Probabilistic Semantics for Categorical Syllogisms of Figure II

Niki Pfeifer[1](✉) and Giuseppe Sanfilippo[2](✉)

[1] Munich Center for Mathematical Philosophy, LMU Munich, Geschwister-Scholl-Platz 1, 80539 Munich, Germany
niki.pfeifer@lmu.de

[2] Department of Mathematics and Computer Science, University of Palermo, Via Archirafi 34, 90123 Palermo, Italy
giuseppe.sanfilippo@unipa.it

Abstract. A coherence-based probability semantics for categorical syllogisms of Figure I, which have transitive structures, has been proposed recently (Gilio, Pfeifer, & Sanfilippo [15]). We extend this work by studying Figure II under coherence. Camestres is an example of a Figure II syllogism: from *Every* P *is* M and *No* S *is* M infer *No* S *is* P. We interpret these sentences by suitable conditional probability assessments. Since the probabilistic inference of $\bar{P}|S$ from the premise set $\{M|P, \overline{M}|S\}$ is not informative, we add $p(S|(S \vee P)) > 0$ as a probabilistic constraint (i.e., an "existential import assumption") to obtain probabilistic informativeness. We show how to propagate the assigned (precise or interval-valued) probabilities to the sequence of conditional events $(M|P, \overline{M}|S, S|(S \vee P))$ to the conclusion $\bar{P}|S$. Thereby, we give a probabilistic meaning to the other syllogisms of Figure II. Moreover, our semantics also allows for generalizing the traditional syllogisms to new ones involving generalized quantifiers (like *Most* S *are* P) and syllogisms in terms of defaults and negated defaults.

Keywords: Categorical syllogisms · Coherence · Conditional events
Defaults · Generalized quantifiers · Imprecise probability

1 Motivation and Outline

There is a long tradition in logic to investigate categorical syllogisms that goes back to Aristotle's *Analytica Priora*. However, not many authors proposed *probabilistic* semantics for categorical syllogisms (see, e.g., [9,10,13,15,23,38]) to overcome formal restrictions imposed by logic, like its *monotonicity* (i.e., the

N. Pfeifer and G. Sanfilippo—Shared first authorship (both authors contributed equally to this work).

N. Pfeifer—Supported by his DFG grant PF 740/2-2 (within the SPP1516).
G. Sanfilippo—Partially supported by the "National Group for Mathematical Analysis, Probability and their Applications" (GNAMPA – INdAM).

D. Ciucci et al. (Eds.): SUM 2018, LNAI 11142, pp. 196–211, 2018.
https://doi.org/10.1007/978-3-030-00461-3_14

inability to retract conclusions in the light of new evidence) or its qualitative nature (i.e., the inability to express *degrees of belief*). The main goal of building a probabilistic semantics is therefore to manage nonmonotonicity and degrees of belief, which are necessary for the formalization of commonsense reasoning. Although this paper is about probabilistic reasoning, applications of our results may include (i) relating ancient syllogisms to nonmonotonic reasoning by proposing a new nonmonotonic rule of inference and (ii) proposing a new rationality framework for the psychology of reasoning, specifically, for reasoning about a particular set of quantified statements (see, e.g., [29–32,35]). Moreover, (iii) our results are also applicable in formal semantics: specifically, our probabilistic approach is scalable in the sense that the proposed semantics allows for managing not only traditional logical quantifiers but also the much bigger superset of generalized quantifiers.

What are classical categorical syllogisms? They are valid argument forms consisting of two premises and a conclusion, which are composed of basic syllogistic sentence types (see, e.g., [28]): (A) *Every a is b*, (E) *No a is b*, (I) *Some a is b*, and (O) *Some a is not b*, where "a" and "b" denote two of the three categorical terms M ("middle term"), P ("predicate term"), or S ("subject term"). The M term appears only in the premises and are combined with P (in the first premise) and S (in the second premise). The predicates contained in the conclusion appear only in the order (S, P). By all possible permutations of the predicate order, four syllogistic *figures* result under the given restrictions. Syllogisms of Figure I, for example, have a transitive structure, i.e., M *is* P, S *is* M, therefore S *is* P. Consider *(Modus) Barbara* as an instance of a syllogism of Figure I: *Every M is P, Every S is M, therefore Every S is P*. The syllogism's name traditionally encodes logical properties. For the present purpose, we only recall that vocals refer to the syllogistic sentence type: for instance, B<u>a</u>rb<u>a</u>r<u>a</u> involves only sentences of type (A) (see, e.g., [28] for details). Our paper is based on [15,19], where a coherence-based probability semantics for categorical syllogisms of Figure I was studied. We extend this work to Figure II, which has the following structure: P *is* M, S *is* M, therefore S *is* P. *Camestres* is an instance of a Figure II syllogism: *Every P is M, No S is M, therefore No S is P*. Camestres involves the sentence types (A) and (E).

While the S, M, and P terms are interpreted as predicate terms in first order logic, we interpret them as events as follows. Imagine a random experiment where the (random) outcome is denoted by X. Consider, for example, the predicate S. Depending on the outcome of the experiment, X may satisfy or not satisfy the predicate S. Then, we denote by E_S the event "X satisfies S" (the event E_S is true if X satisfies the predicate S and E_S is false if X does not satisfy S). We conceive the predicate S as the event E_S, which will be true or false. Thus, we simply identify E_S by S (in this sense S is both a predicate and an event). The same reasoning applies to the P and M terms, which are in our context both predicates and events. On the level of events, we associate pairs of predicates (S, P) with the corresponding conditional event $P|S$. On the level of probability assessments, we interpret the degree of belief in syllogistic sentence

(A) by $p(P|S) = 1$, (E) by $p(\bar{P}|S) = 1$, (I) by $P(P|S) > 0$, and we interpret (O) by $p(\bar{P}|S) > 0$ (see also [9,15] for similar interpretations and [33,34] for basic relations among these probabilistic sentence types). Thus, (A) and (E) are interpreted as precise probability assessments and (I) and (O) by imprecise probability assessments.

We note that, like the probabilistic Modus Barbara [15], the probabilistic Camestres is not probabilistically informative without existential import assumption: indeed, $p(M|P) = 1, p(\overline{M}|S) = 1 \implies 0 \leqslant p(\bar{P}|S) \leqslant 1$. We propose to add the conditional event existential import (i.e., $p(S|(S \vee P)) > 0$, which was originally proposed in the context of Weak Transitivity, see [15]) to the premise set to make Camestres probabilistically informative:

$$\text{(Camestres) } p(M|P) = 1, p(\overline{M}|S) = 1, \text{ and } p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) = 1.$$

After recalling some preliminary notions and results in Sect. 2, we show how to propagate the assigned probabilities to the sequence of conditional events $(M|P, \overline{M}|S, S|(S \vee P))$ to the conclusion $\bar{P}|S$ in Sect. 3. This result is applied in Sect. 4, where we give a probabilistic meaning to the other syllogisms of Figure II. Section 5 concludes by remarks on further applications (generalized quantifiers and nonmonotonic reasoning) and future work.

2 Preliminary Notions and Results

In this section we recall selected key features of coherence (for more details see, e.g., [5,7,11,12,20,21,27,37]). Given two events E and H, with $H \neq \bot$, the *conditional event* $E|H$ is defined as a three-valued logical entity which is *true* if EH (i.e., $E \wedge H$) is true, *false* if $\bar{E}H$ is true, and *void* if H is false. In betting terms, assessing $p(E|H) = x$ means that, for every real number s, you are willing to pay an amount $s \cdot x$ and to receive s, or 0, or $s \cdot x$, according to whether EH is true, or $\bar{E}H$ is true, or $\bar{H}$ is true (i.e., the bet is called off), respectively. In these cases the random gain is $\mathcal{G} = sH(E - x)$. More generally speaking, consider a real-valued function $p : \mathcal{K} \rightarrow \mathbb{R}$, where $\mathcal{K}$ is an arbitrary (possibly not finite) family of conditional events. Let $\mathcal{F} = (E_1|H_1, \ldots, E_n|H_n)$ be a sequence of conditional events, where $E_i|H_i \in \mathcal{K}$, $i = 1, \ldots, n$, and let $\mathcal{P} = (p_1, \ldots, p_n)$ be the vector of values $p_i = p(E_i|H_i)$, where $i = 1, \ldots, n$. We denote by $\mathcal{H}_0$ the disjunction $H_1 \vee \cdots \vee H_n$. With the pair $(\mathcal{F}, \mathcal{P})$ we associate the random gain $\mathcal{G} = \sum_{i=1}^{n} s_i H_i (E_i - p_i)$, where $s_1, \ldots, s_n$ are n arbitrary real numbers. $\mathcal{G}$ represents the net gain of n transactions. Let $G_{\mathcal{H}_0}$ denote the set of possible values of $\mathcal{G}$ restricted to $\mathcal{H}_0$, that is, the values of $\mathcal{G}$ when at least one conditioning event is true.

Definition 1. *Function p defined on $\mathcal{K}$ is* coherent *if and only if, for every integer n, for every sequence $\mathcal{F}$ of n conditional events in $\mathcal{K}$ and for every $s_1, \ldots, s_n$, it holds that:* $\min G_{\mathcal{H}_0} \leqslant 0 \leqslant \max G_{\mathcal{H}_0}$.

Intuitively, Definition 1, means in betting terms that a probability assessment is coherent if and only if, in any finite combination of n bets, it cannot happen that the values in $G_{\mathcal{H}_0}$ are all positive, or all negative (*no Dutch Book*).

We recall the fundamental theorem of de Finetti for conditional events, which states that a coherent assessment of premises can always be coherently extended to a conclusion:

Theorem 1. *Let a coherent probability assessment $\mathcal{P} = (p_1, \dots, p_n)$ on a sequence $\mathcal{F} = (E_1|H_1, \dots, E_n|H_n)$ be given. Moreover, given a further conditional event $E_{n+1}|H_{n+1}$. Then, there exists a suitable closed interval $[z', z''] \subseteq [0, 1]$ such that the extension $(\mathcal{P}, z)$ of $\mathcal{P}$ to $(\mathcal{F}, E_{n+1}|H_{n+1})$ is coherent if and only if $z \in [z', z'']$.*

For applying Theorem 1, we now recall an algorithm which allows for computing the interval of coherent extensions $[z', z'']$ on $E_{n+1}|H_{n+1}$ from a coherent probability assessment $\mathcal{P}$ on a finite sequence of conditional events $\mathcal{F}$ (see [15, Algorithm 1], which is originally based on [5, Algorithm 2]).

Algorithm 1. Let $\mathcal{F} = (E_1|H_1, \dots, E_n|H_n)$ be a sequence of conditional events and $\mathcal{P} = (p_1, \dots, p_n)$ be a coherent precise probability assessment on $\mathcal{F}$, where $p_j = p(E_j|H_j) \in [0, 1]$, $j = 1, \dots, n$. Moreover, let $E_{n+1}|H_{n+1}$ be a further conditional event and denote by J_{n+1} the set $\{1, \dots, n+1\}$. The steps below describe the computation of the lower bound z' (resp., the upper bound z'') for the coherent extensions $z = p(E_{n+1}|H_{n+1})$.

– **Step 0.** Expand the expression $\bigwedge_{j \in J_{n+1}} \left(E_j H_j \vee \bar{E}_j H_j \vee \bar{H}_j\right)$ and denote by $C_1, \dots, C_m$ the constituents contained in $\mathcal{H}_0 = \bigvee_{j \in J_{n+1}} H_j$ associated with $(\mathcal{F}, E_{n+1}|H_{n+1})$. Then, construct the following system in the unknowns $\lambda_1, \dots, \lambda_m, z$

$$\begin{cases} \sum_{r: C_r \subseteq E_{n+1} H_{n+1}} \lambda_r = z \sum_{r: C_r \subseteq H_{n+1}} \lambda_r \; ; \\ \sum_{r: C_r \subseteq E_j H_j} \lambda_r = p_j \sum_{r: C_r \subseteq H_j} \lambda_r, \;\; j \in J_n \; ; \\ \sum_{r \in J_m} \lambda_r = 1; \;\; \lambda_r \geq 0, \; r \in J_m \; . \end{cases} \tag{1}$$

– **Step 1.** Check the solvability of system (1) under the condition $z = 0$ (resp., $z = 1$). If the system (1) is not solvable go to Step 2, otherwise go to Step 3.
– **Step 2.** Solve the following linear programming problem

$$\text{Compute}: \qquad \gamma' = \min \sum_{r: C_r \subseteq E_{n+1} H_{n+1}} \lambda_r$$

$$(\text{respectively}: \qquad \gamma'' = \max \sum_{r: C_r \subseteq E_{n+1} H_{n+1}} \lambda_r \;)$$

subject to:

$$\begin{cases} \sum_{r: C_r \subseteq E_j H_j} \lambda_r = p_j \sum_{r: C_r \subseteq H_j} \lambda_r, \; j \in J_n \; ; \\ \sum_{r: C_r \subseteq H_{n+1}} \lambda_r = 1; \; \lambda_r \geqslant 0, \; r \in J_m . \end{cases}$$

The minimum γ' (respectively the maximum γ'') of the *objective function* coincides with z' (respectively with z'') and the procedure stops.

- **Step 3.** For each subscript $j \in J_{n+1}$, compute the maximum M_j of the function $\Phi_j = \sum_{r:C_r \subseteq H_j} \lambda_r$, subject to the constraints given by the system (1) with $z = 0$ (respectively $z = 1$). We have the following three cases:
  1. $M_{n+1} > 0$;
  2. $M_{n+1} = 0$, $M_j > 0$ for every $j \neq n+1$;
  3. $M_j = 0$ for $j \in I_0 = J \cup \{n+1\}$, with $J \neq \varnothing$.

  In the first two cases $z' = 0$ (respectively $z'' = 1$) and the procedure stops. In the third case, defining $I_0 = J \cup \{n+1\}$, set $J_{n+1} = I_0$ and $(\mathcal{F}, \mathcal{P}) = (\mathcal{F}_J, \mathcal{P}_J)$, where $\mathcal{F}_J = (E_i|H_i : i \in J)$ and $\mathcal{P}_J = (p_i : i \in J)$. Then, go to Step 0.

The procedure ends in a finite number of cycles by computing the value z' (respectively z'').

Remark 1. Assuming $(\mathcal{P}, z)$ on $(\mathcal{F}, E_{n+1}|H_{n+1})$ coherent, each solution $\Lambda = (\lambda_1, \ldots, \lambda_m)$ of System (1) is a coherent extension of the assessment $(\mathcal{P}, z)$ to the sequence $(C_1|\mathcal{H}_0, \ldots, C_m|\mathcal{H}_0)$.

Definition 2. *An* imprecise, or set-valued, assessment $\mathcal{I}$ *on a finite sequence of n conditional events $\mathcal{F}$ is a (possibly empty) set of precise assessments $\mathcal{P}$ on $\mathcal{F}$.*

Definition 2, introduced in [14], states that an *imprecise (probability) assessment* $\mathcal{I}$ on a finite sequence $\mathcal{F}$ of n conditional events is just a (possibly empty) subset of $[0,1]^n$. We recall the notions of g-coherence and total-coherence for imprecise (in the sense of set-valued) probability assessments [15].

Definition 3. *Let a sequence of n conditional events $\mathcal{F}$ be given. An imprecise assessment $\mathcal{I} \subseteq [0,1]^n$ on $\mathcal{F}$ is* g-coherent *if and only if there exists a coherent precise assessment $\mathcal{P}$ on $\mathcal{F}$ such that $\mathcal{P} \in \mathcal{I}$.*

Definition 4. *An imprecise assessment $\mathcal{I}$ on $\mathcal{F}$ is* totally coherent *(t-coherent) if and only if the following two conditions are satisfied: (i) $\mathcal{I}$ is non-empty; (ii) if $\mathcal{P} \in \mathcal{I}$, then $\mathcal{P}$ is a coherent precise assessment on $\mathcal{F}$.*

We denote by Π the set of *all coherent precise* assessments on $\mathcal{F}$. We recall that if there are no logical relations among the events $E_1, H_1, \ldots, E_n, H_n$ involved in $\mathcal{F}$, that is $E_1, H_1, \ldots, E_n, H_n$ are logically independent, then the set Π associated with $\mathcal{F}$ is the whole unit hypercube $[0,1]^n$. If there are logical relations, then the set Π *could be* a strict subset of $[0,1]^n$. As it is well known $\Pi \neq \varnothing$; therefore, $\varnothing \neq \Pi \subseteq [0,1]^n$.

Remark 2. We observe that:

$$\begin{aligned} \mathcal{I} \text{ is g-coherent} &\iff \Pi \cap \mathcal{I} \neq \varnothing \\ \mathcal{I} \text{ is t-coherent} &\iff \varnothing \neq \Pi \cap \mathcal{I} = \mathcal{I}. \end{aligned}$$

Then: $\mathcal{I}$ is t-coherent $\Rightarrow$ $\mathcal{I}$ is g-coherent.

Given a g-coherent assessment $\mathcal{I}$ on a sequence of n conditional events $\mathcal{F}$, for each coherent precise assessment $\mathcal{P}$ on $\mathcal{F}$, with $\mathcal{P} \in \mathcal{I}$, we denote by $[\alpha_{\mathcal{P}}, \beta_{\mathcal{P}}]$ the interval of coherent extensions of $\mathcal{P}$ to $E_{n+1}|H_{n+1}$; that is, the assessment $(\mathcal{P}, z)$ on $(\mathcal{F}, E_{n+1}|H_{n+1})$ is coherent if and only if $z \in [z'_{\mathcal{P}}, z''_{\mathcal{P}}]$. Then, defining the set

$$\Sigma = \bigcup\nolimits_{\mathcal{P} \in \Pi \cap \mathcal{I}} [z'_{\mathcal{P}}, z''_{\mathcal{P}}], \tag{2}$$

for every $z \in \Sigma$, the assessment $\mathcal{I} \times \{z\}$ is a g-coherent extension of $\mathcal{I}$ to $(\mathcal{F}, E_{n+1}|H_{n+1})$; moreover, for every $z \in [0,1] \setminus \Sigma$, the extension $\mathcal{I} \times \{z\}$ of $\mathcal{I}$ to $(\mathcal{F}, E_{n+1}|H_{n+1})$ is not g-coherent. We say that Σ is the *set of coherent extensions* of the imprecise assessment $\mathcal{I}$ on $\mathcal{F}$ to the conditional event $E_{n+1}|H_{n+1}$.

3 Figure II: Propagation of Probability Bounds

In this section, we prove the precise and imprecise probability propagation rules for the inference from $(B|C, \bar{B}|A, A|A \vee C)$ to $\bar{C}|A$. We apply our results in Sect. 4, where we give a probabilistically informative interpretation of categorical syllogisms of Figure II.

Remark 3. Let A, B, C be logically independent events. It can be proved that the assessment (x, y, z) on $(B|C, \bar{B}|A, \bar{C}|A)$ is coherent for every $(x, y, z) \in [0,1]^3$, that is, the imprecise assessment $\mathcal{I} = [0,1]^3$ on $(B|C, \bar{B}|A, \bar{C}|A)$ is totally coherent. For this it is sufficient to check that each of the eight vertices of the unit cube is coherent. Coherence can be checked, for example, by applying Algorithm 1 of [14] or by the CkC-package [3]. Moreover, it can also be proved that the assessment (x, y, t) on $(B|C, \bar{B}|A, A|A \vee C)$ is coherent for every $(x, y, t) \in [0,1]^3$, that is, the imprecise assessment $\mathcal{I} = [0,1]^3$ on $(B|C, \bar{B}|A, A|A \vee C)$ is totally coherent.

Given a coherent probability assessment (x, y, t) on the sequence of conditional events $(B|C, \bar{B}|A, A|A \vee C)$. The next result allows for computing the lower and upper bounds, z' and z'' respectively, for the coherent extension $z = p(\bar{C}|A)$.

Theorem 2. *Let A, B, C be three logically independent events and $(x, y, t) \in [0,1]^3$ be a (coherent) assessment on the family $\big(B|C, \bar{B}|A, A|A \vee C\big)$. Then, the extension $z = P(\bar{C}|A)$ is coherent if and only if $z \in [z', z'']$, where*

$$[z', z''] = \begin{cases} \quad [0,1]\,, & \text{if } t \leqslant x + yt \leqslant 1, \\ \left[\dfrac{x + yt - 1}{t\,x},\ 1\right], & \text{if } x + yt > 1, \\ \left[\dfrac{t - x - yt}{t\,(1 - x)},\ 1\right], & \text{if } x + yt < t. \end{cases}$$

Proof. We now apply Algorithm 1 in a symbolic way.
Computation of the lower probability bound z' on $\bar{C}|A$.
Input. $\mathcal{F} = (B|C, \bar{B}|A, A|A \vee C)$, $E_{n+1}|H_{n+1} = \bar{C}|A$.

Step 0. The constituents associated with $(B|C, \bar{B}|A, A|A\vee C, \bar{C}|A)$ and contained in $\mathcal{H}_0 = A \vee C$ are $C_1 = ABC, C_2 = AB\bar{C}, C_3 = A\bar{B}C, C_4 = A\bar{B}\bar{C}, C_5 = \bar{A}BC$, and $C_6 = \bar{A}\bar{B}C$. We construct the following starting system with unknowns $\lambda_1, \ldots, \lambda_6, z$ (see Remark 1):

$$\begin{cases} \lambda_2 + \lambda_4 = z(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4),\ \lambda_1 + \lambda_5 = x(\lambda_1 + \lambda_3 + \lambda_5 + \lambda_6), \\ \lambda_3 + \lambda_4 = y(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4), \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = t(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6), \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 = 1,\ \lambda_i \geq 0,\ i = 1, \ldots, 6. \end{cases} \tag{3}$$

Step 1. By setting $z = 0$ in System (3), we obtain

$$\begin{cases} \lambda_2 + \lambda_4 = 0,\ \lambda_1 + \lambda_5 = x, \\ \lambda_3 = y(\lambda_1 + \lambda_3),\ \lambda_1 + \lambda_3 = t, \\ \lambda_1 + \lambda_3 + \lambda_5 + \lambda_6 = 1, \\ \lambda_i \geq 0,\ i = 1, \ldots, 6. \end{cases} \iff \begin{cases} \lambda_1 = t(1-y),\ \lambda_2 = 0,\ \lambda_3 = yt, \\ \lambda_4 = 0,\ \lambda_5 = x - t(1-y), \\ \lambda_6 = 1 - x - yt, \\ \lambda_i \geq 0,\ i = 1, \ldots, 6. \end{cases} \tag{4}$$

The solvability of System (4) is a necessary condition for the coherence of the assessment $(x, y, t, 0)$ on $(B|C, \bar{B}|A, A|A \vee C, \bar{C}|A)$. As $(x, y, t) \in [0, 1]^3$, it holds that: $\lambda_1 = t(1-y) \geqslant 0$, $\lambda_3 = yt \geqslant 0$. Thus, System (4) is solvable if and only if $\lambda_5 \geqslant 0$ and $\lambda_6 \geqslant 0$, that is

$$t - yt \leqslant x \leqslant 1 - yt \iff t \leqslant x + yt \leqslant 1.$$

We distinguish two cases: (i) $x + yt > 1 \vee x + yt < t$; (ii) $t \leqslant x + yt \leqslant 1$. In Case (i), System (4) is not solvable (which implies that the coherent extension z of (x, y, t) must be positive). Then, we go to Step 2 of the algorithm where the (positive) lower bound z' is obtained by optimization. In Case (ii), System (4) is solvable and in order to check whether $z = 0$ is a coherent extension, we go to Step 3.

Case (i). We observe that in this case t cannot be 0. By Step 2 we have the following linear programming problem:
Compute $z' = \min(\lambda_2 + \lambda_4)$ *subject to:*

$$\begin{cases} \lambda_1 + \lambda_5 = x(\lambda_1 + \lambda_3 + \lambda_5 + \lambda_6),\ \lambda_3 + \lambda_4 = y(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4), \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = t(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6), \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1,\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases} \tag{5}$$

In this case, the constraints in (5) can be rewritten in the following way

$$\begin{cases} \lambda_1 + \lambda_5 = x(\lambda_1 + \lambda_3 + \lambda_5 + \lambda_6), \\ \lambda_3 + \lambda_4 = y,\ \lambda_5 + \lambda_6 = \frac{1-t}{t}, \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1, \\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6, \end{cases} \Leftrightarrow \begin{cases} 1 - y - \lambda_2 + \lambda_5 = x(1 - \lambda_2 - \lambda_4 + \frac{1-t}{t}), \\ \lambda_3 = y - \lambda_4,\ \lambda_6 = \frac{1-t}{t} - \lambda_5, \\ \lambda_1 = 1 - y - \lambda_2, \\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6, \end{cases}$$

or equivalently

$$\begin{cases} x\lambda_4 + (1-y) + \lambda_5 = \lambda_2(1-x) + \frac{x}{t},\ \lambda_3 = y - \lambda_4, \\ \lambda_5 = \frac{1-t}{t} - \lambda_6,\ \lambda_1 = 1 - y - \lambda_2,\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases}$$

We distinguish two (alternative) cases: $(i.1)$ $x + yt > 1$; $(i.2)$ $x + yt < t$.
Case $(i.1)$. The constraints in (5) can be rewritten in the following way

$$\begin{cases} x(\lambda_2 + \lambda_4) = \frac{x}{t} - (1-y) - \frac{1-t}{t} + \lambda_2 + \lambda_6,\ \lambda_3 = y - \lambda_4, \\ \lambda_5 = \frac{1-t}{t} - \lambda_6,\ \lambda_1 = 1 - y - \lambda_2,\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases}$$

As $x > 1 - ty$, we observe that $x > 0$. Then, the minimum of $z = \lambda_2 + \lambda_4$, obtained when $\lambda_2 = \lambda_6 = 0$, is

$$z' = \frac{1}{x}\left(\frac{x}{t} - (1-y) - \frac{1-t}{t}\right) = \frac{x - t + yt - 1 + t}{xt} = \frac{x + yt - 1}{xt}. \qquad (6)$$

By choosing $\lambda_2 = \lambda_6 = 0$ the constraints in (5) are satisfied with

$$\begin{cases} \lambda_1 = 1 - y,\ \lambda_2 = 0,\ \lambda_3 = y - \frac{x+yt-1}{xt},\ \lambda_4 = \frac{x+yt-1}{xt}, \\ \lambda_5 = \frac{1-t}{t},\ \lambda_6 = 0,\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases}$$

In particular $\lambda_3 \geqslant 0$ is satisfied because the condition $\frac{x+yt-1}{xt} \leqslant y$, which in this case amounts to $yt(1-x) \leq 1 - x$, is always satisfied. Then, the *procedure stops* yielding as *output* $z' = \frac{x+yt-1}{xt}$.
Case $(i.2)$. The constraints in (5) can be rewritten in the following way

$$\begin{cases} (1-y) - \frac{x}{t} + \lambda_5 + \lambda_4 = \lambda_2(1-x) - x\lambda_4 + \lambda_4,\ \lambda_3 = y - \lambda_4, \\ \lambda_6 = \frac{1-t}{t} - \lambda_5,\ \lambda_1 = 1 - y - \lambda_2,\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6, \end{cases}$$

or equivalently

$$\begin{cases} (\lambda_2 + \lambda_4)(1-x) = (1-y) - \frac{x}{t} + \lambda_4 + \lambda_5,\ \lambda_3 = y - \lambda_4, \\ \lambda_6 = \frac{1-t}{t} - \lambda_5,\ \lambda_1 = 1 - y - \lambda_2,\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases}$$

As $t - yt - x > 0$, that is $x < t(1-y)$, it holds that $x < 1$. Then, the minimum of $z = \lambda_2 + \lambda_4$, obtained when $\lambda_4 = \lambda_5 = 0$, is

$$z' = \frac{1}{1-x}\left(1 - y - \frac{x}{t}\right) = \frac{t - yt - x}{(1-x)t} \geqslant 0.$$

We observe that by choosing $\lambda_4 = \lambda_5 = 0$ the constraints in (5) are satisfied, indeed they are

$$\begin{cases} \lambda_1 = 1 - y,\ \lambda_2 = \frac{t-yt-x}{(1-x)t},\ \lambda_3 = y,\ \lambda_4 = 0, \\ \lambda_5 = 0,\ \lambda_6 = \frac{1-t}{t},\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases}$$

Then, the *procedure stops* yielding as *output* $z' = \frac{t-yt-x}{(1-x)t}$.

Case (ii). We take Step 3 of the algorithm. We denote by Λ and $\mathcal{S}$ the vector of unknowns $(\lambda_1, \ldots, \lambda_6)$ and the set of solution of System (4), respectively. We consider the following linear functions (associated with the conditioning events $H_1 = C, H_2 = H_4 = A, H_3 = A \vee C$) and their maxima in $\mathcal{S}$:

$$\begin{array}{l} \Phi_1(\Lambda) = \sum_{r:C_r \subseteq C} \lambda_r = \lambda_1 + \lambda_3 + \lambda_5 + \lambda_6, \\ \Phi_2(\Lambda) = \Phi_4(\Lambda) = \sum_{r:C_r \subseteq A} \lambda_r = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4, \\ \Phi_3(\Lambda) = \sum_{r:C_r \subseteq A \vee C} \lambda_r = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6, \\ M_i = \max_{\Lambda \in \mathcal{S}} \Phi_i(\Lambda),\ \ i = 1, 2, 3, 4\ . \end{array} \qquad (7)$$

By (4) we obtain: $\Phi_1(\Lambda) = 1$, $\Phi_2(\Lambda) = \Phi_4(\Lambda) = t$, $\Phi_3(\Lambda) = 1$, $\forall \Lambda \in \mathcal{S}$. Then, $M_1 = 1$, $M_2 = M_4 = t$, and $M_3 = 1$. We consider two subcases: $t > 0$; $t = 0$. If $t > 0$, then $M_4 > 0$ and we are in the first case of Step 3. Thus, the *procedure stops* and yields $z' = 0$ as *output*.
If $t = 0$, then $M_1 > 0, M_3 > 0$ and $M_2 = M_4 = 0$. Hence, we are in third case of Step 3 with $J = \{2\}, I_0 = \{2, 4\}$ and the procedure restarts with Step 0, with $\mathcal{F}$ replaced by $\mathcal{F}_J = (\bar{B}|A)$.

(2^{nd} cycle) Step 0. The constituents associated with $(\bar{B}|A, \bar{C}|A)$, contained in A, are $C_1 = ABC, C_2 = AB\bar{C}, C_3 = A\bar{B}C, C_4 = A\bar{B}\bar{C}$. The starting system is

$$\begin{cases} \lambda_3 + \lambda_4 = y(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4),\ \lambda_2 + \lambda_4 = z(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4), \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1,\ \lambda_i \geqslant 0, \quad i = 1, \ldots, 4. \end{cases} \tag{8}$$

(2^{nd}*cycle*) *Step*1. By setting $z = 0$ in System (8), we obtain

$$\{\lambda_1 = 1 - y, \quad \lambda_2 = \lambda_4 = 0, \quad \lambda_3 = y, \quad \lambda_i \geqslant 0, \quad i = 1, \ldots, 4. \tag{9}$$

As $y \in [0, 1]$, System (9) is always solvable; thus, we go to Step 3.
(2^{nd} *cycle*) *Step* 3. We denote by Λ and $\mathcal{S}$ the vector of unknowns $(\lambda_1, \ldots, \lambda_4)$ and the set of solution of System (9), respectively. The conditioning events are $H_2 = A$ and $H_4 = A$; then the associated linear functions are: $\Phi_2(\Lambda) = \Phi_4(\Lambda) = \sum_{r:C_r \subseteq A} \lambda_r = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$. From System (9), we obtain: $\Phi_2(\Lambda) = \Phi_4(\Lambda) = 1$, $\forall \Lambda \in \mathcal{S}$; so that $M_2 = M_4 = 1$. We are in the first case of Step 3 of the algorithm; then the *procedure stops* and yields $z' = 0$ as *output*.

To summarize, for any $(x, y, t) \in [0, 1]^3$ on $(B|C, \bar{B}|A, A|A \vee C)$, we have computed the coherent lower bound z' on $\bar{C}|A$. In particular, if $t = 0$, then $z' = 0$. We also have $z' = 0$, when $t > 0$ and $t \leqslant x + yt \leqslant 1$, that is when $0 < t \leqslant x + yt \leqslant 1$. Then, we can write that $z' = 0$, when $t \leqslant x + yt \leqslant 1$. Otherwise, we have two cases: $(i.1)$ $z' = \frac{x+yt-1}{xt}$, if $x + yt > 1$; $(i.2)$ $z' = \frac{t-yt-x}{(1-x)t}$, if $x + yt < t$.

Computation of the Upper Probability Bound z'' on $\bar{C}|A$
Input and *Step 0* are the same as in the proof of z'.
Step 1. By setting $z = 1$ in System (3), we obtain

$$\begin{cases} \lambda_1 + \lambda_3 = 0,\ \lambda_5 = x(\lambda_5 + \lambda_6), \\ \lambda_4 = y(\lambda_2 + \lambda_4),\ \lambda_2 + \lambda_4 = t, \\ \lambda_2 + \lambda_4 + \lambda_5 + \lambda_6 = 1, \\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases} \Longleftrightarrow \begin{cases} \lambda_1 = \lambda_3 = 0,\ \lambda_2 = t(1 - y), \\ \lambda_4 = yt,\ \lambda_5 = x(1 - t), \\ \lambda_6 = (1 - x)(1 - t), \\ \lambda_i \geqslant 0,\ i = 1, \ldots, 6. \end{cases} \tag{10}$$

As $(x, y, t) \in [0, 1]^3$, System (10) is solvable and we go to Step 3.
Step 3. We denote by Λ and $\mathcal{S}$ the vector of unknowns $(\lambda_1, \ldots, \lambda_6)$ and the set of solution of System (10), respectively. We consider the functions given in (7). From System (10), we obtain $M_1 = x(1 - t) + (1 - x)(1 - t) = 1 - t$, $M_2 = M_4 = t$, and $M_3 = 1$. If $t > 0$, then $M_4 > 0$ and we are in the first case of Step 3. Thus, the *procedure stops* and yields $z'' = 1$ as *output*. If $t = 0$,

then $M_1 > 0, M_3 > 0$ and $M_2 = M_4 = 0$. Hence, we are in the third case of Step 3 with $J = \{2\}, I_0 = \{2, 4\}$ and the procedure restarts with Step 0, with $\mathcal{F}$ replaced by $\mathcal{F}_J = (E_2|H_2) = (\bar{B}|A)$ and $\mathcal{P}$ replaced by $\mathcal{P}_J = y$.
(2^{nd} *cycle*) *Step* 0. This is the same as the (2^{nd} cycle) Step 0 in the proof of z'.
(2^{nd} *cycle*) *Step* 1. By setting $z = 1$ in System (3), we obtain

$$\left\{ \lambda_1 + \lambda_3 = 0, \quad \lambda_4 = y, \quad \lambda_2 = 1 - y, \quad \lambda_i \geqslant 0, \quad i = 1, \dots, 4. \right. \tag{11}$$

As $y \in [0, 1]$, System (11) is always solvable; thus, we go to Step 3.
(2^{nd} *cycle*) *Step* 3. Like in the (2^{nd} cycle) Step 3 of the proof of z', we obtain $M_4 = 1$. Thus, the procedure stops and yields $z'' = 1$ as output. To summarize, for any assessment $(x, y, t) \in [0, 1]^3$ on $(B|C, \bar{B}|A, A|A \vee C)$, we have computed the coherent upper probability bound z'' on $\bar{C}|A$, which is always $z'' = 1$. $\square$

Remark 4. We observe that in Theorem 2 we do not presuppose, differently from the classical approach, positive probability for the conditioning events (A and C). For example, even if we assume $p(A|A \vee C) = t > 0$ we do not require positive probability for the conditioning event A, and $p(A)$ could be zero (indeed, since $p(A) = p(A \wedge (A \vee C)) = p(A|A \vee C)p(A \vee C)$, $p(A) > 0$ implies $p(A|A \vee C) > 0$, but not vice versa). Moreover, we used a general and global approach for obtaining the inference rule in Theorem 2 (see [1,2,6,8,24,25] on local versus global approaches).

The next result is based on Theorem 2 and presents the set of the coherent extensions of a given interval-valued probability assessment $\mathcal{I} = ([x_1, x_2] \times [y_1, y_2] \times [t_1, t_2]) \subseteq [0, 1]^3$ on the sequence on $\big(B|C, \bar{B}|A, A|A \vee C\big)$ to the further conditional event $\bar{C}|A$.

Theorem 3. *Let A, B, C be three logically independent events and $\mathcal{I} = ([x_1, x_2] \times [y_1, y_2] \times [t_1, t_2]) \subseteq [0, 1]^3$ be an imprecise assessment on $\big(B|C, \bar{B}|A, A|A \vee C\big)$. Then, the set Σ of the coherent extensions of $\mathcal{I}$ on $\bar{C}|A$ is the interval $[z^*, z^{**}]$, where*

$$[z^*, z^{**}] = \begin{cases} [0, 1]\,, & \textit{if } (x_2 + y_2 t_1 \geqslant t_1) \wedge (x_1 + y_1 t_1 \leqslant 1), \\ \Big[\dfrac{x_1 + y_1 t_1 - 1}{t_1 x_1}, 1\Big], & \textit{if } x_1 + y_1 t_1 > 1, \\ \Big[\dfrac{t_1 - x_2 - y_2 t_1}{t_1(1 - x_2)}, 1\Big], & \textit{if } x_2 + y_2 t_1 < t_1. \end{cases}$$

Proof. As from Remark 3 the set $[0, 1]^3$ on $\big(B|C, \bar{B}|A, A|A \vee C\big)$ is totally coherent, then $\mathcal{I}$ is totally coherent too. Then, $\Sigma = \bigcup_{\mathcal{P} \in \mathcal{I}} [z'_{\mathcal{P}}, z''_{\mathcal{P}}] = [z^*, z^{**}]$, where $z^* = \inf_{\mathcal{P} \in \mathcal{I}} z'_{\mathcal{P}}$ and $z^{**} = \sup_{\mathcal{P} \in \mathcal{I}} z''_{\mathcal{P}}$. We distinguish three alternative cases: (i) $x_1 + y_1 t_1 > 1$; (ii) $x_2 + y_2 t_1 < t_1$; (iii) $(x_2 + y_2 t_1 \geqslant t_1) \wedge (x_1 + y_1 t_1 \leqslant 1)$.
Of course, for all three cases $z^{**} = \sup_{\mathcal{P} \in \mathcal{I}} z''_{\mathcal{P}} = 1$.
Case (i). We observe that the function $x + yt : [0, 1]^3$ is non-decreasing in the arguments x, y, t. Then, in this case, $x + yt \geqslant x_1 + y_1 t_1 > 1$ for every $\mathcal{P} = (x, y, t) \in \mathcal{I}$ and hence by Theorem 2 $z'_{\mathcal{P}} = f(x, y, t) = \frac{x+yt-1}{t\,x}$ for every $\mathcal{P} \in \mathcal{I}$.

Moreover, $f(x, y, t) : [0, 1]^3$ is non-decreasing in the arguments x, y, t, thus $z^* = \frac{x_1+y_1t_1-1}{t_1x_1}$.
Case (ii). We observe that the function $x + yt - t : [0, 1]^3$ is non-decreasing in the arguments x, y and non-increasing in the argument t. Then, in this case, $x+yt-t \leqslant x_2+y_2t_1-t_1 < 0$ for every $\mathcal{P} = (x, y, t) \in \mathcal{I}$ and hence by Theorem 2 $z'_{\mathcal{P}} = g(x, y, t) = \frac{t-x-yt}{t(1-x)}$ for every $\mathcal{P} \in \mathcal{I}$. Moreover, $g(x, y, t) : [0, 1]^3$ is non-increasing in the arguments x, y and non-decreasing in the argument t. Thus, $z^* = \frac{t_1-x_2-y_2t_1}{t_1(1-x_2)}$. Case (iii). In this case there exists a vector $(x, y, t) \in \mathcal{I}$ such that $t \leqslant x + yt \leqslant 1$ and hence by Theorem 2 $z'_{\mathcal{P}} = 0$. Thus, $z^* = 0$. $\square$

Remark 5. By instantiating Theorem 3 with the imprecise assessment $\mathcal{I} = \{1\} \times [y_1, 1] \times [t_1, 1]$, where $t_1 > 0$, we obtain the following lower and upper bounds for the conclusion $[z^*, z^{**}] = [y_1, 1]$. Thus, for every $t_1 > 0$: z^* depends only on the value of y_1.

4 Some Categorical Syllogisms of Figure II

In this section we consider examples of probabilistic categorical syllogisms of Figure II (Camestres, Camestrop, Baroco, Cesare, Cesaro, Festino) by suitable instantiations in Theorem 2. We consider three events P, M, S corresponding to the predicate, middle, and the subject term, respectively.

Camestres. The direct probabilistic interpretation of the categorical syllogism "*Every P is M, No S is M, therefore No S is P*" would correspond to infer $p(\bar{P}|S) = 1$ from the premises $p(M|P) = 1$ and $p(\overline{M}|S) = 1$; however, this inference is not justified. Indeed, by Remark 3, a probability assessment $(1, 1, z)$ on $(M|P, \overline{M}|S, \bar{P}|S)$ is coherent for every $z \in [0, 1]$. In order to construct a probabilistically informative version of Camestres, a further constraint of the premise set is needed. Like in [15], we use the *conditional event existential import* for further constraining the premise set: this is defined by the conditional probability of the conditioning event of the conclusion given the disjunction of all conditioning events. For categorical syllogisms of Figure II the conditional event existential import is $p(S|(S \vee P)) > 0$. Then, by instantiating S, M, P in Theorem 2 for A, B, C with $x = y = 1$ and $t > 0$ it follows that $z' = \frac{x+yt-1}{t\,x} = 1$. Then,

$$p(M|P) = 1,\ p(\overline{M}|S) = 1,\ \text{and}\ p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) = 1. \tag{12}$$

Therefore, inference (12) is a probabilistically informative version of Camestres.

By instantiating S, M, P in Theorem 3 for A, B, C with $x_1 = x_2 = 1$, $y_1 = .6$, $y_2 = .9$, $t_1 > 0$, and $t_2 = 1$, we obtain $z^* = y_1 = .6$ and $z^{**} = 1$, i.e.,

$$p(M|P) = 1,\ .6 \leqslant p(\overline{M}|S) \leqslant .9,\ \text{and}\ p(S|(S \vee P)) \geqslant t_1 > 0 \Longrightarrow p(\bar{P}|S) \geqslant .6. \tag{13}$$

This can be seen as an extension of Camestres to generalized quantifiers. Specifically, the second premise can be used to represent a generalized quantified statement like *At least most but not all S are not-M* and the conclusion can represent

At least most S are not-P. Of course, the specific values involved in the premises are context dependent (see also [4,26]).

We observe that, by Remark 3, every direct probabilistic interpretation of the other categorical syllogisms of Figure II are probabilistically non-informative without the further probabilistic constraint $p(S|(S \vee P)) > 0$. In what follows, we show how to construct probabilistically informative versions of other categorical syllogisms of Figure II by suitable instantiations of Theorem 2.

Camestrop. From (12) it also follows that

$$p(M|P) = 1,\ p(\overline{M}|S) = 1,\ \text{and}\ p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) > 0, \tag{14}$$

which is a probabilistic informative interpretation of Camestrop (*Every P is M, No S is M, therefore Some S is not P*) under the existential import assumption ($p(S|(S \vee P)) > 0$).

Baroco. By instantiating S, M, P in Theorem 2 for A, B, C with $x = 1, y > 0$ and $t > 0$ it follows that $z' = \frac{x+yt-1}{t\,x} = \frac{1+yt-1}{t} = y > 0$. Then,

$$p(M|P) = 1,\ p(\overline{M}|S) > 0,\ \text{and}\ p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) > 0. \tag{15}$$

Therefore, inference (15) is a probabilistically informative version of Baroco (*Every P is M, Some S is not M, therefore Some S is not P*) under the existential import.

Cesare. The direct probabilistic interpretation of the categorical syllogism "*No P is M, Every S is M, therefore No S is P*" would correspond to infer $p(\bar{P}|S) = 1$ from the premises $p(\overline{M}|P) = 1$ and $p(M|S) = 1$; However, this inference is not probabilistically informative because it is obtained from Camestres when M is replaced by $\overline{M}$. By instantiating S, M, P in Theorem 2 for A, B, C with $x = y = 0$ and $t > 0$ it follows that $z' = \frac{t-x-yt}{t\,(1-x)} = 1$. Then,

$$p(M|P) = 0,\ p(\overline{M}|S) = 0,\ \text{and}\ p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) = 1.$$

or equivalently,

$$p(\overline{M}|P) = 1,\ p(M|S) = 1,\ \text{and}\ p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) = 1. \tag{16}$$

Therefore, inference (16) is a probabilistically informative version of Cesare under the existential import assumption.

By instantiating S, M, P in Theorem 3 for A, B, C with $x_1 = x_2 = 0$, $y_1 = .1$, $y_2 = .4$, $t_1 > 0$, and $t_2 = 1$, we obtain $z^* = 1 - y_2 = .6$ and $z^{**} = 1$; i.e.,

$$p(M|P) = 0,\ .1 \leqslant p(\overline{M}|S) \leqslant .4,\ \text{and}\ p(S|(S \vee P)) \geqslant t_1 > 0 \Longrightarrow p(\bar{P}|S) \geqslant .6,$$

which is equivalent to

$$p(\overline{M}|P) = 1,\ .6 \leqslant p(M|S) \leqslant .9,\ \text{and}\ p(S|(S \vee P)) \geqslant t_1 > 0 \Longrightarrow p(\bar{P}|S) \geqslant .6. \tag{17}$$

Equation (17) is a generalized version of Cesare, where the second premise can represent a generalized quantified statement like *At least most but not all S are M* and the conclusion can represent *At least many S are not-P*.

Cesaro. From (16) it also follows that

$$p(\overline{M}|P) = 1,\ p(M|S) = 1,\ \text{and } p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) > 0, \tag{18}$$

which is a probabilistically informative interpretation of Cesaro (*No P is M, Every S is M, therefore Some S is not P*) under the existential import.

Festino. By instantiating S, M, P in Theorem 2 for A, B, C with $x = 0, y < 1$ and $t > 0$, as $x + yt < t$, it follows that $z' = \frac{t-x-yt}{t\,(1-x)} = \frac{t-yt}{t} > 1 - y > 0$. Then,

$$p(\overline{M}|P) = 1,\ p(M|S) > 0,\ \text{and } p(S|(S \vee P)) > 0 \Longrightarrow p(\bar{P}|S) > 0. \tag{19}$$

Therefore, inference (19) is a probabilistically informative version of Festino (*No P is M, Some S is M, therefore Some S is not P*) under the existential import.

Remark 6. We observe that, traditionally, the conclusions of logically valid categorical syllogisms of Figure II are neither in the form of sentence type I (*some*) nor of A (*every*). In terms of our probability semantics, indeed, this must be the case even if the existential import assumption $p(S|(S \vee P)) > 0$ is made: according to Theorem 2, the upper the bound for the conclusion $p(\bar{P}|S)$ is always 1; thus, neither sentence type I ($p(P|S) > 0$, i.e. $p(\bar{P}|S) < 1$) nor sentence type A ($p(P|S) = 1$, i.e. $p(\bar{P}|S) = 0$) can be validated.

Remark 7. We recall that $p(S) = p(S \wedge (S \vee P)) = p(S|(S \vee P))P(S \vee P)$. Hence, if we assume that $p(S)$ is positive, then $p(S|(S \vee P))$ must be positive too (the converse, however, does not hold). Therefore, the inferences (12)–(19) hold if $p(S|(S \vee P)) > 0$ is replaced $p(S) > 0$. The constraint $p(S) > 0$ can be seen as a stronger version of an existential import assumption compared to the conditional event existential import.

5 Concluding Remarks

In this paper we observed that an existential import assumption is required for the probabilistic validity of syllogisms of Figure II, which we expressed in terms of a probability constraint. Then, we proved probability propagation rules for categorical syllogisms of Figure II. We applied the probability propagation rules to show the validity of the probabilistic versions of the traditional categorical syllogisms of Figure II (i.e., Camestres, Camestrop, Baroco, Cesare, Cesaro, Festino).

We note that, by setting appropriate thresholds, our semantics also allows for generalizing the traditional syllogisms to new ones involving generalized quantifiers (like interpreting *Most S are P* by $p(P|S) \geqslant t$, where t is a given—usually context dependent—threshold like $>.5$). Probabilistic syllogisms are a much more plausible rationality framework for studying categorical syllogisms compared to the traditional syllogisms, as "truly" *all-* and *existentially quantified* statements are hardly ever used in commonsense contexts (even if people mention words like "all", they usually don't mean *all* in a strictly universal sense). Indeed, quantified statements are usually not falsified by one exception (while the universal quantifier $\forall$ does not allow for exceptions) and quantify over more than just *at least one thing* (while the existential quantifier $\exists$ is weak in the sense that it is true when it holds for at least one thing).

Furthermore, as proposed in [15], the basic syllogistic sentence types can also be interpreted as instances of defaults (i.e., (A) by $S \mid\!\sim P$ and (E) by $S \mid\!\sim \bar{P}$) and negated defaults (i.e., (I) by $S \not\mid\!\sim \bar{P}$ and (O) by $S \not\mid\!\sim P$): thus, we can also build a bridge from categorical syllogisms of Figure II to default reasoning. Camestres, for example, has the following form in terms of defaults: (A) $P \mid\!\sim M$, (E) $S \mid\!\sim \overline{M}$, and the existential import $(S \vee P) \not\mid\!\sim \bar{S}$ implies (E) $S \mid\!\sim \bar{P}$. This version of Camestres can serve as a valid inference rule for nonmonotonic reasoning.

We will devote future work to deepen our results and to extend our coherence-based probability semantics to other categorical syllogisms. In particular, we plan to further generalize categorical syllogisms by applying the theory of compounds of conditionals under coherence (see, e.g. [16,17,22]): as shown in the context of conditional syllogisms [18,36,37], this theory allows for managing logical operations on conditionals and iterated conditionals. Iterated conditionals can be used for interpreting categorical syllogisms with statements like *If* S_1 *are* P_1, *then* S_2 *are* P_2 (i.e., $(P_2|S_2)|(P_1|S_1)$).

Acknowledgments. We thank three anonymous reviewers for their useful comments and suggestions.

References

1. Amarger, S., Dubois, D., Prade, H.: Constraint propagation with imprecise conditional probabilities. In: Proceedings of UAI 1991, pp. 26–34. Morgan Kaufmann, Burlington (1991)
2. Amarger, S., Dubois, D., Prade, H.: Handling imprecisely-known conditional probabilities. In: Hand, D. (ed.) AI and Computer Power: the Impact on Statistics, pp. 63–97. Chapman & Hall, London (1994)
3. Baioletti, M., Capotorti, A., Galli, L., Tognoloni, S., Rossi, F., Vantaggi, B.: CkC (Check Coherence package; version e6, November 2016). http://www.dmi.unipg.it/~upkd/paid/software.html
4. Barwise, J., Cooper, R.: Generalized quantifier and natural language. Linguist. Philos. **4**, 159–219 (1981)
5. Biazzo, V., Gilio, A.: A generalization of the fundamental theorem of de Finetti for imprecise conditional probability assessments. Int. J. Approximate Reason. **24**(2–3), 251–272 (2000)
6. Biazzo, V., Gilio, A., Lukasiewicz, T., Sanfilippo, G.: Probabilistic logic under coherence: complexity and algorithms. Ann. Math. Artif. Intell. **45**(1–2), 35–81 (2005)

7. Biazzo, V., Gilio, A., Sanfilippo, G.: Coherent conditional previsions and proper scoring rules. In: Greco, S., Bouchon-Meunier, B., Coletti, G., Fedrizzi, M., Matarazzo, B., Yager, R.R. (eds.) IPMU 2012. CCIS, vol. 300, pp. 146–156. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31724-8_16
8. Capotorti, A., Galli, L., Vantaggi, B.: Locally strong coherence and inference with lower-upper probabilities. Soft Comput. **7**(5), 280–287 (2003)
9. Chater, N., Oaksford, M.: The probability heuristics model of syllogistic reasoning. Cogn. Psychol. **38**, 191–258 (1999)
10. Cohen, A.: Generics, frequency adverbs, and probability. Linguist. Philos. **22**, 221–253 (1999)
11. Coletti, G., Scozzafava, R.: Probabilistic Logic in a Coherent Setting. Kluwer, Dordrecht (2002)
12. Coletti, G., Scozzafava, R., Vantaggi, B.: Possibilistic and probabilistic logic under coherence: default reasoning and system P. Math. Slovaca **65**(4), 863–890 (2015)
13. Dubois, D., Godo, L., López De Màntaras, R., Prade, H.: Qualitative reasoning with imprecise probabilities. J. Intell. Inf. Syst. **2**(4), 319–363 (1993)
14. Gilio, A., Ingrassia, S.: Totally coherent set-valued probability assessments. Kybernetika **34**(1), 3–15 (1998)
15. Gilio, A., Pfeifer, N., Sanfilippo, G.: Transitivity in coherence-based probability logic. J. Appl. Logic **14**, 46–64 (2016)
16. Gilio, A., Sanfilippo, G.: Conditional random quantities and iterated conditioning in the setting of coherence. In: van der Gaag, L.C. (ed.) ECSQARU 2013. LNCS (LNAI), vol. 7958, pp. 218–229. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39091-3_19
17. Gilio, A., Sanfilippo, G.: Generalized logical operations among conditional events. Appl. Intell. (in press). https://doi.org/10.1007/s10489-018-1229-8
18. Gilio, A., Over, D.E., Pfeifer, N., Sanfilippo, G.: Centering and compound conditionals under coherence. In: Ferraro, M.B., et al. (eds.) Soft Methods for Data Science. AISC, vol. 456, pp. 253–260. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-42972-4_32
19. Gilio, A., Pfeifer, N., Sanfilippo, G.: Transitive reasoning with imprecise probabilities. In: Destercke, S., Denoeux, T. (eds.) ECSQARU 2015. LNCS (LNAI), vol. 9161, pp. 95–105. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20807-7_9
20. Gilio, A., Sanfilippo, G.: Probabilistic entailment in the setting of coherence: the role of quasi conjunction and inclusion relation. Int. J. Approximate Reason. **54**(4), 513–525 (2013)
21. Gilio, A., Sanfilippo, G.: Quasi conjunction, quasi disjunction, t-norms and t-conorms: probabilistic aspects. Inf. Sci. **245**, 146–167 (2013)
22. Gilio, A., Sanfilippo, G.: Conditional random quantities and compounds of conditionals. Studia Logica **102**(4), 709–729 (2014)
23. Lambert, J.H.: Neues Organon oder Gedanken über die Erforschung und Bezeichnung des Wahren und dessen Unterscheidung vom Irrthum und Schein. Wendler, Leipzig (1764)
24. Lukasiewicz, T.: Local probabilistic deduction from taxonomic and probabilistic knowledge-bases over conjunctive events. Int. J. Approximate Reason. **21**, 23–61 (1999)
25. Lukasiewicz, T.: Probabilistic deduction with conditional constraints over basic events. J. Artif. Intell. Res. **10**, 199–241 (1999)
26. Peters, S., Westerståhl, D.: Quantifiers in Language and Logic. Oxford University Press, Oxford (2006)

27. Petturiti, D., Vantaggi, B.: Envelopes of conditional probabilities extending a strategy and a prior probability. Int. J. Approximate Reason. **81**, 160–182 (2017)
28. Pfeifer, N.: Contemporary syllogistics: comparative and quantitative syllogisms. In: Kreuzbauer, G., Dorn, G.J.W. (eds.) Argumentation in Theorie und Praxis: Philosophie und Didaktik des Argumentierens, pp. 57–71. Lit Verlag, Wien (2006)
29. Pfeifer, N.: The new psychology of reasoning: a mental probability logical perspective. Thinking Reason. **19**(3–4), 329–345 (2013)
30. Pfeifer, N.: Reasoning about uncertain conditionals. Studia Logica **102**(4), 849–866 (2014)
31. Pfeifer, N., Douven, I.: Formal epistemology and the new paradigm psychology of reasoning. Rev. Philos. Psychol. **5**(2), 199–221 (2014)
32. Pfeifer, N., Kleiter, G.D.: Towards a mental probability logic. Psychol. Belgica **45**(1), 71–99 (2005)
33. Pfeifer, N., Sanfilippo, G.: Probabilistic squares and hexagons of opposition under coherence. Int. J. Approximate Reason. **88**, 282–294 (2017)
34. Pfeifer, N., Sanfilippo, G.: Square of opposition under coherence. In: Ferraro, M.B. (ed.) Soft Methods for Data Science. AISC, vol. 456, pp. 407–414. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-42972-4_50
35. Pfeifer, N., Tulkki, L.: Conditionals, counterfactuals, and rational reasoning. An experimental study on basic principles. Minds Mach. **27**(1), 119–165 (2017)
36. Sanfilippo, G., Pfeifer, N., Gilio, A.: Generalized probabilistic modus ponens. In: Antonucci, A., Cholvy, L., Papini, O. (eds.) ECSQARU 2017. LNCS (LNAI), vol. 10369, pp. 480–490. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61581-3_43
37. Sanfilippo, G., Pfeifer, N., Over, D.E., Gilio, A.: Probabilistic inferences from conjoined to iterated conditionals. Int. J. Approximate Reason. **93**, 103–118 (2018)
38. Martin, T.: J.-H. Lambert's theory of probable syllogisms. Int. J. Approximate Reason. **52**(2), 144–152 (2011)