Uncertainty in Computer Science 2021/22

DAVIDE CIUCCI UNIVERSITY OF MILANO-BICOCCA

UNCERTAINTY THEORIES BEYOND PROBABILITY

BEYOND PROBABILITY

Beyond probability

Extending probability

- Belief functions
- Imprecise probabilities

Different from probability

- Possibility Theory/Logic
- Fuzzy Sets/Logic
- Modal Logic
- Interval Sets
- Rough Sets

DIFFERENT TOOLS FOR DIFFERENT UNCERTAINTIES

- Possibility theory
- Belief function
- Rough Sets

- Starting from the notion of capacity

CAPACITY: EXAMPLES

A capacity is a number associated to a set

- Example 1
  - X = possible answers to a questions, only 1 answer is correct
  - The capacity of a set of answers $A \subseteq X$ quantifies the uncertainty that A contains the correct answer
  - The bigger is A, the greater its capacity

CAPACITY: EXAMPLES

Not necessarily related to uncertainty

- Example 2
  - X = a set of goods
  - The capacity of a set of goods $A \subseteq X$ is the total price
  - Usually, it is additive: the total price is the sum of the single prices
    - …not always: the whole collection of Naruto volumes has much value than the sum of single volumes

CAPACITY: FORMAL DEFINITION

A **capacity** on a universe X is a function $\mu : 2^X \mapsto L$

- L can be $\Re$, [0,1], …
- grounded $\mu(\emptyset) = 0$
- monotone $A \subseteq B \quad implies \quad \mu(A) \leq \mu(B)$

A capacity is **normalized** if $\mu(X) = 1$

CAPACITY: COMMENTS

- Our interpretation will mostly be $\mu(A)$ represents the confidence of an agent wrt A
- Monotony $A \subseteq B \quad implies \quad \mu(A) \leq \mu(B)$
  - If A implies B then the agent has at least confidence in B as in A
  - $\mu(A \cup B) \geq \max\{\mu(A), \mu(B)\}$
  - $\mu(A \cap B) \leq \min\{\mu(A), \mu(B)\}$
- Capacities are sometimes called fuzzy measures or confidence measures

CAPACITY AND PROBABILITY

More general than a probability measure

- An additive capacity is a probability measure
- Additivity: $P(A \cup B) = P(A) + P(B)$

CAPACITIES AND POSSIBILITY THEORY

More general than a possibility measure

- A **maxitive** capacity is a possibility measure
- Maxitivity: $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$
- What is a possibility measure? More generally, what is possibility theory?

POSSIBILITY THEORY

- The term was coined in 1978 by L. Zadeh: imprecise linguistic statements modeled by **fuzzy sets interpreted as possibility distributions**
- The theory was formalized and developed by Dubois and Prade
- Spohn (1988): independently define similar ideas

POSSIBILITY DISTRIBUTION

- A **possibility distribution** is a function $\pi : X \mapsto [0,1]$
- it represents the state of knowledge of an agent
  - the state a is **impossible** $\pi(a) = 0$
  - the state a is **totally possible** $\pi(a) = 1$
    - more that one state can be totally possible $\pi(a) = 1, \pi(b) = 1$
  - **complete knowledge** $\exists a, \pi(a) = 1 \quad \forall b \neq a, \pi(b) = 0$
  - **complete ignorance** $\forall a, \pi(a) = 1$
- Usually **normalization** is assumed: $\exists a, \pi(a) = 1$

POSSIBILITY & NECESSITY MEASURES

- **Possibility** measure $\Pi(A) = \sup_{a \in A} \pi(A)$
- **Necessity** measure $N(A) = 1 - \Pi(A^c) = \inf_{a \notin A} (1 - \pi(A))$
- Necessity implies possibility: $N(A) \leq \Pi(A)$
- Necessity is a **minitive** capacity:

$$N(A \cap B) = \min\{N(A), N(B)\}$$

EXAMPLES

- " I am highly confident that the thief's car is blue or black"
  - N({blue, black}) = 0.8
- "I exclude that she likes tomatoes, but she might like carrots and salad"
  $\Pi(\{tomatoes\}) = 0, N(\{carrots\} = N(\{salad\}) = 0.6$
- "I have no idea about A", "I do not know A"
  $\Pi(A) = 1, N(A) = 0$

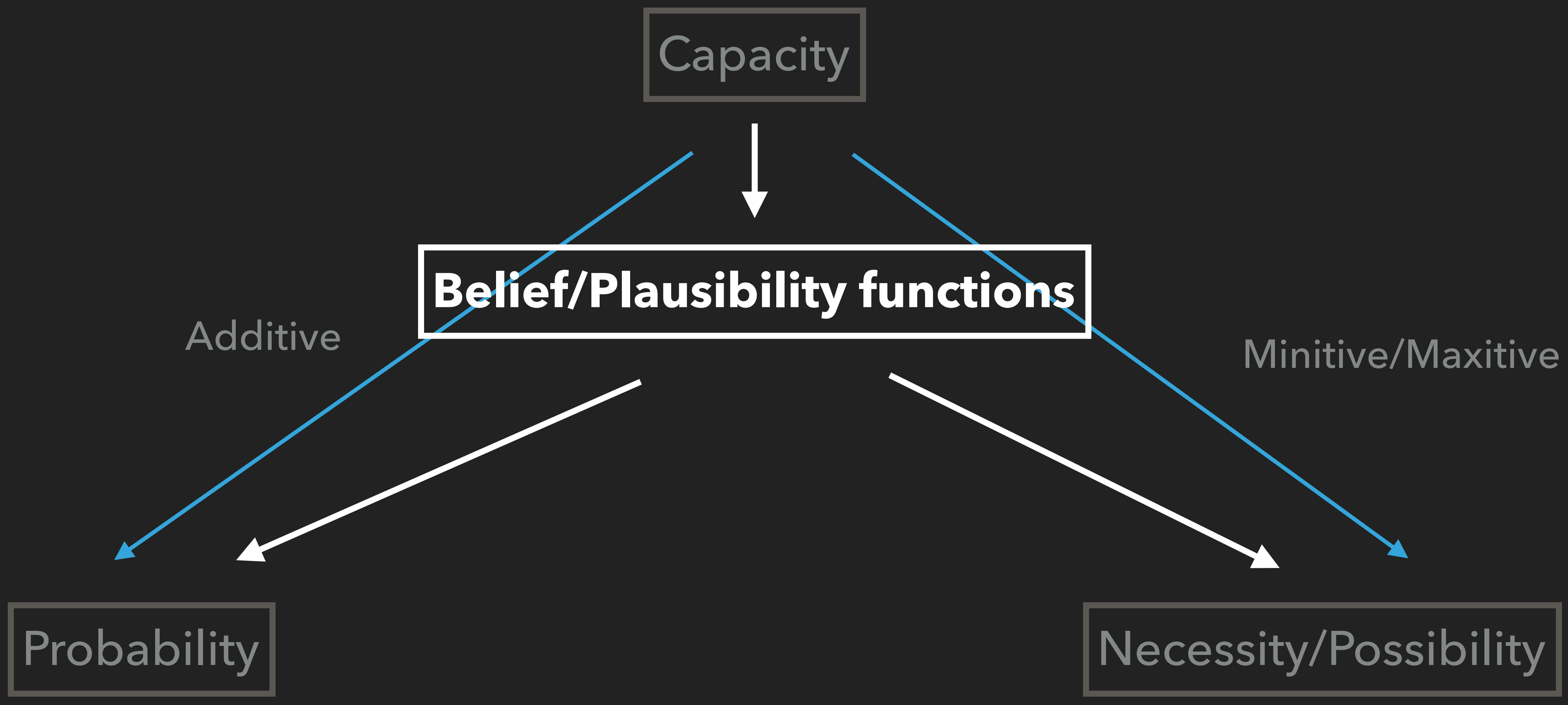
POSSIBILITY VS PROBABILITY

Probability	Possibility
$P(A) = \sum_{a \in A} p(a)$	$\Pi(A) = \sup_{a \in A}\{\pi(a)\}$
$\sum_{a \in X} p(a) = 1$	$\exists a, \pi(a) = 1$
$P(A) = 1 - P(A^c)$	$\Pi(A) = 1 - N(A^c)$ $\Pi(A) + \Pi(A^c) \geq 1$

BOOLEAN POSSIBILITY DISTRIBUTION

- A Boolean possibility distribution is a function $\pi : X \mapsto \{0,1\}$
- If X is a set of propositional variables, then π is **a partial truth assignment**
- X ={v1,v2,v3,v4}
- $\pi(v1) = 1(true), \pi(v2) = 1(true), \pi(v3) = 0(false)$
- v4 undefined

BELIEF FUNCTIONS AND CAPACITIES



BELIEF FUNCTION

- Evidence theory
  - “A mathematical theory of evidence”, Shafer, 1976
  - belief/plausibility functions: totally monotone/ alternating normalized capacities
- Dempster-Shafer theory
  - Dempster (1967) used the names lower and upper probabilities

THE IDEA

- To associate probabilities to incomplete and imprecise observations
  - A radar measuring the speed of vehicles is **imprecise,** it only returns an interval
  - I have a king of clubs in my hand and I reveal that it is a black card. Concerning the set {♠,♣,♡,♢} you have an **incomplete** information
- The true outcome lies in a set, but the exact value is unknown

MASS FUNCTION

- X a set of outcomes of an experiment
- $A \subseteq X$ an **event**
- We associate to an event a **(belief) mass distribution** $m : 2^X \mapsto [0,1]$ such that
  - $m(\emptyset) = 0$ (normalization)
  - $\sum_{A \subseteq X} m(A) = 1$ (m is a probability distribution)

MASS FUNCTION

- The card is black: m({♠,♣}) = 1
  - I am certain that the card is spade or club
- m(A) represents the belief committed to A, not to its subsets
  - The fact that the card is black does not tell anything on the fact it is a club or a spade
- **Focal sets**: m(A) >0

ON THE INTERPRETATION OF M

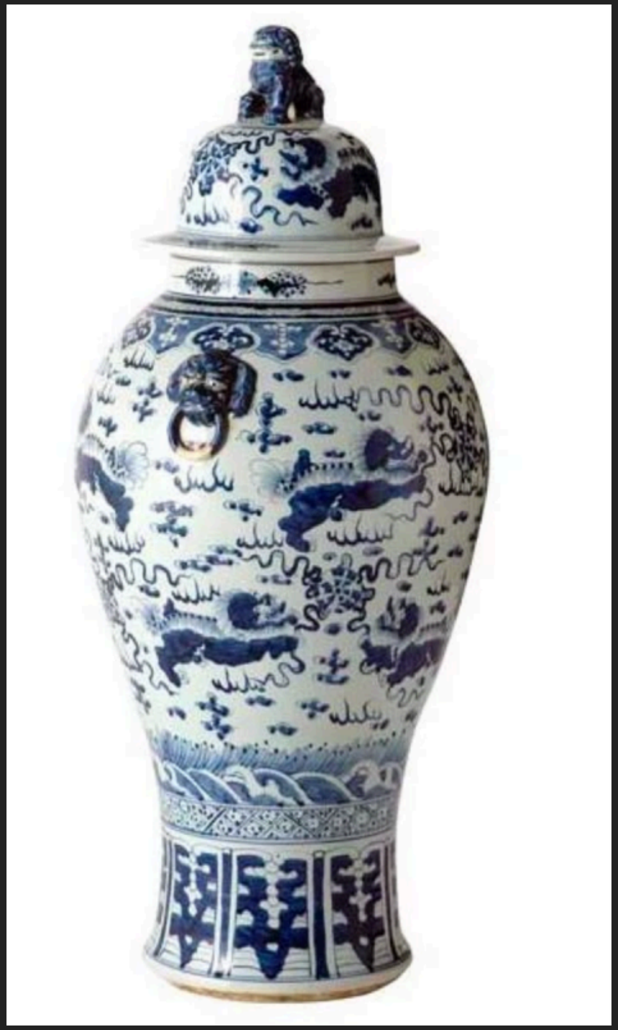
- m(A) is the probability that the agent only knows that x is in A –> Epistemic point of view
- It can be non-monotonic: $A \subseteq B, m(B) \leq m(A)$ if the agent is sure enough that x in A
- m(X) = 1 the agent knows nothing

BELIEF AND PLAUSIBILITY FUNCTIONS

- Belief function $Bel : 2^X \mapsto [0,1]$
  - $$Bel(A) = \sum_{B \subseteq A} m(B)$$
  - The total mass of belief certainly committed to A
- Plausibility function $Pl : 2^X \mapsto [0,1]$
  - $$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$$
  - The total mass of belief possibly related to A
- $Bel(A) \leq Pl(A)$, $Pl(A) = 1 - Bel(A^c)$ (as it was for possibility/necessity)

EXAMPLE: CHINESE VASE

- In an antique shop you see a wonderful, expensive vase



oroa.com, 1355€

amazon.com, 9.39€

EXAMPLE: CHINESE VASE

- X= {genuine, counterfeit}
- CASE 1: not an expert, total ignorance
  - total ignorance: m({genuine})=m({counterfeit}) = 0, m(X) =1
  - Bel({genuine})=Bel({counterfeit})= 0
  - Pl({genuine})=Pl({counterfeit}) = 1

EXAMPLE: CHINESE VASE

- X= {genuine, counterfeit}
- CASE 2: expert, there are as many clues in favor of genuine than of counterfeit
    - m({genuine})=m({counterfeit}) = 1/2, m(X) =0
    - Bel({genuine})=Bel({counterfeit})= 1/2
    - Pl({genuine})=Pl({counterfeit}) = 1/2
- In probability theory the two cases (expert/non expert) are not distinguishable
    - P(genuine)=P(counterfeit)=1/2

COMBINING EVIDENCE: DEMPSTER'S RULE

- m1, m2 two mass distributions, both reliable
- $m_{1,2}(\varnothing) = 0$
- $m_{1,2}(A) = \frac{1}{1-k} \sum_{B \cap C = A} m_1(B) m_2(C)$ where

  $K = \sum_{B \cap C = \varnothing} m_1(B) m_2(C)$
- K is a measure of conflict

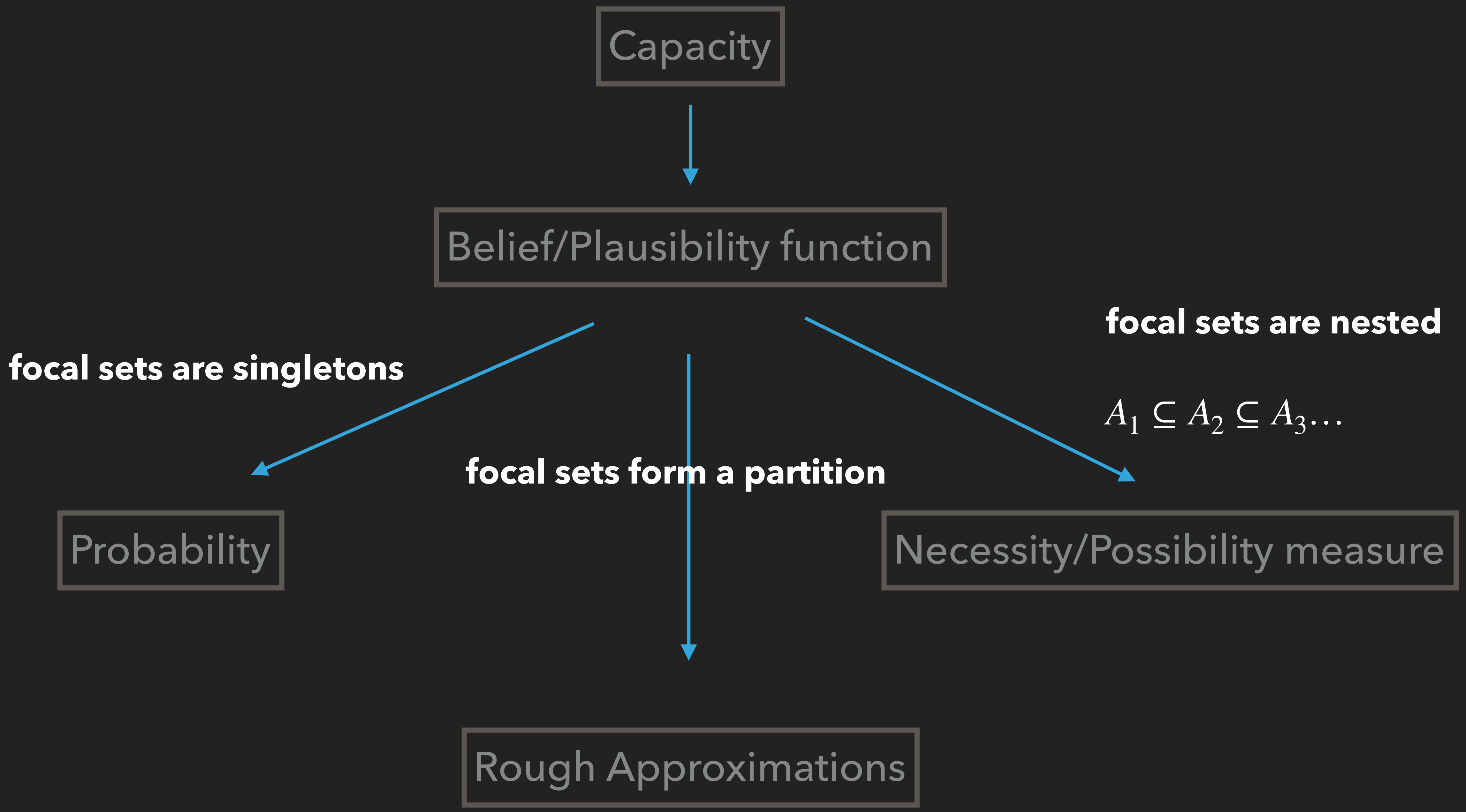
DEMPSTER'S RULE: (COUNTERINTUITIVE) EXAMPLE

- Doctor A: illness a with 99%, illness b with 1% certainty
- Doctor B: illness c with 99%, illness b with 1% certainty
- X = {a,b,c}
- m1({a})=0.99 m1({b})=0.01
- m2({c})=0.99 m1({b})=0.01
- K = 0.99*0.99+0.99*0.01+0.01*0.99 = 0.9999
- $m_{1,2}(\{a\}) = m_{1,2}(\{c\}) = 0, m_{1,2}(\{b\}) = 1$

OTHER RULES

- Not normalized: $m_{1,2}(A) = \sum_{B \cap C = A} m_1(B) m_2(C)$
- $m_{1,2}(\{a\}) = m_{1,2}(\{c\}) = 0,\ m_{1,2}(\{b\}) = 0.0001$
- Rule of thumb:
  - normalized in closed world hypothesis
  - Non-normalized in open world hypothesis
- m1, m2 two mass distributions, at least one reliable
- $(m_1 \cup m_2)(A) = \sum_{B \cup C = A} m_1(B) m_2(C)$
- Many others…

BELIEF FUNCTIONS AND CAPACITIES



QUALITATIVE CAPACITY

- A capacity on a universe X is a function $\mu : 2^X \mapsto L$
- L a non-numerical totally ordered set representing terms pertaining to belief
- L = {zero, very low, low, medium, high, very high, top}
- Similar theories to the quantitative case can be developed

SOME BIBLIOGRAPHY

- M. Grabisch, “Set Functions, Games and Capacities in Decision Making”, Springer, 2016
- Da “A Guided Tour of Artificial Intelligence Research”, (P. Marquis O. Papini, H. Prade, Eds), Springer, 2020
  - T. Denoeux, D. Dubois, H. Prade ”Representations of Uncertainty in AI: probability and possibility”, pp. 69-117 https://doi.org/10.1007/978-3-030-06164-7_3
  - T. Denoeux, D. Dubois, H. Prade”Representations of Uncertainty in AI: beyond probability and possibility”, pp. 119-150 https://doi.org/10.1007/978-3-030-06164-7_4

TOTALLY MONOTONE CAPACITY

- A capacity μ is totally monotone if

$\forall k \geq 2$ it is k-monotone

$$A_1, \ldots, A_k \quad \mu(\cup A_i) \geq \sum_{I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1}(\mu(\cap_{i \in I} A_i))$$

- A generalized version of

$$|\cup A_i| = \sum_{I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1}(\cap_{i \in I} A_i)$$