'Logic and Applications'
September 19-23, 2016, Dubrovnik, Croatia

Soundness and completeness of a sequent calculus with high probabilities

Marija Boričić
Faculty of Organizational Sciences, University of Belgrade
marija.boricic@fon.bg.ac.rs

§0. **Introduction.**

We define a sequent calculus **LKprob**(ε) as a probabilistic extension of classical propositional calculus **LK**, enabling one to manipulate with sequents of the form $\Gamma \vdash^n \Delta$ with the intended meaning that 'the probability of the sequent $\Gamma \vdash \Delta$ belongs to the interval $[1 - n\varepsilon, 1]$' for a given small real $\varepsilon > 0$ of the form $\varepsilon = \frac{1}{k}$ for some fixed $k \in \mathbf{N}$ and any $n \in \mathbf{N}$, $n \leq k$. The system **LKprob**(ε) is sound and complete with respect to a kind of Carnap–Popper–Leblanc–type probability logic semantics.

M. Boričić, *Suppes–style sequent calculus for probability logic,* **J. Log. Comput,** (to appear)

§1. **The probabilistic sequent calculus LKprob(ε).**

The sequent $\Gamma \vdash \Delta$, as introduced by Gentzen, consists of two finite (possibly empty) sequences (or words) of formulae Γ — the antecedent, and Δ — the consequent, with the main interpretation as $\bigwedge \Gamma \rightarrow \bigvee \Delta$, where $\bigwedge \Gamma$ denotes the conjunction of all formulae appearing in Γ, and $\bigvee \Delta$ denotes the disjunction of all formulae appearing in Δ; particularly, if Γ or Δ is an empty sequence, then $\vdash \Delta$ is interpreted as $\bigvee \Delta$, $\Gamma \vdash$ as $\neg \bigwedge \Gamma$, and $\vdash$ can be understood as a pure contradiction. Propositional formulae are defined over propositional language consisting of a denumerable set of propositional letters: $\{p_1, p_2, ...\}$, logical connectives: $\neg$, $\wedge$, $\vee$ and $\rightarrow$, and two auxiliary symbols:) and (. The set of formulae is the smallest set containing propositional letters closed under the following formation rule: if A and B are formulae, then $(\neg A)$, $(A \wedge B)$, $(A \vee B)$ and $(A \rightarrow B)$ are formulae as well.

The axioms of **LKprob**(ε) are the following two forms of sequents:

$$A \vdash^0 A$$
$$\Gamma \vdash^k \Delta,$$

for any words Γ and Δ, and any formula A.

The structural rules of **LKprob**(ε) are as follows:

permutation: $\dfrac{\Gamma A B \Pi \vdash^n \Delta}{\Gamma B A \Pi \vdash^n \Delta}(P\vdash)$ $\quad \dfrac{\Gamma \vdash^n \Delta A B \Lambda}{\Gamma \vdash^n \Delta B A \Lambda}(\vdash P)$

contraction: $\dfrac{\Gamma A A \vdash^n \Delta}{\Gamma A \vdash^n \Delta}(C\vdash)$ $\quad \dfrac{\Gamma \vdash^n A A \Delta}{\Gamma \vdash^n A \Delta}(\vdash C)$

weakening: $\dfrac{\Gamma \vdash^n \Delta}{\Gamma A \vdash^n \Delta}(W\vdash)$ $\quad \dfrac{\Gamma \vdash^n \Delta}{\Gamma \vdash^n A \Delta}(\vdash W)$

the cut rule:

$$\dfrac{\Gamma \vdash^n A\Delta \qquad \Pi A \vdash^m \Lambda}{\Gamma \Pi \vdash^{m+n} \Delta \Lambda}(\text{cut})$$

Some specific structural rules treating monotonicity:

$$\frac{\Gamma \vdash^n \Delta}{\Gamma \vdash^m \Delta}(M \uparrow)$$

for any m $(m \geq n)$, and additivity:

$$\frac{AB \vdash^0 \quad \vdash^m A \quad \vdash^n B}{\vdash^{m+n-k} AB}(ADD)$$

where $k\varepsilon = 1$.

The logical rules of **LKprob**(ε) are as follows:

$$\frac{\Gamma \vdash^n A\Delta}{\Gamma\neg A \vdash^n \Delta}(\neg\vdash) \qquad \frac{\Gamma A \vdash^n \Delta}{\Gamma \vdash^n \neg A\Delta}(\vdash\neg)$$

$$\frac{\Gamma AB \vdash^n \Delta}{\Gamma A\wedge B \vdash^n \Delta}(\wedge\vdash) \qquad \frac{\Gamma \vdash^n A\Delta \quad \Gamma \vdash^m B\Delta}{\Gamma \vdash^{m+n} A\wedge B\Delta}(\vdash\wedge)$$

$$\frac{\Gamma A \vdash^n \Delta \quad \Gamma B \vdash^m \Delta}{\Gamma A\vee B \vdash^{m+n} \Delta}(\vee\vdash) \qquad \frac{\Gamma \vdash^n AB\Delta}{\Gamma \vdash^n A\vee B\Delta}(\vdash\vee)$$

$$\frac{\Gamma \vdash^n A\Delta \quad \Pi B \vdash^m \Lambda}{\Gamma\Pi A\to B \vdash^{m+n} \Delta\Lambda}(\to\vdash) \qquad \frac{\Gamma A \vdash^n B\Delta}{\Gamma \vdash^n A\to B\Delta}(\vdash\to)$$

Example 1. Let the formulas A, B, C, D, E, F and G represent the following symptoms and diagnosis respectively: cough, sniffle, dermal rash, allergy 1, allergy 2, cold and flu. Suppose that we have information about patients during n years in n databases. We will consider a set of hypothesis obtained from the given databases: $B \vdash^5 D$, $BG \vdash^{10} E$, $A \vdash^3 FG$ and $AC \vdash^8 (D \wedge E)F$, where $\varepsilon = 10^{-2}$. For example, $B \vdash^5 D$ means that the probability that a patient has the allergy 1 given that he has sniffles, is at least 0.95, where 0.95 is the lowest probability for that event through the years. Suppose that we are interested in probability of patient having both allergies $(D \wedge E)$ or a cold (F) given that he is coughing (A) and, sneezing or having a dermal rash $(B \vee C)$. Using our system **LKprob**(10^{-2}), we calculate it as follows:

$$\dfrac{\dfrac{\dfrac{B \vdash^5 D}{AB \vdash^5 DF}(W\vdash, \vdash W) \quad \dfrac{A \vdash^3 FG \quad BG \vdash^{10} E}{AB \vdash^{13} FE}(\text{cut})}{AB \vdash^{18} (D \wedge E)F}(\wedge \vdash) \qquad AC \vdash^8 (D \wedge E)F}{A(B \vee C) \vdash^{26} (D \wedge E)F}(\vdash \vee)$$

That means that the probability of patient having both allergies or a cold given that he is coughing and, sneezing or having a dermal rash is at least $1 - 26\varepsilon = 0.74$.

§2. **Models for Probabilized Sequents.**

Let $I = \{1 - n\varepsilon | n \in \mathbf{N}\}$. A *model* for $\mathbf{LKprob}(\varepsilon)$ is a mapping $p : \mathrm{Seq} \to I \cap [0, 1]$ satisfying the following conditions:

(i) $p(A \vdash A) = 1$, for any formula A;

(ii) if $p(AB \vdash) = 1$, then $p(\vdash AB) = p(\vdash A) + p(\vdash B)$, for any formulae A and B;

(iii) if sequents $\Gamma \vdash \Delta$ and $\Pi \vdash \Lambda$ are equivalent in **LK**, in sense that there are proofs for both sequents $\bigwedge \Gamma \to \bigvee \Delta \vdash \bigwedge \Pi \to \bigvee \Lambda$ and $\bigwedge \Pi \to \bigvee \Lambda \vdash \bigwedge \Gamma \to \bigvee \Delta$ in **LK**, then $p(\Gamma \vdash \Delta) = p(\Pi \vdash \Lambda)$.

We say that the sequent $\Gamma \vdash^n \Delta$ is satisfied in a model p, i.e. $\models_p \Gamma \vdash^n \Delta$, if and only if $p(\Gamma \vdash \Delta) \geq 1 - n\varepsilon$. A sequent $\Gamma \vdash^n \Delta$ is *valid* if and only if it is satisfied in each model, denoted by $\models \Gamma \vdash^n \Delta$.

Lemma. *For any formulae A and B the following holds:*
(a) $p(\vdash \neg A) = 1 - p(\vdash A)$*;*
(b) $p(\vdash AB) = p(\vdash A) + p(\vdash B) - p(\vdash A \wedge B)$*;*
(c) $p(\vdash AB) \geq p(\vdash A)$*;*
(d) $p(A \vdash B) \leq p(A \vdash) + p(\vdash B)$*;*
(e) $p(A \vdash) + p(\vdash A) = 1$*.*

Lemma. *The axioms of* $\mathbf{LKprob}(\varepsilon)$ *are valid.*

Lemma. *For every A, B, Γ, Δ, Π and Λ, we have:*

(a) if $p(\Gamma \vdash A\Delta) \geq 1 - n\varepsilon$ and $p(\Gamma \vdash B\Delta) \geq 1 - m\varepsilon$, then

$$p(\Gamma \vdash A \wedge B\Delta) \geq 1 - (m + n)\varepsilon;$$

(b) if $p(\Gamma A \vdash \Delta) \geq 1 - n\varepsilon$ and $p(\Gamma B \vdash \Delta) \geq 1 - m\varepsilon$, then

$$p(\Gamma A \vee B \vdash \Delta) \geq 1 - (m + n)\varepsilon;$$

(c) if $p(\Gamma \vdash A\Delta) \geq 1 - n\varepsilon$ and $p(\Pi B \vdash \Lambda) \geq 1 - m\varepsilon$, then

$$p(\Gamma\Pi A \to B \vdash \Delta\Lambda) \geq 1 - (m + n)\varepsilon.$$

Proof. (c) Suppose that $p(\Gamma \vdash A\Delta) \geq 1 - n\varepsilon$ and $p(\Pi B \vdash \Lambda) \geq 1 - m\varepsilon$, then

$$p(\Gamma\Pi A \to B \vdash \Delta\Lambda) = p(\vdash (A \wedge \neg B)\Delta\Lambda\neg(\bigwedge\Gamma)\neg(\bigwedge\Pi)) =$$

$$= p(\vdash (A \vee \Delta \vee \Lambda \vee \neg(\bigwedge\Gamma) \vee \neg(\bigwedge\Pi)) \wedge (\neg B \vee \Delta \vee \Lambda \vee \neg(\bigwedge\Gamma) \vee \neg(\bigwedge\Pi))) =$$

$$= p(\vdash A\Delta\Lambda\neg(\bigwedge\Gamma)\neg(\bigwedge\Pi)) + p(\vdash \neg B\Delta\Lambda\neg(\bigwedge\Gamma)\neg(\bigwedge\Pi)) -$$

$$-p(\vdash A\neg B\Delta\Lambda\neg(\bigwedge\Gamma)\neg(\bigwedge\Pi)) \geq$$

$$\geq p(\vdash A\Delta\neg(\bigwedge\Gamma)) + p(\vdash \neg B\Lambda\neg(\bigwedge\Pi)) - 1 =$$

$$= p(\Gamma \vdash A\Delta) + p(\Pi B \vdash \Lambda) - 1$$

Therefore, $p(\Gamma\Pi A \to B \vdash \Delta\Lambda) \geq 1 - (m + n)\varepsilon$.□

Lemma. *Let* $p(\vdash A) \geq 1 - i\varepsilon$, $p(\vdash B) \geq 1 - j\varepsilon$, $p(\vdash C) \geq 1 - k\varepsilon$, $p(A \vdash B) \geq 1 - m\varepsilon$ *and* $p(B \vdash C) \geq 1 - n\varepsilon$, *with* $2(x+y)\varepsilon > 1$ *for* $x, y \in \{i, j, k, m, n\}$. *Then:*

(a) [28] *(modus ponens probabilized)*

$$p(\vdash B) \geq 1 - (i + m)\varepsilon$$

(b) [33] *(modus tollens probabilized)*

$$p(A \vdash) \geq 1 - (j + m)\varepsilon$$

(c) [5] *(hypothetical syllogism rule probabilized)*

$$p(A \vdash C) \geq 1 - (j + k)\varepsilon$$

(d) [5] *(hypothetical syllogism rule probabilized)*

$$p(A \vdash C) \geq 1 - (m + n)\varepsilon$$

The bounds in (a), (b), (c) and (d) are the best possible.

§*3.* **The notion of consistency.**

A theory $\mathbf{LKprob}(\varepsilon)(\sigma_1, \ldots, \sigma_n)$ is said to be consistent iff there exists a sequent $\Gamma \vdash^0 \Delta$ which is unprovable in $\mathbf{LKprob}(\varepsilon)(\sigma_1, \ldots, \sigma_n)$. Otherwise, we say that it is inconsistent, and in this case every sequent is provable. A consistent theory is called a maximal consistent theory if each its proper extension is inconsistent.

Lemma. *Each consistent theory can be extended to a maximal consistent theory.*

Proof. Let $\mathcal{T}$ be a consistent theory, and let $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$ be the sequence of all unlabelled sequents, i.e. α_n is $\Gamma_n \vdash \Delta_n$. Let for each $k \in \{0, 1, \ldots, m\}$, where $m = \min\{p | p \in \mathbf{N}$ and $1 - p\varepsilon \leq 0\}$, α_n^k be the sequence of the corresponding labelled sequents, i.e. α_n^k is $\Gamma_n \vdash^k \Delta_n$. We define, for each $n \in \{0, 1, \ldots\}$, a theory $\mathcal{T}_n$ so that: $\mathcal{T}_0 = \mathcal{T}$, and $\mathcal{T}_{n+1} = \mathcal{T}_n \cup \{\alpha_n^0\}$ if it is consistent, else, $\mathcal{T}_{n+1} = \mathcal{T}_n \cup \{\alpha_n^1\}$ if it is consistent, else, $\ldots$, $\mathcal{T}_{n+1} = \mathcal{T}_n \cup \{\alpha_n^{m-1}\}$, else, $\mathcal{T}_{n+1} = \mathcal{T}_n \cup \{\alpha_n^m\}$. Having the previous results in mind, we have that the theory $\mathcal{T}' = \cup_{n \in \mathbf{N}} \mathcal{T}_n$ is consistent as well. Also, theory $\mathcal{T}'$ is a maximal consistent extension of $\mathcal{T}$. Suppose that we extend $\mathcal{T}'$ by the sequent $\Gamma_k \vdash^t \Delta_k$. If the sequent $\Gamma_k \vdash^s \Delta_k$, for some $s < t$, is already in $\mathcal{T}'$, than by monotonicity inference rule $(M \uparrow)$, we infer $\Gamma_k \vdash^t \Delta_k$, which means that $\mathcal{T}' \cup \{\Gamma_k \vdash^t \Delta_k\}$ isn't a proper extension of $\mathcal{T}'$. Else, if any of sequents $\Gamma_k \vdash^s \Delta_k$, for $s < t$, are not in $\mathcal{T}'$, having in mind the definition of $\mathcal{T}'$, there is a sequent $\Gamma_k \vdash^l \Delta_k$ in $\mathcal{T}'$, for some $l > t$. But that is impossible, because that means that the theory $\mathcal{T}_k \cup \{\Gamma_k \vdash^t \Delta_k\}$ is inconsistent, and consequently, $\mathcal{T}' \cup \{\Gamma_k \vdash^t \Delta_k\}$ is inconsistent as well. $\square$

§4. **Soundness and Completeness.**

Soundness and Completeness Theorem. **LKprob**(ε)*–theory has a model iff it is consistent.*

Proof. Soundness is justified immediately by Lemmata 1—5. Completeness part can be justified as follows: Suppose that a theory **LKprob**$(\varepsilon)(\sigma_1, \ldots, \sigma_n)$ is consistent. We construct a model for it as follows. Let CE(**LKprob**$(\sigma_1, \ldots, \sigma_n)$) be the class of all maximal consistent extensions of the set of all **LKprob**$(\varepsilon)(\sigma_1, \ldots, \sigma_n)$ -provable sequents. For any $X \in$ CE(**LKprob**$(\varepsilon)(\sigma_1, \ldots, \sigma_n)$) we define $\models_{p^X} \Gamma \vdash^m \Delta$ iff $1 - m\varepsilon \leq p_X(\Gamma \vdash \Delta)$, where $p_X(\Gamma \vdash \Delta) = 1 - \varepsilon \min\{n | \Gamma \vdash^n \Delta \in X\}$, meaning that $\models_{p^X} \Gamma \vdash^m \Delta$ iff $m \geq \min\{n | \Gamma \vdash^n \Delta \in X\}$. From the previous definition and the monotonicity rule $(M \uparrow)$ we can conclude that $\models_{p^X} \Gamma \vdash^m \Delta$ iff $\Gamma \vdash^m \Delta \in X$, which is a distinctive feature of the canonical model. Also, let us note that the constructed canonical model is a model, because we have: (i) $p(A \vdash A) = 1$, since $A \vdash^0 A$ is an axiom; (ii) if $p(AB \vdash) = 1$, then $p(\vdash AB) = p(\vdash A) + p(\vdash B)$, for any formulae A and B, follows from the additivity rule; (iii) equality $p(\Gamma \vdash \Delta) = p(\Pi \vdash \Lambda)$, where sequents $\Gamma \vdash \Delta$ and $\Pi \vdash \Lambda$ are equivalent in **LK**, holds because of the following. Namely, directly by Lemma 5(a) we have that if $\vdash^n A$ and $A \vdash^0 B$ are provable in **LKprob**(ε), then $\vdash^n B$ is provable in **LKprob**(ε). Finally, since X is a maximal consistent set and every theory can be extended to a maximal consistent theory, there is a model for **LKprob**(ε).$\square$