Probabilistic logics and probabilistic networks[*]
Reading Group Meeting 1: Chapters 1 and 2, pp. 3 – 20.

Version 1.0 (29 October 2019)

Aleksandra Samonek, aleksandra.samonek@uclouvain.be
https://asamonek.github.io/events/2019/10/29/progic.html

Our primary reading
Haenni, R., Romeijn, J. W., Wheeler, G., & Williamson, J. (2010). *Probabilistic logics and probabilistic networks* (Vol. 350). Springer Science & Business Media.

Chapter 1. Introduction

Progic. There are at least two important differences between probabilisitic logic (progic) and other logics:
1. In progic propositions have *probabilities* attached to them. For a proposition φ, we have a progic-proposition of the form φ^X, where $X \subseteq [0,1]$ is the set of probabilities. We spell out φ^X as *the probability of* φ *lies in* X.

2. The fundamental question of progic differs from that of other logics.

Definition (Fundamental question of logic).
Let $\varphi_1, \ldots, \varphi_n, \psi$ *be propositions. We are interested in whether* ψ *follows from the set* $\varphi_1, \ldots, \varphi_n$*, that is:*

$$\varphi_1, \ldots, \varphi_n \approx^? \psi$$

Definition (Fundamental question of progic).
Let $\varphi_1^{X_1}, \ldots, \varphi_n^{X_n}, \psi^Y$ *be progic-propositions. We are interested what set* Y *of probabilities we should attach to* ψ *given* $\varphi_1, \ldots, \varphi_n$*, that is:*

$$\varphi_1^{X_1}, \ldots, \varphi_n^{X_n} \approx \psi^?$$

Progicnet programme. The goal of the progicnet programme is to show how the fundamental question of progic:
– involves a variety of inferential procedures, and
– can be answered using probabilisitic networks.

Notation and basic definitions.

$\mathscr{L}$	a logical language
$A, B, C, \ldots$	*propositional* variables
x, y, z	*logical* variables
U, V, W	predicates
$t, t_1, t_2, \ldots$	constants
$\approx$	generic entailment
$\mid\!\approx$	non-monotonic entailment
$\models\!\!\!\approx$	monotonic entailment
$\models$	decomposable, monotonic entailment
φ, ψ	sentences of the logical language
$\Gamma, \Delta, \Theta, \Phi$	sets of sentences of the logical language
P, Q, R, S	probability functions
$\mathbb{P}$	a set of probability functions
X, Y, Z	sets of probabilities
ζ, η, θ	parameter in a probabilistic model
E, F, G, H	subsets of the outcome space Ω
$\mathscr{E}, \mathscr{F}$	algebras of subsets of Ω
$\omega, \omega_1, \omega_2, \ldots$	elements of the outcome space Ω

[*]This handout has been prepared by Aleksandra Samonek for the purposes of the *Probabilistic logics and probabilistic networks* reading group at UCLouvain in the academic year 2019/2020. Should you find mistakes in it, please report them at aleksandra.samonek@uclouvain.be

Definition (Propositional variable.). *A propositional variable is a variable* A *that takes one of two values, true or false. The assignment of true to a variable* A *is denoted* a *or* a^1*, while the assignment of false is denoted* $\bar{a}$ *or* a^0*.*

Definition (Propositional sentence.). *Given propositional variables* $A_1, \ldots, A_n$ *a propositional language contains sentences built in the standard way from the assignments* $a_1, \ldots, a_n$ *and the logical connectives* $\neg, \wedge, \vee, \rightarrow$ *and* $\leftrightarrow$*.*

Definition (Elementary outcome ω.). *An elementary outcome* ω *is an assignment* $a_1^{e_1}, \ldots, a_n^{e_n}$*, where* $e_1, \ldots, e_n \in \{0,1\}$.

Definition (Atomic state.). *A atomic state* α *is a sentence that denotes an elementary outcome.* α *is a conjuction* $\pm a_1 \wedge \ldots \wedge \pm a_n$*, where* $\pm a_i$ *is of the form* a_i *if* $e_i = 1$ *in the elementary outcome and* $\neg a_i$ *otherwise. More generally, let* $e = e_1, \ldots, e_n$*. Then let* α^e *denotes the atomic state describing the elementary outcome* $a_1^{e_1}, \ldots, a_n^{e_n}$*.*

Definition (Propositional language.). *Expressions of the form* $U(t)$ *and* $V(x)$ *determine (single-case and, respectively, repeatably-instantiatable) propositional variables.* Ut *and* Vx *denote the positive assignments* $U(t) = true$ *and* $V(x) = true$*. Finitely many such expressions yield a propositional language.*

Definition (Atomic state of propositional literals.). *The letter* α *denotes an atomic state or state description, which is a conjunction of atomic literals, e. g.,* $Ut_1 \wedge \neg Vt_2 \wedge \neg Wt_3$*, where each predicate in the language (or in a given finite sublanguage) features in* α*.*

Definition (Credal set $\mathbb{K}$.). *A credal set* $\mathbb{K}$ *is a closed convex set of probability functions.*

Chapter 2. Standard Probabilistic Semantics

We will use the following notion underlying the standard (probabilistic) semantics to compare different interpretations of the fundamental question of progic.

Namely, according to the standard semantics, an entailment relation $\varphi_1^{X_1}, \ldots, \varphi_n^{X_n} \approx \psi^Y$ holds if all probability functions which satisfy the constraints imposed by the left-hand side also satisfy the right-hand side of the entailment.

Kolgomorov probabilities.

Definition (Probability space.). *A probability space is a triple* $(\Omega, \mathscr{F}, P)$*, where* Ω *is a sample space of elementary events,* $\mathscr{F}$ *is a σ-algebra of subsets of* Ω*, and* $P : \mathscr{F} \to [0,1]$ *is a probability measure satisfying the Kolgomorov axioms:*
(P1) $P(E) \geqslant 0$*, for all* $E \in \mathscr{F}$*;*
(P2) $P(\Omega) = 1$*;*
(P3) $P(E_1 \cup E_2 \cup \ldots) = \Sigma_i P(E_i)$*, for any countable sequence* $E_1, E_2, \ldots$ *of pairwise disjoint events* $E_i \in \mathscr{F}$*.*

Definition (Probability structure.). *A probability structure is a quadruple* $M = (\Omega, \mathscr{F}, P, I)$*, where* $(\Omega, \mathscr{F}, P)$ *is a probability space and* I *is an interpretation function associating each elementary event* $\omega \in \Omega$ *with a truth assignment on the propositional variables* Φ *in a language* $\mathscr{L}$ *such that* $I(\omega, A) \in \{true, false\}$ *for each* $\omega \in \Omega$ *and for every* $A, B, C, \cdots \in \Phi$*.*

How to relate events to formulas? Since P is defined on events, in order to link events in M to formulas in Ω, we associate $[\![\varphi]\!]_M$ with the set of elementary events within (finite) Ω in M where φ is true.

Proposition. *For arbitrary propositional formulas φ, ψ:*
1. $[\![\varphi \wedge \psi]\!]_M = [\![\varphi]\!]_M \cap [\![\psi]\!]_M$,
2. $[\![\varphi \vee \psi]\!]_M = [\![\varphi]\!]_M \cup [\![\psi]\!]_M$,
3. $[\![\neg\varphi]\!]_M = \Omega \setminus [\![\varphi]\!]_M$.

Under this interpretation, the assignments of probability to sets in the algebra are effectively the assignments of probability to expressions in $\mathscr{L}$, so we have that $P(\varphi \Leftrightarrow P([\![\varphi]\!]_M)$.
Notation. We may compress notation and use capital letters for both, propositions and events. Moreover, we will omit M in the subscript when the context is clear.

Interval-valued probabilities.

Proposition (Wheeler, 2006). *If $P(E)$ and $P(F)$ are defined in M, then:*
1. *$P(E \cap F)$ lies within the interval* $[max(0, (P(E) + P(F)) - 1), min(P(E), P(F))]$, *and*
2. *$P(E \cup F)$ lies within the interval* $[max(P(E), P(F)), min((P(E) + P(F), 1)]$.

Here the premises are seen as restrictions on a set of probability assignments. But how to extend the use of interval-valued assignments to premises? One of the strategies is to rely on assignments by means of inner and outer measure.

Example. Let $F \notin \mathscr{F}$ but also let F be such that F is logically related to events in $\mathscr{F}$. For example, let $E \subset F$ and $F \subset G$ and $E, G \in \mathscr{F}$. We do not have a sharp probability value for F, because F is outside our probability structure $\mathscr{F}$.

Call E a kernel event for F if there does not exist in $\mathscr{F}$ a measurable event that dominates E. Call G a covering event for F if every measurable event containing F dominates G. We obtain non-trivial bounds on F with respect to M by taking the measures of F's kernel and cover (otherwise $P(F)$ would be undefined).

If P is defined on $\mathscr{F}$ of M and E' is not in $\mathscr{F}$, then $P(E')$ is not defined since E' is not in the domain of P. However, E' may be an element of an algebra $\mathscr{F}'$ such that $\mathscr{F}$ is a subalgebra of $\mathscr{F}'$. Then we may extend the measure P to the set E' by defining *inner and other measures* to represent our uncertainty with respect to the precise measure of E'.

Definition (Inner and Outer Measure.). *Let $\mathscr{F}$ be a subalgebra of an algebra $\mathscr{F}'$[1] ; let $P : \mathscr{F} \to [0, 1]$ be a probability measure defined on the space $(\Omega, \mathscr{F}, P)$ and E be an arbitrary set in $\mathscr{F}'$. Then define an inner measure (IM) $\underline{P}$ induced by P and the outer measure (OM) $\overline{P}$ induced by P as:*

$\underline{P}(E) = sup\{P(F) : F \subseteq E, F \in \mathscr{F}\}$ *(IM of E);*

$\overline{P}(E) = inf\{P(F) : F \supseteq E, F \in \mathscr{F}\}$ *(OM of E).*

Properties of IMs and OMs.

(P4) $\underline{P}(E \cup F) \geq \underline{P}(E) + \underline{P}(F)$, when E and F are disjoint (superadditivity);

(P5) $\underline{P}(E \cup F) \leq \overline{P}(E) + \overline{P}(F)$, when E and F are disjoint (subadditivity);

(P6) $\underline{P}(E) = 1 - \overline{P}(E)$;

(P7) $\underline{P}(E) = \overline{P}(E) = P(E)$ if $E \in \mathscr{F}$.

(P4) is generalized by (P2) and (P3) in the following way:

(P4') $\underline{P}(E \cup F) \geq \underline{P}(E) + \underline{P}(F) - \underline{P}(E \cap F)$, when E and F are disjoint (generalized superadditivity);

[1] There is a mistake in the definition in the original text, which says that $\mathscr{F}'$ is a subalgebra of $\mathscr{F}$. Note that if it were so, there would be no point in defining inner and outer measure, as P would range over $\mathscr{F}'$ and all events in $\mathscr{F}'$ would also be in $\mathscr{F}$ and thus would have strict measures.

A positive function which satisfies (P2) and (P4') is called *2-monotone Choquet capacity*. It can be generalized to an *n-monotone Choquet capacity*.

Imprecise probabilities. The following theorem concerns the relationship between IMs and sets of probabilities and it links the IM of an event E to the lower probability $\underline{\mathbb{P}}(E)$ for a particular set of probability measures, namely those which extend $\mathscr{F}'$ to events in $\mathscr{F}$.

Theorem (Horn and Tarski, 1948.). *Suppose P is a measure on a (finitely additive) probability structure $\mathscr{M}$ such that $\mathscr{F} \subseteq \mathscr{F}'$. Define $\mathbb{P}$ as the set of all extensions P' of P to $\mathscr{F}'$. Then for all $E \in \mathscr{F}'$:*
1. $\underline{P}(E) = \underline{\mathbb{P}}(E) = inf\{P'(E) : P' \in \mathbb{P}\}$, *and*
2. $\overline{P}(E) = \overline{\mathbb{P}}(E) = sup\{P'(E) : P' \in \mathbb{P}\}$.

Convexity.

Definition (Convex set.). *A set X is called convex iff X is closed under all binary operations $b_{(\lambda,1-\lambda)}$ for $\lambda \in [0, 1]$.*

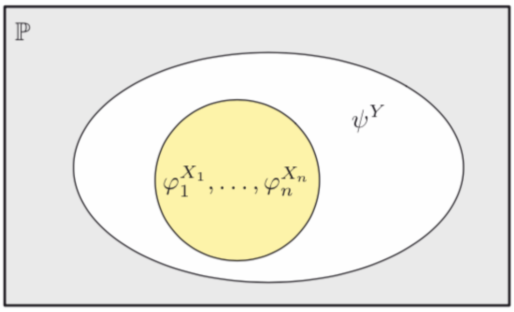
Definition (Credal set.). *A credal set $\mathbb{K}$ is a convex set of probability functions. That is $P_1, P2 \in \mathbb{K}$ implies $\lambda P_1 + (1 - \lambda)P_2 \in \mathbb{K}$.*

See: (Levi, 1983) on convex bayesianism.

The following theorem relates a convex set of probability functions to lower probability.

Theorem (Walley, 1991). *If $\mathbb{K}$ is a convex set of probability functions, then*
1. $\underline{P}(E) = \underline{\mathbb{K}}(E) = inf\{P(E) : P \in \mathbb{K}\}$, *and*
2. $\overline{P}(E) = \overline{\mathbb{K}}(E) = sup\{P(E) : P \in \mathbb{K}\}$.

Representation. According to the standard semantics an inference in valid iff the set of all models (probability assignments) satisfying the premises is included in the set of all models satisfying the conclusion (which in turn is always included in the set of all probability assignments over a given language.)
This fact is represented by Fig. 2.1, p. 18, quoted below.



Supplementary literature.

Levi, I. (1983). *The enterprise of knowledge: An essay on knowledge, credal probability, and chance.* MIT Press.