coletti, R. Scozzafava: Stochastic independence in a **coherent setting**
- A. Gilio: Probabilistic Reasoning Under **Coherence** in System-P
- A. Gilio, V. Biazzo: On the linear structure of betting criterion and the checking of **coherence**

# Some "seeds" in the ground of AI

**2005**:

- A. Gilio, V. Biazzo, T. Lukasiewicz, G. Sanfilippo: Probabilistic logic under **coherence**: Complexity and **algorithms**, *Annals of Mathematics and* **Artificial Intelligence**

- A. Gilio, D. Dubois, G. Kern-Isberner: Probabilistic Abduction **Without Priors**, Dagstuhl Seminar Proc. on "*Belief Change in Rational Agents: Perspectives from* **Artificial Intelligence**, *Philosophy, and Economics*" (eds. J.P. Delgrande, J. Lang, H. Rott, and J.M. Tallon)

# Some "seeds" in the ground of AI

**2005**:

- A. Gilio, V. Biazzo, T. Lukasiewicz, G. Sanfilippo: Probabilistic logic under **coherence**: Complexity and **algorithms**, *Annals of Mathematics and* **Artificial Intelligence**

- A. Gilio, D. Dubois, G. Kern-Isberner: Probabilistic Abduction **Without Priors**, Dagstuhl Seminar Proc. on "*Belief Change in Rational Agents: Perspectives from* **Artificial Intelligence**, *Philosophy, and Economics*" (eds. J.P. Delgrande, J. Lang, H. Rott, and J.M. Tallon)

# And nowadays...

These authors have, as usual in scienze, "**anticipated the times**" because, meanwhile **A.I.**, **Big Data**, **Machine Learning**, **Alghoritms**, **IoT** have became "**reality**"... ... with the main-stream models running after the "*chimera*" of **full information**, **deep learning** on everything, **SMART** cities, etc. ... some more realistic position starts to appear, e.g....

# And nowadays...

These authors have, as usual in scienze, "**anticipated the times**" because, meanwhile **A.I.**, **Big Data**, **Machine Learning**, **Alghoritms**, **IoT** have became "**reality**"... ... with the main-stream models running after the "*chimera*" of **full information**, **deep learning** on everything, **SMART** cities, etc.

... some more realistic position starts to appear, e.g....

# And nowadays...

These authors have, as usual in scienze, "**anticipated the times**" because, meanwhile **A.I.**, **Big Data**, **Machine Learning**, **Alghoritms**, **IoT** have became "**reality**"... ... with the main-stream models running after the "*chimera*" of **full information**, **deep learning** on everything, **SMART** cities, etc.

... some more realistic position starts to appear, e.g....

# And nowadays...

from

R. She, S. Liu, S. Wan, K. Xiong and P. Fan, "Importance of Small Probability Events in Big Data: Information Measures, Applications, and Challenges," in IEEE Access, vol. 7, pp. 100363–100382, 2019

"In many applications (e.g., anomaly detection and security systems) of smart cities, **rare events** **dominate the importance of the total information of big data** collected by Internet of Things (IoTs). That is, it is pretty crucial to explore the **valuable information** associated with the rare events involved in **minority subsets** of the voluminous amounts of data."

# And nowadays...

from

Epistemic AI project https://www.epistemic-ai.eu/:
"Although artificial intelligence (AI) has improved remarkably over the last years, **its inability to deal with fundamental uncertainty** severely **limits its application**. This proposal re-imagines AI with a proper **treatment of the uncertainty stemming from our forcibly partial knowledge of the world**."
**(see the "Epistemic AI schema" available in their website)**

# And nowadays...

from

Q. Ning, H. He, C. Fan, D. Roth: Partial Or Complete, That Is The Question, *Proceedings of NAACL-HLT* 2019

"Many machine learning tasks require structured outputs, and the goal is to assign values to a set of variables coherently. ... ... An important implication is that **once some variables are determined**, the values **taken by other variables are constrained**. ...**it is important to study partialness systematically**, **before we hastily assume that completeness should always be favored** in data collection ... ... **Many recent datasets** that were collected via crowdsourcing **are already partial**..."

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0, 1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0, 1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0,1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0,1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0,1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \dots, X_k, H_1, \dots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \rightarrow [0,1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \dots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \rightarrow [0, 1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0,1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0,1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# Conditional probability assessment in a "nutshell"

- $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$, where:
  - $\mathcal{V} = \{X_1, \ldots, X_k, H_1, \ldots, H_k\}$ is a finite set of propositional variables $X_i$'s,
    representing any potential event of interest and their associated scenarios $H_i$'s;
  - $\mathcal{U}|\mathcal{H}$ is a set of $n \leq k$ effective conditional events $X_i|H_i$'s taken into consideration;
  - $p : \mathcal{U}|\mathcal{H} \to [0, 1]^n$ is a vector which assigns a "potential" probability value $p_i$, $i = 1, \ldots, n$, to each conditional event $X_i|H_i$;
  - $\mathcal{C}$ is a finite set of **logical constraints** which lie among all the variables in $V$.
- probability values for the elements in $\mathcal{U}|\mathcal{H}$ are assessed on the base of data or expert evaluations;
- but logical constraints $\mathcal{C}$ can be written in terms of all the potential events in $\mathcal{V}$, permitting extension to larger domains without redefining the whole model.

# A focus on the logical constraints

- Constraints in $\mathcal{C}$ are usually given through, or can be easily translated into, logical notation.
- These constraints can be used to represent any kind of compound ("meta") event, for instance that an event is the conjunction of other two events, or denote the implications or incompatibilities among the elements of $\mathcal{V}$.
- To avoid a foolish solution (i.e. concentrated on $\bigwedge_{i=1}^{n} \neg H_i$), a further constraint must be added to $\mathcal{C}$.
- Without loss of generality, $\mathcal{C}$ can be expressed (explicitly or through automatized procedures) in conjunctive normal form (CNF), Hence $\mathcal{C}$ can be seen as the conjunction of $\{c_1, \ldots, c_m\}$, where each element $c_i$ of $\mathcal{C}$ is a disjunctive clause.

# A focus on the logical constraints

- Constraints in $\mathcal{C}$ are usually given through, or can be easily translated into, logical notation.
- These constraints can be used to represent any kind of compound ("meta") event, for instance that an event is the conjunction of other two events, or denote the implications or incompatibilities among the elements of $\mathcal{V}$.
- To avoid a foolish solution (i.e. concentrated on $\bigwedge_{i=1}^{n} \neg H_i$), a further constraint must be added to $\mathcal{C}$.
- Without loss of generality, $\mathcal{C}$ can be expressed (explicitly or through automatized procedures) in conjunctive normal form (CNF), Hence $\mathcal{C}$ can be seen as the conjunction of $\{c_1, \ldots, c_m\}$, where each element $c_i$ of $\mathcal{C}$ is a disjunctive clause.

# A focus on the logical constraints

- Constraints in $\mathcal{C}$ are usually given through, or can be easily translated into, logical notation.
- These constraints can be used to represent any kind of compound ("meta") event, for instance that an event is the conjunction of other two events, or denote the implications or incompatibilities among the elements of $\mathcal{V}$.
- To avoid a foolish solution (i.e. concentrated on $\bigwedge_{i=1}^{n} \neg H_i$), a further constraint must be added to $\mathcal{C}$.
- Without loss of generality, $\mathcal{C}$ can be expressed (explicitly or through automatized procedures) in conjunctive normal form (CNF), Hence $\mathcal{C}$ can be seen as the conjunction of $\{c_1, \ldots, c_m\}$, where each element $c_i$ of $\mathcal{C}$ is a disjunctive clause.

# A focus on the logical constraints

- Constraints in $\mathcal{C}$ are usually given through, or can be easily translated into, logical notation.
- These constraints can be used to represent any kind of compound ("meta") event, for instance that an event is the conjunction of other two events, or denote the implications or incompatibilities among the elements of $\mathcal{V}$.
- To avoid a foolish solution (i.e. concentrated on $\bigwedge_{i=1}^{n} \neg H_i$), a further constraint must be added to $\mathcal{C}$.
- Without loss of generality, $\mathcal{C}$ can be expressed (explicitly or through automatized procedures) in conjunctive normal form (CNF), Hence $\mathcal{C}$ can be seen as the conjunction of $\{c_1, \ldots, c_m\}$, where each element $c_i$ of $\mathcal{C}$ is a disjunctive clause.

# A focus on the logical constraints

- Constraints in $\mathcal{C}$ are usually given through, or can be easily translated into, logical notation.
- These constraints can be used to represent any kind of compound ("meta") event, for instance that an event is the conjunction of other two events, or denote the implications or incompatibilities among the elements of $\mathcal{V}$.
- To avoid a foolish solution (i.e. concentrated on $\bigwedge_{i=1}^{n} \neg H_i$), a further constraint must be added to $\mathcal{C}$.
- Without loss of generality, $\mathcal{C}$ can be expressed (explicitly or through automatized procedures) in conjunctive normal form (CNF), Hence $\mathcal{C}$ can be seen as the conjunction of $\{c_1, \ldots, c_m\}$, where each element $c_i$ of $\mathcal{C}$ is a disjunctive clause.

# A focus on the logical constraints

- Constraints in $\mathcal{C}$ are usually given through, or can be easily translated into, logical notation.
- These constraints can be used to represent any kind of compound ("meta") event, for instance that an event is the conjunction of other two events, or denote the implications or incompatibilities among the elements of $\mathcal{V}$.
- To avoid a foolish solution (i.e. concentrated on $\bigwedge_{i=1}^{n} \neg H_i$), a further constraint must be added to $\mathcal{C}$.
- Without loss of generality, $\mathcal{C}$ can be expressed (explicitly or through automatized procedures) in conjunctive normal form (CNF), Hence $\mathcal{C}$ can be seen as the conjunction of $\{c_1, \ldots, c_m\}$, where each element $c_i$ of $\mathcal{C}$ is a disjunctive clause.

# A focus on the logical constraints

- Constraints in $\mathcal{C}$ are usually given through, or can be easily translated into, logical notation.
- These constraints can be used to represent any kind of compound ("meta") event, for instance that an event is the conjunction of other two events, or denote the implications or incompatibilities among the elements of $\mathcal{V}$.
- To avoid a foolish solution (i.e. concentrated on $\bigwedge_{i=1}^{n} \neg H_i$), a further constraint must be added to $\mathcal{C}$.
- Without loss of generality, $\mathcal{C}$ can be expressed (explicitly or through automatized procedures) in conjunctive normal form (CNF), Hence $\mathcal{C}$ can be seen as the conjunction of $\{c_1, \ldots, c_m\}$, where each element $c_i$ of $\mathcal{C}$ is a disjunctive clause.

# Coherence "in a nutshell"

- Coherence of the assessments equates consistency with a full conditional probability distribution;
- Coherence is noting more than specific constraints satisfactions on different "unexpectedeness" scenarios (the so called "zero layers");
- When a probability assessment $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$ is not coherent, then it is possible to "correct" it in order to obtain a coherent $\pi'$ which is as close as possible to $\pi$, according to some distance or pseudo-distance function between probability assessments;
- Any coherent assessment can be extended, through constraints propagation, to the probability of a conditional event $X|H$ being a new inference target. Extension, in general, leads to closed intervals $[lb(X|H), ub(X|H)]$, with extremes induced by the "worst" and by the "best" cases.

# Coherence "in a nutshell"

- Coherence of the assessments equates consistency with a full conditional probability distribution;
- Coherence is noting more than specific constraints satisfactions on different "unexpectedeness" scenarios (the so called "zero layers");
- When a probability assessment $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$ is not coherent, then it is possible to "correct" it in order to obtain a coherent $\pi'$ which is as close as possible to $\pi$, according to some distance or pseudo-distance function between probability assessments;
- Any coherent assessment can be extended, through constraints propagation, to the probability of a conditional event $X|H$ being a new inference target. Extension, in general, leads to closed intervals $[lb(X|H), ub(X|H)]$, with extremes induced by the "worst" and by the "best" cases.

# Coherence "in a nutshell"

- Coherence of the assessments equates consistency with a full conditional probability distribution;
- Coherence is noting more than specific constraints satisfactions on different "unexpectedeness" scenarios (the so called "zero layers");
- When a probability assessment $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$ is not coherent, then it is possible to "correct" it in order to obtain a coherent $\pi'$ which is as close as possible to $\pi$, according to some distance or pseudo-distance function between probability assessments;
- Any coherent assessment can be extended, through constraints propagation, to the probability of a conditional event $X|H$ being a new inference target. Extension, in general, leads to closed intervals $[lb(X|H), ub(X|H)]$, with extremes induced by the "worst" and by the "best" cases.

# Coherence "in a nutshell"

- Coherence of the assessments equates consistency with a full conditional probability distribution;
- Coherence is noting more than specific constraints satisfactions on different "unexpectedeness" scenarios (the so called "zero layers");
- When a probability assessment $\pi = (\mathcal{V}, \mathcal{U}|\mathcal{H}, p, \mathcal{C})$ is not coherent, then it is possible to "correct" it in order to obtain a coherent $\pi'$ which is as close as possible to $\pi$, according to some distance or pseudo-distance function between probability assessments;
- Any coherent assessment can be extended, through constraints propagation, to the probability of a conditional event $X|H$ being a new inference target. Extension, in general, leads to closed intervals $[lb(X|H), ub(X|H)]$, with extremes induced by the "worst" and by the "best" cases.

# Coherence approach peculiarities...

- **Logical (structural) constraints** among events well encompass **compulsory aggregations**, **implications**, **dependencies** among variables;
- **Direct assessments** on conditional events seen as "**primitives**" logical entities;
- **Partial assessments** by focusing only on the information at hand, leaving room for whatever it is unknown;
- Possibility of conditioning to events with (assessed or implied) **zero probability**, representing **"unexpected scenarios"**, by building on a representation of conditional probability as a sequence of unconditional probability distributions.

# Coherence approach peculiarities...

- **Logical (structural) constraints** among events well encompass **compulsory aggregations**, **implications**, **dependencies** among variables;
- **Direct assessments** on conditional events seen as "**primitives**" logical entities;
- **Partial assessments** by focusing only on the information at hand, leaving room for whatever it is unknown;
- Possibility of conditioning to events with (assessed or implied) **zero probability**, representing **"unexpected scenarios"**, by building on a representation of conditional probability as a sequence of unconditional probability distributions.

# Coherence approach peculiarities...

- **Logical (structural) constraints** among events well encompass **compulsory aggregations**, **implications**, **dependencies** among variables;
- **Direct assessments** on conditional events seen as "**primitives**" logical entities;
- **Partial assessments** by focusing only on the information at hand, leaving room for whatever it is unknown;
- Possibility of conditioning to events with (assessed or implied) **zero probability**, representing **"unexpected scenarios"**, by building on a representation of conditional probability as a sequence of unconditional probability distributions.

# Coherence approach peculiarities...

- **Logical (structural) constraints** among events well encompass **compulsory aggregations**, **implications**, **dependencies** among variables;
- **Direct assessments** on conditional events seen as "**primitives**" logical entities;
- **Partial assessments** by focusing only on the information at hand, leaving room for whatever it is unknown;
- Possibility of conditioning to events with (assessed or implied) **zero probability**, representing **"unexpected scenarios"**, by building on a representation of conditional probability as a sequence of unconditional probability distributions.

# Coherence approach peculiarities...

- **Logical (structural) constraints** among events well encompass **compulsory aggregations**, **implications**, **dependencies** among variables;
- **Direct assessments** on conditional events seen as "**primitives**" logical entities;
- **Partial assessments** by focusing only on the information at hand, leaving room for whatever it is unknown;
- Possibility of conditioning to events with (assessed or implied) **zero probability**, representing **"unexpected scenarios"**, by building on a representation of conditional probability as a sequence of unconditional probability distributions.

# Coherence approach peculiarities...

- **Logical (structural) constraints** among events well encompass **compulsory aggregations**, **implications**, **dependencies** among variables;
- **Direct assessments** on conditional events seen as "**primitives**" logical entities;
- **Partial assessments** by focusing only on the information at hand, leaving room for whatever it is unknown;
- Possibility of conditioning to events with (assessed or implied) **zero probability**, representing **"unexpected scenarios"**, by building on a representation of conditional probability as a sequence of unconditional probability distributions.

# Coherence approach peculiarities...

- **Logical (structural) constraints** among events well encompass **compulsory aggregations**, **implications**, **dependencies** among variables;
- **Direct assessments** on conditional events seen as "**primitives**" logical entities;
- **Partial assessments** by focusing only on the information at hand, leaving room for whatever it is unknown;
- Possibility of conditioning to events with (assessed or implied) **zero probability**, representing **"unexpected scenarios"**, by building on a representation of conditional probability as a sequence of unconditional probability distributions.

# Coherence approach peculiarities...

- With **partial conditional probability assessments** we can manage "**open worlds**" and "**on-demand**" association rules: (since the huge volume and heterogeneity of the data and the need to perform not previously designed analyses (e.g. in data-lake architectures), instead of a full and complex model, several smaller and specifically tailored tools are increasingly needed);
- Through **coherent extensions** based on partial conditional probabilistic models (generalizing de Finetti's fundamental theorem of previsions) we are be able to **infer conclusions** for **new "targets"**, and possibly to take into consideration **different scenarios**(possibly "**unexpected**") where to make analysis;

# Coherence approach peculiarities...

- With **partial conditional probability assessments** we can manage "**open worlds**" and "**on-demand**" association rules: (since the huge volume and heterogeneity of the data and the need to perform not previously designed analyses (e.g. in data-lake architectures), instead of a full and complex model, several smaller and specifically tailored tools are increasingly needed);
- Through **coherent extensions** based on partial conditional probabilistic models (generalizing de Finetti's fundamental theorem of previsions) we are be able to **infer conclusions** for **new "targets"**, and possibly to take into consideration **different scenarios**(possibly "**unexpected**") where to make analysis;

# Coherence approach peculiarities...

- With **partial conditional probability assessments** we can manage "**open worlds**" and "**on-demand**" association rules: (since the huge volume and heterogeneity of the data and the need to perform not previously designed analyses (e.g. in data-lake architectures), instead of a full and complex model, several smaller and specifically tailored tools are increasingly needed);
- Through **coherent extensions** based on partial conditional probabilistic models (generalizing de Finetti's fundamental theorem of previsions) we are be able to **infer conclusions** for **new "targets"**, and possibly to take into consideration **different scenarios**(possibly "**unexpected**") where to make analysis;

# Coherence approach peculiarities...

- With **partial conditional probability assessments** we can manage "**open worlds**" and "**on-demand**" association rules: (since the huge volume and heterogeneity of the data and the need to perform not previously designed analyses (e.g. in data-lake architectures), instead of a full and complex model, several smaller and specifically tailored tools are increasingly needed);
- Through **coherent extensions** based on partial conditional probabilistic models (generalizing de Finetti's fundamental theorem of previsions) we are be able to **infer conclusions** for **new "targets"**, and possibly to take into consideration **different scenarios**(possibly "**unexpected**") where to make analysis;

# Coherence approach peculiarities...

- With **partial conditional probability assessments** we can manage "**open worlds**" and "**on-demand**" association rules: (since the huge volume and heterogeneity of the data and the need to perform not previously designed analyses (e.g. in data-lake architectures), instead of a full and complex model, several smaller and specifically tailored tools are increasingly needed);
- Through **coherent extensions** based on partial conditional probabilistic models (generalizing de Finetti's fundamental theorem of previsions) we are be able to **infer conclusions** for **new "targets"**, and possibly to take into consideration **different scenarios**(possibly "**unexpected**") where to make analysis;

# Coherence approach peculiarities...

- With **partial conditional probability assessments** we can manage "**open worlds**" and "**on-demand**" association rules: (since the huge volume and heterogeneity of the data and the need to perform not previously designed analyses (e.g. in data-lake architectures), instead of a full and complex model, several smaller and specifically tailored tools are increasingly needed);
- Through **coherent extensions** based on partial conditional probabilistic models (generalizing de Finetti's fundamental theorem of previsions) we are be able to **infer conclusions** for **new "targets"**, and possibly to take into consideration **different scenarios**(possibly "**unexpected**") where to make analysis;

# Prototypical examples

By borrowing examples from some, more or less recent, contributions we can emphasize some of such peculiarities...

# On the crucial role of structural constraints

Example taken from

A. Brozzi, A. Capotorti, B. Vantaggi: Incoherence correction strategies in statistical matching, *Int. Journ. of Approximate Reasoning* 2012

The data are a subset of 2313 employees (people at least 15 years old) extracted from **two distinct** databases: the pilot survey of the Italian Population and Household Census in the year 2000.
Three categorical variables have been analyzed: **Age**, **Educational Level** and **Professional Status**.

# On the crucial role of structural constraints

In database A, containing 1148 units, the variables **Age** and **Professional Status** are observed, while database B, consisting of 1165 observations, the variables **Age** and **Educational Level** are considered.

Variables are grouped in homogeneous response categories as follows:

$H_1$ = 15–17 years old, $H_2$ = 18–22 years old, $H_3$ = 23–64 years old, $H_4$ = more than 65;

$X_1$ = None or compulsory school, $X_2$ = Vocational school, $X_3$=High school, $X_4$= Master;

$Y_1$=Manager, $Y_2$=Office worker, $Y_3$=Basic Worker.

# On the crucial role of structural constraints

In database A, containing 1148 units, the variables **Age** and **Professional Status** are observed, while database B, consisting of 1165 observations, the variables **Age** and **Educational Level** are considered.

Variables are grouped in homogeneous response categories as follows:
$H_1$ = 15–17 years old, $H_2$ = 18–22 years old, $H_3$ = 23–64 years old, $H_4$ = more than 65;
$X_1$ = None or compulsory school, $X_2$ = Vocational school, $X_3$=High school, $X_4$= Master;
$Y_1$=Manager, $Y_2$=Office worker, $Y_3$=Basic Worker.

# On the crucial role of structural constraints

Logical constraints between the variables (at least in Italy in the 2000) are:

- $H_1 \wedge (X_3 \vee X_4) = \emptyset$ (underages cannot have HS diplomas or Master degrees)
- $H_2 \wedge X_4 = \emptyset$ (under 22 cannot have Master degrees)
- $H_1 \wedge (Y_1 \vee Y_2) = \emptyset$ (underages cannot have office or manager positions)
- $H_2 \wedge Y_1 = \emptyset$ (under 22 cannot have manager positions)
- $Y_1 \wedge (X_1 \vee X_2) = \emptyset$ (a manager should have at least a HS diploma)
- $Y_2 \wedge X_1 = \emptyset$ (an office worker should have at least a vocational school diploma).

# On the crucial role of structural constraints

Observed frequencies give rise to the following (ML) estimates:

for the variable Age:

- $p(H_1) = \frac{15}{2313}, p(H_2) = \frac{55}{2313}, p(H_3) = \frac{2219}{2313}, p(X_4) = \frac{24}{2313}$;

for the Professional Status given the Age:

- $p(Y_2|H_2) = \frac{5}{22}, p(Y_3|H_2) = \frac{17}{22}$,
- $p(Y_1|H_3) = \frac{179}{1108}, p(Y_2|H_3) = \frac{443}{1108}, p(Y_3|H_3) = \frac{486}{1108}$,
- $p(Y_1|H_4) = \frac{2}{3}, p(Y_2|H_4) = \frac{1}{9}, p(Y_3|H_4) = \frac{2}{9}$;

for the Educational level given the Age:

- $p(X_1|H_1) = 1, p(X_2|H_1) = 0$,
- $p(X_1|H_2) = \frac{14}{33}, p(X_2|H_2) = \frac{6}{33}, p(X_3|H_2) = \frac{13}{33}$,
- $p(X_1|H_3) = \frac{387}{1111}, p(X_2|H_3) = \frac{102}{1111}, p(X_3|H_3) = \frac{464}{1111}, p(X_4|H_3) = \frac{158}{1111}$,
- $p(X_1|H_4) = \frac{2}{3}, p(X_2|H_4) = 0, p(X_3|H_4) = \frac{1}{5}, p(X_4|H_4) = \frac{2}{15}$.

# On the crucial role of structural constraints

Observed frequencies give rise to the following (ML) estimates:
for the variable Age:

- $p(H_1) = \frac{15}{2313}, p(H_2) = \frac{55}{2313}, p(H_3) = \frac{2219}{2313}, p(X_4) = \frac{24}{2313};$

for the Professional Status given the Age:

- $p(Y_2|H_2) = \frac{5}{22}, p(Y_3|H_2) = \frac{17}{22},$
- $p(Y_1|H_3) = \frac{179}{1108}, p(Y_2|H_3) = \frac{443}{1108}, p(Y_3|H_3) = \frac{486}{1108},$
- $p(Y_1|H_4) = \frac{2}{3}, p(Y_2|H_4) = \frac{1}{9}, p(Y_3|H_4) = \frac{2}{9};$

for the Educational level given the Age:

- $p(X_1|H_1) = 1, p(X_2|H_1) = 0,$
- $p(X_1|H_2) = \frac{14}{33}, p(X_2|H_2) = \frac{6}{33}, p(X_3|H_2) = \frac{13}{33},$
- $p(X_1|H_3) = \frac{387}{1111}, p(X_2|H_3) = \frac{102}{1111}, p(X_3|H_3) = \frac{464}{1111}, p(X_4|H_3) = \frac{158}{1111},$
- $p(X_1|H_4) = \frac{2}{3}, p(X_2|H_4) = 0, p(X_3|H_4) = \frac{1}{5}, p(X_4|H_4) = \frac{2}{15}.$

# On the crucial role of structural constraints

Observed frequencies give rise to the following (ML) estimates:
for the variable Age:

- $p(H_1) = \frac{15}{2313}, p(H_2) = \frac{55}{2313}, p(H_3) = \frac{2219}{2313}, p(X_4) = \frac{24}{2313}$;

for the Professional Status given the Age:

- $p(Y_2|H_2) = \frac{5}{22}, p(Y_3|H_2) = \frac{17}{22}$,
- $p(Y_1|H_3) = \frac{179}{1108}, p(Y_2|H_3) = \frac{443}{1108}, p(Y_3|H_3) = \frac{486}{1108}$,
- $p(Y_1|H_4) = \frac{2}{3}, p(Y_2|H_4) = \frac{1}{9}, p(Y_3|H_4) = \frac{2}{9}$;

for the Educational level given the Age:

- $p(X_1|H_1) = 1, p(X_2|H_1) = 0$,
- $p(X_1|H_2) = \frac{14}{33}, p(X_2|H_2) = \frac{6}{33}, p(X_3|H_2) = \frac{13}{33}$,
- $p(X_1|H_3) = \frac{387}{1111}, p(X_2|H_3) = \frac{102}{1111}, p(X_3|H_3) = \frac{464}{1111}, p(X_4|H_3) = \frac{158}{1111}$,
- $p(X_1|H_4) = \frac{2}{3}, p(X_2|H_4) = 0, p(X_3|H_4) = \frac{1}{5}, p(X_4|H_4) = \frac{2}{15}$.

# On the crucial role of structural constraints

Observed frequencies give rise to the following (ML) estimates:
for the variable Age:

- $p(H_1) = \frac{15}{2313}, p(H_2) = \frac{55}{2313}, p(H_3) = \frac{2219}{2313}, p(X_4) = \frac{24}{2313};$

for the Professional Status given the Age:

- $p(Y_2|H_2) = \frac{5}{22}, p(Y_3|H_2) = \frac{17}{22},$
- $p(Y_1|H_3) = \frac{179}{1108}, p(Y_2|H_3) = \frac{443}{1108}, p(Y_3|H_3) = \frac{486}{1108},$
- $p(Y_1|H_4) = \frac{2}{3}, p(Y_2|H_4) = \frac{1}{9}, p(Y_3|H_4) = \frac{2}{9};$

for the Educational level given the Age:

- $p(X_1|H_1) = 1, p(X_2|H_1) = 0,$
- $p(X_1|H_2) = \frac{14}{33}, p(X_2|H_2) = \frac{6}{33}, p(X_3|H_2) = \frac{13}{33},$
- $p(X_1|H_3) = \frac{387}{1111}, p(X_2|H_3) = \frac{102}{1111}, p(X_3|H_3) = \frac{464}{1111}, p(X_4|H_3) = \frac{158}{1111},$
- $p(X_1|H_4) = \frac{2}{3}, p(X_2|H_4) = 0, p(X_3|H_4) = \frac{1}{5}, p(X_4|H_4) = \frac{2}{15}.$

# On the crucial role of structural constraints

Even $p(H_i)$, $p(Y_j|H_i)$ and $p(X_l|H_i)$ being separately coherent, **the whole assessment is not coherent**.

Incoherence is localized in $p(\cdot|H_4)$ since **from logical constraints** it follows that $X_1 \wedge Y_1 = \emptyset$ and $X_1 \subseteq Y_3$, while we have $p(X_1|H_4) + p(Y_1|H_4) > 1$ and $p(X_1|H_4) > p(Y_3|H_4)$.

Different coherent corrections are possible, e.g. through a suitable dicrepancy measure we have:

$p(Y_1|H_4) = .498, p(Y_2|X_4) = .094, p(Y_3|H_4) = .408,$
$p(X_1|H_4) = .408, p(X_2|H_4) = .094, p(X_3|H_4) = .294, p(X_4|H_4) = .204$

and the coherent bounds

$p(Y_3|X_4) = p(\text{Basic Worker}|\text{Master}) \in [0, .625]$ are determined.

# On the crucial role of structural constraints

Even $p(H_i)$, $p(Y_j|H_i)$ and $p(X_l|H_i)$ being separately coherent, **the whole assessment is not coherent**.

Incoherence is localized in $p(\cdot|H_4)$ since **from logical constraints** it follows that $X_1 \wedge Y_1 = \emptyset$ and $X_1 \subseteq Y_3$, while we have $p(X_1|H_4) + p(Y_1|H_4) > 1$ and $p(X_1|H_4) > p(Y_3|H_4)$.

Different coherent corrections are possible, e.g. through a suitable dicrepancy measure we have:

$p(Y_1|H_4) = .498, p(Y_2|X_4) = .094, p(Y_3|H_4) = .408,$
$p(X_1|H_4) = .408, p(X_2|H_4) = .094, p(X_3|H_4) = .294, p(X_4|H_4) = .204$

and the coherent bounds

$p(Y_3|X_4) = p(\text{Basic Worker}|\text{Master}) \in [0, .625]$ are determined.

# On the crucial role of structural constraints

Even $p(H_i)$, $p(Y_j|H_i)$ and $p(X_l|H_i)$ being separately coherent, **the whole assessment is not coherent**.

Incoherence is localized in $p(\cdot|H_4)$ since **from logical constraints** it follows that $X_1 \wedge Y_1 = \emptyset$ and $X_1 \subseteq Y_3$, while we have $p(X_1|H_4) + p(Y_1|H_4) > 1$ and $p(X_1|H_4) > p(Y_3|H_4)$.

Different coherent corrections are possible, e.g. through a suitable dicrepancy measure we have:

$p(Y_1|H_4) = .498, p(Y_2|X_4) = .094, p(Y_3|H_4) = .408,$
$p(X_1|H_4) = .408, p(X_2|H_4) = .094, p(X_3|H_4) = .294, p(X_4|H_4) = .204$

and the coherent bounds

$p(Y_3|X_4) = p(\text{Basic Worker}|\text{Master}) \in [0, .625]$ are determined.

# On the crucial role of structural constraints

Even $p(H_i)$, $p(Y_j|H_i)$ and $p(X_l|H_i)$ being separately coherent, **the whole assessment is not coherent**.

Incoherence is localized in $p(\cdot|H_4)$ since **from logical constraints** it follows that $X_1 \wedge Y_1 = \emptyset$ and $X_1 \subseteq Y_3$, while we have $p(X_1|H_4) + p(Y_1|H_4) > 1$ and $p(X_1|H_4) > p(Y_3|H_4)$.

Different coherent corrections are possible, e.g. through a suitable dicrepancy measure we have:

$p(Y_1|H_4) = .498, p(Y_2|X_4) = .094, p(Y_3|H_4) = .408,$
$p(X_1|H_4) = .408, p(X_2|H_4) = .094, p(X_3|H_4) = .294, p(X_4|H_4) = .204$

and the coherent bounds

$p(Y_3|X_4) = p(\text{Basic Worker}|\text{Master}) \in [0, .625]$ are determined.

# Benefits zero probability for rare events...

Example taken from

L. van der Gaag, A. Capotorti: Naive Bayesian Classifiers with Extreme Probability Features, Proc. of *Int. Conf. on Probabilistic Graphical Models, PGM* 2018

Naive Bayesian classifiers are proving **less suited for real-world applications** in which **quite rare features play an important role**, such as the **medical diagnosis of serious disease with rare (pathognomonic) findings**.

# Benefits zero probability for rare events...

When the classifier's parameter probabilities are learned from data, the estimate obtained for the **rare feature** at hand will most likely be zero, but, generally, some smoothing arbitrary method is employed to forestall the inclusion of zero probability parameters in the model... leading to various unwanted effects (e.g. inaccuracies in small parameter probabilities can strongly influence output probabilities of interest)

# Benefits zero probability for rare events...

When the classifier's parameter probabilities are learned from data, the estimate obtained for the **rare feature** at hand will most likely be zero, but, generally, some smoothing arbitrary method is employed to forestall the inclusion of zero probability parameters in the model... leading to various unwanted effects (e.g. inaccuracies in small parameter probabilities can strongly influence output probabilities of interest)

# Benefits zero probability for rare events...

When the classifier's parameter probabilities are learned from data, the estimate obtained for the **rare feature** at hand will most likely be zero, but, generally, some smoothing arbitrary method is employed to forestall the inclusion of zero probability parameters in the model... leading to various unwanted effects (e.g. inaccuracies in small parameter probabilities can strongly influence output probabilities of interest)

# Benefits zero probability for rare events...

Let $A$ be a binary the class variable;

$B$ and $C$ two binary feature variables, the value $b$ of $B$ representing the rare feature;

Logical constraints $(B = b) \subseteq (A = a)$ (the rare feature $(B = b)$ appears only in class $(A = a)$);

Probability classification threshold $\delta = .75$.

# Benefits zero probability for rare events...

The "smooth" assessment:

$p(A = a) = .8, \quad p(C = c|A = \neg a) = .2,$
$p(C = c|A = a) = .4, \quad \mathbf{p(B = b|A = a) = .001}$

implies

$p(A = a|(B = \neg b \wedge C = \neg c)) = .7498 < \delta \quad \Rightarrow \quad \text{Class} = \neg\mathbf{a}$

$p(A = a|C = \neg c) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$

Hence adding the feature $B = \neg b$ to the instance changes the output, even though such a feature is in essence uninformative. (**Note**: the same counterintuitive effect even reducing $p(B = b|A = a)$ by several orders of magnitude in size)

# Benefits zero probability for rare events...

The "smooth" assessment:

$$p(A = a) = .8, \quad p(C = c|A = \neg a) = .2,$$
$$p(C = c|A = a) = .4, \quad \mathbf{p(B = b|A = a) = .001}$$

implies

$$p(A = a|(B = \neg b \land C = \neg c)) = .7498 < \delta \quad \Rightarrow \quad \text{Class} = \neg \mathbf{a}$$

$$p(A = a|C = \neg c) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$$

Hence adding the feature $B = \neg b$ to the instance changes the output, even though such a feature is in essence uninformative. (**Note**: the same counterintuitive effect even reducing $p(B = b|A = a)$ by several orders of magnitude in size)

# Benefits zero probability for rare events...

The "smooth" assessment:

$$p(A = a) = .8, \quad p(C = c|A = \neg a) = .2,$$
$$p(C = c|A = a) = .4, \quad \mathbf{p(B = b|A = a) = .001}$$

implies

$$p(A = a|(B = \neg b \wedge C = \neg c)) = .7498 < \delta \quad \Rightarrow \quad \text{Class} = \neg\mathbf{a}$$

$$p(A = a|C = \neg c) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$$

Hence adding the feature $B = \neg b$ to the instance changes the output, even though such a feature is in essence uninformative. (**Note**: the same counterintuitive effect even reducing $p(B = b|A = a)$ by several orders of magnitude in size)

# Benefits zero probability for rare events...

The "smooth" assessment:

$$p(A = a) = .8, \quad p(C = c|A = \neg a) = .2,$$
$$p(C = c|A = a) = .4, \quad \mathbf{p(B = b|A = a) = .001}$$

implies

$$p(A = a|(B = \neg b \wedge C = \neg c)) = .7498 < \delta \quad \Rightarrow \quad \text{Class} = \neg\mathbf{a}$$

$$p(A = a|C = \neg c) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$$

Hence adding the feature $B = \neg b$ to the instance changes the output, even though such a feature is in essence uninformative. (**Note**: the same counterintuitive effect even reducing $p(B = b|A = a)$ by several orders of magnitude in size)

# Benefits zero probability for rare events...

The "natural" assessment:

$p(A = a) = .8, \quad p(C = c|A = \neg a) = .2,$
$p(C = c|A = a) = .4, \quad \mathbf{p(B = b|A = a) = 0} \quad (p(B = b) = 0)$

implies

$p(A = a|(B = \neg b \wedge C = \neg c)) = .75 = \delta \quad \Rightarrow \quad$ Class $= \mathbf{a}$

$p(A = a|C = \neg c) = .75 = \delta \quad \Rightarrow \quad$ Class $= \mathbf{a}$

Hence adding the feature $B = \neg b$ to the instance doesn't change the output, confirming the usualness of such a feature
(**Note**: although we have $p(B = b|A = a) = p(B = b|A = \neg a)$, in the coherent setting $A$ and $B$ aren't - correctly - independent...)

# Benefits zero probability for rare events...

The "natural" assessment:

$$p(A=a)=.8, \quad p(C=c|A=\neg a)=.2,$$
$$p(C=c|A=a)=.4, \quad \mathbf{p(B=b|A=a)=0} \quad (p(B=b)=0)$$

implies

$$p(A=a|(B=\neg b \wedge C=\neg c))=.75=\delta \quad \Rightarrow \quad \text{Class} \ =\mathbf{a}$$

$$p(A=a|C=\neg c)=.75=\delta \quad \Rightarrow \quad \text{Class} \ =\mathbf{a}$$

Hence adding the feature $B=\neg b$ to the instance doesn't change the output, confirming the usualness of such a feature
(**Note**: although we have $p(B=b|A=a)=p(B=b|A=\neg a)$, in the coherent setting $A$ and $B$ aren't - correctly - independent...)

# Benefits zero probability for rare events...

The "natural" assessment:

$$p(A = a) = .8, \quad p(C = c|A = \neg a) = .2,$$
$$p(C = c|A = a) = .4, \quad \mathbf{p(B = b|A = a) = 0} \quad (p(B = b) = 0)$$

implies

$$p(A = a|(B = \neg b \wedge C = \neg c)) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$$

$$p(A = a|C = \neg c) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$$

Hence adding the feature $B = \neg b$ to the instance doesn't change the output, confirming the usualness of such a feature
(**Note**: although we have $p(B = b|A = a) = p(B = b|A = \neg a)$, in the coherent setting $A$ and $B$ aren't - correctly - independent...)

# Benefits zero probability for rare events...

The "natural" assessment:

$$p(A = a) = .8, \quad p(C = c|A = \neg a) = .2,$$
$$p(C = c|A = a) = .4, \quad \mathbf{p(B = b|A = a) = 0} \quad (p(B = b) = 0)$$

implies

$$p(A = a|(B = \neg b \wedge C = \neg c)) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$$

$$p(A = a|C = \neg c) = .75 = \delta \quad \Rightarrow \quad \text{Class} = \mathbf{a}$$

Hence adding the feature $B = \neg b$ to the instance doesn't change the output, confirming the usualness of such a feature (**Note**: although we have $p(B = b|A = a) = p(B = b|A = \neg a)$, in the coherent setting $A$ and $B$ aren't - correctly - independent...)

# What we can really infer with (very) partial information ...

Two examples taken from

N. Pfeifer, A. Capotorti: What society can and cannot learn from coherence: theoretical and practical considerations, *Adapting Human Thinking and Moral Reasoning in Contemporary Society (§8)* 2020

Honest and right description of the information at hand, *even being just on a "bunch" of events*, can lead to meaningful an sometimes unexpected conclusions...

# Inference with partial information: a way to reveal wrong conclusions

Let us see a prototypical example of how coherence can help in discovering one of the first known "fake new": autism induced by anti measles ("morbillo") vaccination.

The fake new started with a wrong (and actually dishonest) study where partial data about 12 vaccinated patients were reported. Among them, 9 were diagnosed to have developed some kind of autism.

Let us see how a right and sound formalization could lead instead to reasonable conclusion and reveal the wrong ones.

# Inference with partial information: a way to reveal wrong conclusions

Let us see a prototypical example of how coherence can help in discovering one of the first known "fake new": autism induced by anti measles ("morbillo") vaccination.

The fake new started with a wrong (and actually dishonest) study where partial data about 12 vaccinated patients were reported. Among them, 9 were diagnosed to have developed some kind of autism.

Let us see how a right and sound formalization could lead instead to reasonable conclusion and reveal the wrong ones.

# Inference with partial information: a way to reveal wrong conclusions

Let us see a prototypical example of how coherence can help in discovering one of the first known “fake new”: autism induced by anti measles (“morbillo”) vaccination.

The fake new started with a wrong (and actually dishonest) study where partial data about 12 vaccinated patients were reported. Among them, 9 were diagnosed to have developed some kind of autism.

Let us see how a right and sound formalization could lead instead to reasonable conclusion and reveal the wrong ones.

# Inference with partial information: a way to reveal wrong conclusions

Let us see a prototypical example of how coherence can help in discovering one of the first known "fake new": autism induced by anti measles ("morbillo") vaccination.

The fake new started with a wrong (and actually dishonest) study where partial data about 12 vaccinated patients were reported. Among them, 9 were diagnosed to have developed some kind of autism.
Let us see how a right and sound formalization could lead instead to reasonable conclusion and reveal the wrong ones.

# NO-VAX FAKE NEW

Events on the ground:

- V (Vaccinated) = person recently vaccinated against MMR
- A (Autistic) = person suffering real form of autism
- DA (Diagnosed Autistic) = person diagnosed as autistic

Logical relations (constraints): $A \subseteq DA$ (sure diagnosis in case of real presence of autistic disorder)
Partial probabilistic evaluations:

- $p(DA|V) = 9/12$ (taken from the mentioned initial study);
- $p(A) = 1.47\% = 1/68$ (autism prevalence on the reference population - official data)
- $91\% \leq p(V) \leq 92\%$ (realistic prior estimate on vaccination)

# NO-VAX FAKE NEW

Events on the ground:

- V (Vaccinated) = person recently vaccinated against MMR
- A (Autistic) = person suffering real form of autism
- DA (Diagnosed Autistic) = person diagnosed as autistic

Logical relations (constraints): $A \subseteq DA$ (sure diagnosis in case of real presence of autistic disorder)

Partial probabilistic evaluations:

- $p(DA|V) = 9/12$ (taken from the mentioned initial study);
- $p(A) = 1.47\% = 1/68$ (autism prevalence on the reference population - official data)
- $91\% \leq p(V) \leq 92\%$ (realistic prior estimate on vaccination)

# NO-VAX FAKE NEW

Events on the ground:

- V (Vaccinated) = person recently vaccinated against MMR
- A (Autistic) = person suffering real form of autism
- DA (Diagnosed Autistic) = person diagnosed as autistic

Logical relations (constraints): $A \subseteq DA$ (sure diagnosis in case of real presence of autistic disorder)

Partial probabilistic evaluations:

- $p(DA|V) = 9/12$ (taken from the mentioned initial study);
- $p(A) = 1.47\% = 1/68$ (autism prevalence on the reference population - official data)
- $91\% \leq p(V) \leq 92\%$ (realistic prior estimate on vaccination)

# NO-VAX FAKE NEW

Events on the ground:

- V (Vaccinated) = person recently vaccinated against MMR
- A (Autistic) = person suffering real form of autism
- DA (Diagnosed Autistic) = person diagnosed as autistic

Logical relations (constraints): $A \subseteq DA$ (sure diagnosis in case of real presence of autistic disorder)

Partial probabilistic evaluations:

- $p(DA|V) = 9/12$ (taken from the mentioned initial study);
- $p(A) = 1.47\% = 1/68$ (autism prevalence on the reference population - official data)
- $91\% \leq p(V) \leq 92\%$ (realistic prior estimate on vaccination)

# NO-VAX FAKE NEW

Events on the ground:

- V (Vaccinated) = person recently vaccinated against MMR
- A (Autistic) = person suffering real form of autism
- DA (Diagnosed Autistic) = person diagnosed as autistic

Logical relations (constraints): $A \subseteq DA$ (sure diagnosis in case of real presence of autistic disorder)
Partial probabilistic evaluations:

- $p(DA|V) = 9/12$ (taken from the mentioned initial study);
- $p(A) = 1.47\% = 1/68$ (autism prevalence on the reference population - official data)
- $91\% \leq p(V) \leq 92\%$ (realistic prior estimate on vaccination)

Coherent extensions:

- $0 \leq p(A|V) \leq 1.62\%$ (slight risk increasing and only for the more "extreme" models, i.e. "worst scenarios");
- $0 \leq p(A|\neg V) \leq 18.382\%$ (unreasonable possible increase of the risk also for those who will not vaccinate, induced by worst scenarios that allow positive correlation between $\neg V$ and $A$);
- $66.779\% \leq p(\neg A \wedge DA) \leq 75.779\%$ (high risk of wrong diagnosis, as discovered afterward when 2/3 of the results were proved to be frauds).

# NO-VAX FAKE NEW

Coherent extensions:

- $0 \leq p(A|V) \leq 1.62\%$ (slight risk increasing and only for the more "extreme" models, i.e. "worst scenarios");
- $0 \leq p(A|\neg V) \leq 18.382\%$ (unreasonable possible increase of the risk also for those who will not vaccinate, induced by worst scenarios that allow positive correlation between $\neg V$ and $A$);
- $66.779\% \leq p(\neg A \wedge DA) \leq 75.779\%$ (high risk of wrong diagnosis, as discovered afterward when 2/3 of the results were proved to be frauds).

# NO-VAX FAKE NEW

Coherent extensions:

- $0 \leq p(A|V) \leq 1.62\%$ (slight risk increasing and only for the more "extreme" models, i.e. "worst scenarios");
- $0 \leq p(A|\neg V) \leq 18.382\%$ (unreasonable possible increase of the risk also for those who will not vaccinate, induced by worst scenarios that allow positive correlation between $\neg V$ and $A$);
- $66.779\% \leq p(\neg A \wedge DA) \leq 75.779\%$ (high risk of wrong diagnosis, as discovered afterward when 2/3 of the results were proved to be frauds).

# NO-VAX FAKE NEW

Coherent extensions:

- $0 \leq p(A|V) \leq 1.62\%$ (slight risk increasing and only for the more "extreme" models, i.e. "worst scenarios");
- $0 \leq p(A|\neg V) \leq 18.382\%$ (unreasonable possible increase of the risk also for those who will not vaccinate, induced by worst scenarios that allow positive correlation between $\neg V$ and $A$);
- $66.779\% \leq p(\neg A \wedge DA) \leq 75.779\%$ (high risk of wrong diagnosis, as discovered afterward when 2/3 of the results were proved to be frauds).

# Inference with partial information: a warning on "sure" proofs

Coherence can also help in avoiding hasty conclusion when "incontestable" proofs are given in a court.

We have all heard about "sure reliability" of DNA proofs in legal trials.

Let us see what we can actually infer on the base of only the laboratory results and preliminary investigations...

# Inference with partial information: a warning on "sure" proofs

Coherence can also help in avoiding hasty conclusion when "incontestable" proofs are given in a court.
We have all heard about "sure reliability" of DNA proofs in legal trials.
Let us see what we can actually infer on the base of only the laboratory results and preliminary investigations...

# Inference with partial information: a warning on "sure" proofs

Coherence can also help in avoiding hasty conclusion when "incontestable" proofs are given in a court.
We have all heard about "sure reliability" of DNA proofs in legal trials.
Let us see what we can actually infer on the base of only the laboratory results and preliminary investigations...

# DNA proof

Events on the ground:

- Exclusion, Inclusion, InConclusive = possible results of the DNA test performed on a crime suspect
- Suspect = person under investigation for a crime
- Guilty = real author of the crime

Logical relations:

Exclusion, Inclusion, InConclusive form a "partition" (exhaustive and incompatible alternatives)

Exclusion and Guilty are incompatible (full specificity of the DNA test)

Partial probability assessments:

- p(Guilty|Exclusion) = 0 (surely innocent whenever the test is "negative");
- p(Guilty|InConclusive) = .5 (maximum uncertainty whenever the test result is unsound)
- p(Guilty|Suspect) = 95% (reasonable high accuracy of the investigation = high confidence into detectives' skill)

# DNA proof

Events on the ground:

- Exclusion, Inclusion, InConclusive = possible results of the DNA test performed on a crime suspect
- Suspect = person under investigation for a crime
- Guilty = real author of the crime

Logical relations:

Exclusion, Inclusion, InConclusive form a "partition" (exhaustive and incompatible alternatives)

Exclusion and Guilty are incompatible (full specificity of the DNA test)

Partial probability assessments:

- p(Guilty|Exclusion) = 0 (surely innocent whenever the test is "negative");
- p(Guilty|InConclusive) = .5 (maximum uncertainty whenever the test result is unsound)
- p(Guilty|Suspect) = 95% (reasonable high accuracy of the investigation = high confidence into detectives' skill)

# DNA proof

Events on the ground:

- Exclusion, Inclusion, InConclusive = possible results of the DNA test performed on a crime suspect
- Suspect = person under investigation for a crime
- Guilty = real author of the crime

Logical relations:

Exclusion, Inclusion, InConclusive form a "partition" (exhaustive and incompatible alternatives)

Exclusion and Guilty are incompatible (full specificity of the DNA test)

Partial probability assessments:

- p(Guilty|Exclusion) = 0 (surely innocent whenever the test is "negative");
- p(Guilty|InConclusive) = .5 (maximum uncertainty whenever the test result is unsound)
- p(Guilty|Suspect) = 95% (reasonable high accuracy of the investigation = high confidence into detectives' skill)

# DNA PROOF

Events on the ground:

- Exclusion, Inclusion, InConclusive = possible results of the DNA test performed on a crime suspect
- Suspect = person under investigation for a crime
- Guilty = real author of the crime

Logical relations:

Exclusion, Inclusion, InConclusive form a “partition” (exhaustive and incompatible alternatives)

Exclusion and Guilty are incompatible (full specificity of the DNA test)

Partial probability assessments:

- p(Guilty|Exclusion) = 0 (surely innocent whenever the test is “negative”);
- p(Guilty|InConclusive) = .5 (maximum uncertainty whenever the test result is unsound)
- p(Guilty|Suspect) = 95% (reasonable high accuracy of the investigation = high confidence into detectives’ skill)

# DNA proof

Events on the ground:

- Exclusion, Inclusion, InConclusive = possible results of the DNA test performed on a crime suspect
- Suspect = person under investigation for a crime
- Guilty = real author of the crime

Logical relations:

Exclusion, Inclusion, InConclusive form a “partition” (exhaustive and incompatible alternatives)

Exclusion and Guilty are incompatible (full specificity of the DNA test)

Partial probability assessments:

- p(Guilty|Exclusion) = 0 (surely innocent whenever the test is “negative”);
- p(Guilty|InConclusive) = .5 (maximum uncertainty whenever the test result is unsound)
- p(Guilty|Suspect) = 95% (reasonable high accuracy of the investigation = high confidence into detectives’ skill)

# DNA proof

Events on the ground:

- Exclusion, Inclusion, InConclusive = possible results of the DNA test performed on a crime suspect
- Suspect = person under investigation for a crime
- Guilty = real author of the crime

Logical relations:

Exclusion, Inclusion, InConclusive form a “partition” (exhaustive and incompatible alternatives)

Exclusion and Guilty are incompatible (full specificity of the DNA test)

Partial probability assessments:

- p(Guilty|Exclusion) = 0 (surely innocent whenever the test is “negative”);
- p(Guilty|InConclusive) = .5 (maximum uncertainty whenever the test result is unsound)
- p(Guilty|Suspect) = 95% (reasonable high accuracy of the investigation = high confidence into detectives’ skill)

# DNA proof

Events on the ground:

- Exclusion, Inclusion, InConclusive = possible results of the DNA test performed on a crime suspect
- Suspect = person under investigation for a crime
- Guilty = real author of the crime

Logical relations:

Exclusion, Inclusion, InConclusive form a "partition" (exhaustive and incompatible alternatives)

Exclusion and Guilty are incompatible (full specificity of the DNA test)

Partial probability assessments:

- p(Guilty|Exclusion) = 0 (surely innocent whenever the test is "negative");
- p(Guilty|InConclusive) = .5 (maximum uncertainty whenever the test result is unsound)
- p(Guilty|Suspect) = 95% (reasonable high accuracy of the investigation = high confidence into detectives' skill)

# DNA proof

Coherent extensions:

- $0 \leq p(\text{Guilty}|\text{Suspect and Inclusion}) \leq 100\%$

  (maximum vagueness for inference based only on laboratory results and preliminary investigations);

- $0 \leq p(\neg\ \text{Guilty} \wedge \text{Suspect} \wedge \text{Inclusion}) \leq 5\%$

  (rare, but possible, "false positive" scenario where other elements should be investigated, like, e.g., the possibility of contamination during the DNA profiling procedure).

(**Note**: such conclusions are insensitive with respect to the sensibility $p(\text{Inclusion}|\text{Guilty})$ of the DNA test...)

# DNA proof

Coherent extensions:

- $0 \leq$ p(Guilty|Suspect and Inclusion) $\leq 100\%$

  (maximum vagueness for inference based only on laboratory results and preliminary investigations);

- $0 \leq$ p($\neg$ Guilty $\wedge$ Suspect $\wedge$ Inclusion) $\leq 5\%$

  (rare, but possible, "false positive" scenario where other elements should be investigated, like, e.g., the possibility of contamination during the DNA profiling procedure).

(**Note**: such conclusions are insensitive with respect to the sensibility p(Inclusion|Guilty) of the DNA test...)

# DNA PROOF

Coherent extensions:

- $0 \leq$ p(Guilty|Suspect and Inclusion) $\leq 100\%$

  (maximum vagueness for inference based only on laboratory results and preliminary investigations);

- $0 \leq$ p($\neg$ Guilty $\wedge$ Suspect $\wedge$ Inclusion) $\leq 5\%$

  (rare, but possible, "false positive" scenario where other elements should be investigated, like, e.g., the possibility of contamination during the DNA profiling procedure).

(**Note**: such conclusions are insensitive with respect to the sensibility p(Inclusion|Guilty) of the DNA test...)

# DNA PROOF

Coherent extensions:

- $0 \leq$ p(Guilty|Suspect and Inclusion) $\leq 100\%$

  (maximum vagueness for inference based only on laboratory results and preliminary investigations);

- $0 \leq$ p($\neg$ Guilty $\wedge$ Suspect $\wedge$ Inclusion) $\leq 5\%$

  (rare, but possible, "false positive" scenario where other elements should be investigated, like, e.g., the possibility of contamination during the DNA profiling procedure).

(**Note**: such conclusions are insensitive with respect to the sensibility p(Inclusion|Guilty) of the DNA test...)

# Conclusions

- Renewed importance of the coherence approach has been outlined through a review of some prototypical examples ;
- Small, well suited and on-demand models can solve the unavoidable partiality of our information;
- Even few, but well formalized, information could lead to rational, sound, significant,and sometimes unexpected, inferences if guided by the coherence approach.

# Conclusions

- Renewed importance of the coherence approach has been outlined through a review of some prototypical examples ;
- Small, well suited and on-demand models can solve the unavoidable partiality of our information;
- Even few, but well formalized, information could lead to rational, sound, significant,and sometimes unexpected, inferences if guided by the coherence approach.

# Conclusions

- Renewed importance of the coherence approach has been outlined through a review of some prototypical examples ;
- Small, well suited and on-demand models can solve the unavoidable partiality of our information;
- Even few, but well formalized, information could lead to rational, sound, significant,and sometimes unexpected, inferences if guided by the coherence approach.