Probabilized Unranked Sequent Calculus

Mikheil Rukhaia

joint work with: Merium Bishara, Lia Kurtanidze, Lali Tibua

Logic and Applications, Dubrovnik, Croatia

September 24, 2024

Overview

Motivation

- Artificial Intelligence is developing using logical and stataistical methods.
- In early days of AI, logical and probabilistic methods have been used independently.
- Modern AI is area where computer science, mathematics (probability theory, numerical analysis), and logic come together.

Motivation

- Nowadays, researchers started combining logical and probabilistic methods in a single framework and developed several formalisms and programming tools.
- Applications: web mining, natural language processing, robotics, transportation systems, medicine, electronic games, activity recognition, etc.

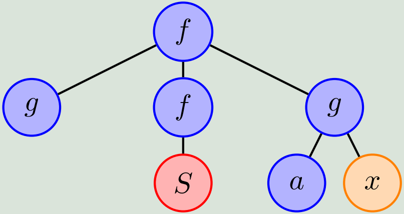
Motivation

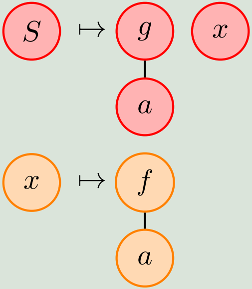
- All probabilistic logic formalisms studied so far are either propositional, or permit only individual variables that can be instantiated by a single term.
- On the other hand, there are very useful theories of symbolic logic, which are using sequence variables and unranked function symbols.
- Applications: in XML modelling, automated reasoning, knowledge representation, etc.

Intuition behind individual (x) and sequence variables (S)

Example

$f(g, f(S), g(a, x))$ $\{S \mapsto (g(a), x),\ x \mapsto f(a)\}$

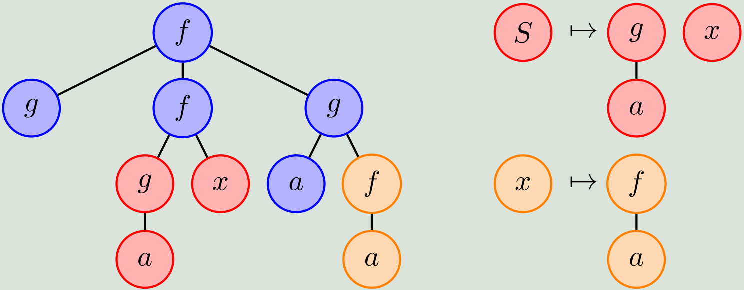




Intuition behind individual (x) and sequence variables (S)

Example

$f(g, f(g(a), x), g(a, f(a)))$ $\{S \mapsto (g(a), x),\ x \mapsto f(a)\}$



Step forward

- Develop probabilized first-order sequent calculus with sequence variables and unranked function symbols.
- So far we are knowledge-oriented and did not put much emphasis on the applications.
- In the future, we believe to find a good application in medical diagnoses.

Syntax

- We use the follwoing signature:
  - Individual variables: $x^i, y^i, z^i, \ldots,$
  - Sequence variables: $\overline{x}, \overline{y}, \overline{z}, \ldots,$
  - Ranked individual function symbols: $f^r, g^r, \ldots,$
  - Unranked individual function symbols: $f^u, g^u, \ldots,$
  - Ranked sequence function symbols: $\overline{f}^r, \overline{g}^r, \ldots,$
  - Unranked sequence function symbols: $\overline{f}^u, \overline{g}^u, \ldots,$
  - Ranked predicate symbols: $p^r, q^r, \ldots,$
  - Unranked predicate symbols: $p^u, q^u, \ldots,$
  - logical connectives and quantifiers: $\neg, \wedge, \vee, \rightarrow, \exists, \forall$.
  - Unary probability operators $P_{[a,b]}$, for every $[a, b] \subseteq [0, 1]$.

Syntax

- Terms:

$$t ::= x^i \mid f^r(t_1, \ldots, t_n) \mid f^u(\overline{s}_1, \ldots, \overline{s}_n)$$
$$\overline{s} ::= t \mid \overline{x} \mid \overline{f}^r(t_1, \ldots, t_n) \mid \overline{f}^u(\overline{s}_1, \ldots, \overline{s}_n)$$

- Atoms: $p^r(t_1, \ldots, t_n)$ and $p^u(\overline{s}_1, \ldots, \overline{s}_n)$
- Formulas: built from atoms, unary probability operators, logical connectives and quantifiers.
- Quantifications are allowed on both, individual and sequence variables.

Lottery Paradox

- Lottery Paradox: for each ticket the winning chance is very low, but there is a high chance that some tickets win.

Example

$$(\forall x^i) P_{[0,0.000001]} Win(x^i) \wedge P_{[0.999999,1]} (\exists \overline{x}) Win(\overline{x})$$

Semantics

- Model is a structure $\mathbf{M} = \langle W, D, I, Prob \rangle$, where:
  - W is a nonempty set of worlds,
  - $D = D_i \cup D_s$ is a domain for every world $w \in W$ and $D_i \neq \emptyset$,
  - I is an interpretation that for each $w \in W$, associates $I(w)$ for every symbol to the corresponding element or function:
    - Example: unranked sequence function symbol $\overline{f}^u$ to a flexible arity multi-valued function $I(w)(\overline{f}^u) \colon D^\infty \to D^\infty$
  - $Prob$ is a probability assignment that for every world $w \in W$ assigns a probability space $Prob(w) = \langle W(w), H(w), \mu(w) \rangle$.

Semantics

- $W(w) \subseteq W$ and $W(w) \neq \emptyset$,
- $H(w)$ is an algebra over subsets of $W(w)$:
  - $W(w) \in H(w)$, and
  - if $\alpha, \beta \in H(w)$, then $W(w) \setminus \alpha \in H(w)$ and $\alpha \cup \beta \in H(w)$
- $\mu(w)\colon H(w) \to [0, 1]$ is a finitely additive probability measure:
  - $\mu(w)(W(w)) = 1$, and
  - if $\alpha \cap \beta = \emptyset$, then $\mu(w)(\alpha \cup \beta) = \mu(w)(\alpha) + \mu(w)(\beta)$.

Semantics

- $\mathbf{M}$-evaluation is a mapping e, which assigns to each individual variable x^i an element $e(x^i) \in D_i$ and to each sequence variable $\overline{x}$ an element $e(\overline{x}) \in D^\infty$.
- Term evaluation:
  - $\|I(w)(x)\|_e^{\mathbf{M}} = e(x)$, for x being individual or sequence variable,
  - $\|I(w)(f(t_1, \dots, t_n))\|_e^{\mathbf{M}} = I(w)(f)(\|t_1\|_e^{\mathbf{M}}, \dots, \|t_n\|_e^{\mathbf{M}})$, for $f \in \{f^r, f^u\}$ and $t_1, \dots, t_n$ being individual terms,
  - $\|I(w)(f(\overline{s}_1, \dots, \overline{s}_n))\|_e^{\mathbf{M}} = I(w)(f)(\|\overline{s}_1\|_e^{\mathbf{M}}, \dots, \|\overline{s}_n\|_e^{\mathbf{M}})$, for $f \in \{\overline{f}^r, \overline{f}^u\}$ and $\overline{s}_1, \dots, \overline{s}_n$ being sequence terms.

Semantics

- Formula evaluation:
  - $\|I(w)(p^r(t_1,\dots,t_n))\|_e^{\mathbf{M}} = \mathit{True}$, if $\langle \|t_1\|_e^{\mathbf{M}},\dots,\|t_n\|_e^{\mathbf{M}}\rangle \in I(w)(p^r)$,
  - $\|I(w)(p^u(\overline{s}_1,\dots,\overline{s}_n))\|_e^{\mathbf{M}} = \mathit{True}$, if $\langle \|\overline{s}_1\|_e^{\mathbf{M}},\dots,\|\overline{s}_n\|_e^{\mathbf{M}}\rangle \in I(w)(p^u)$,
  - $\|I(w)(P_{[a,b]}A)\|_e^{\mathbf{M}} = \mathit{True}$, if $\mu(w)(\{u \mid u \in W(w) \text{ and } \|I(u)(A)\|_e^{\mathbf{M}} = \mathit{True}\}) \in [a,b]$,

Semantics

- Formula evaluation:
  - $\|I(w)(\neg A)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{False}$,
  - $\|I(w)(A \wedge B)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{True}$ and $\|I(w)(B)\|_e^{\mathbf{M}} = \mathit{True}$,
  - $\|I(w)(A \vee B)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{True}$ or $\|I(w)(B)\|_e^{\mathbf{M}} = \mathit{True}$,
  - $\|I(w)(A \rightarrow B)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{False}$ or $\|I(w)(B)\|_e^{\mathbf{M}} = \mathit{True}$,

Semantics

- Formula evaluation:
  - $\|I(w)((\forall x^i)A)\|_e^{\mathbf{M}} = \mathit{True}$, if for every $d \in D_i$, $\|I(w)(A)\|_{e[x^i \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
  - $\|I(w)((\forall \overline{x})A)\|_e^{\mathbf{M}} = \mathit{True}$, if for every $d \in D^\infty$, $\|I(w)(A)\|_{e[\overline{x} \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
  - $\|I(w)((\exists x^i)A)\|_e^{\mathbf{M}} = \mathit{True}$, if for some $d \in D_i$, $\|I(w)(A)\|_{e[x^i \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
  - $\|I(w)((\exists \overline{x})A)\|_e^{\mathbf{M}} = \mathit{True}$, if for some $d \in D^\infty$, $\|I(w)(A)\|_{e[\overline{x} \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
  - in all other cases $I(w)(A) = \mathit{False}$

Semantics

- $\mathbf{M}, w \models A$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{True}$ for every valuation e.
- $\mathbf{M} \models A$, if $\mathbf{M}, w \models A$ for every $w \in W$.
- $\models A$, if $\mathbf{M} \models A$ for every model $\mathbf{M}$.
- $T \models A$, if in every model $\mathbf{M}$, where $\mathbf{M} \models A_j$ for all $A_j \in T$, also $\mathbf{M} \models A$.

Probabilized Sequent

- A probabilized sequent is an expression $\Gamma \vdash_a^b \Delta$, for $[a, b] \subseteq [0, 1]$.
- Intended meaning of a probabilized sequent $\Gamma \vdash_a^b \Delta$ is $P_{[a,b]}(\bigwedge \Gamma \to \bigvee \Delta)$.
- If there is a case when $a > b$ or $a, b \notin [0, 1]$, then we write $\Gamma \vdash^{\emptyset} \Delta$ and treat it as a contradiction.

Inference Rules of **Gp**

- *The axioms*

$$\frac{}{\Gamma \vdash^1_0 \Delta}\textsf{Ax} \qquad \frac{}{A \vdash^1_1 A}\textsf{Ax1} \qquad \frac{}{\vdash^0_0}\textsf{Ax0}$$

- *The cut rule*

$$\frac{\Gamma \vdash^b_a \Delta, A \quad A, \Pi \vdash^d_c \Lambda}{\Gamma, \Pi \vdash^{\mathtt{min}(b+d,1)}_{\mathtt{max}(0,a+c-1)} \Delta, \Lambda}\textsf{cut}$$

Inference Rules of **Gp**

- *The propositional rules*

$$\frac{\Gamma \vdash_a^b \Delta, A}{\neg A, \Gamma \vdash_a^b \Delta}\neg:\mathsf{l} \qquad \frac{A, \Gamma \vdash_a^b \Delta}{\Gamma \vdash_a^b \Delta, \neg A}\neg:\mathsf{r}$$

$$\frac{A, B, \Gamma \vdash_a^b \Delta}{A \wedge B, \Gamma \vdash_a^b \Delta}\wedge:\mathsf{l} \qquad \frac{\Gamma \vdash_a^b \Delta, A \quad \Gamma \vdash_c^d \Delta, B}{\Gamma \vdash_{\mathtt{max}(0,a+c-1)}^{\mathtt{min}(b,d)} \Delta, A \wedge B}\wedge:\mathsf{r}$$

$$\frac{A, \Gamma \vdash_a^b \Delta \quad B, \Gamma \vdash_c^d \Delta}{A \vee B, \Gamma \vdash_{\mathtt{max}(0,a+c-1)}^{\mathtt{min}(b,d)} \Delta}\vee:\mathsf{l} \qquad \frac{\Gamma \vdash_a^b \Delta, A, B}{\Gamma \vdash_a^b \Delta, A \vee B}\vee:\mathsf{r}$$

$$\frac{\Gamma \vdash_a^b \Delta, A \quad B, \Gamma \vdash_c^d \Delta}{A \rightarrow B, \Gamma \vdash_{\mathtt{max}(0,a+c-1)}^{\mathtt{min}(b,d)} \Delta}\rightarrow:\mathsf{l} \qquad \frac{A, \Gamma \vdash_a^b \Delta, B}{\Gamma \vdash_a^b \Delta, A \rightarrow B}\rightarrow:\mathsf{r}$$

Inference Rules of **Gp**

- *The quantifier rules*

$$\frac{A(t), \Gamma \vdash_a^b \Delta}{(\forall x^i) A(x^i), \Gamma \vdash_a^b \Delta} \forall^{\mathsf{i}} : \mathsf{l} \qquad \frac{\Gamma \vdash_a^b \Delta, A(y^i)}{\Gamma \vdash_a^b \Delta, (\forall x^i) A(x^i)} \forall^{\mathsf{i}} : \mathsf{r}$$

$$\frac{A(y^i), \Gamma \vdash_a^b \Delta}{(\exists x^i) A(x^i), \Gamma \vdash_a^b \Delta} \exists^{\mathsf{i}} : \mathsf{l} \qquad \frac{\Gamma \vdash_a^b \Delta, A(t)}{\Gamma \vdash_a^b \Delta, (\exists x^i) A(x^i)} \exists^{\mathsf{i}} : \mathsf{r}$$

$$\frac{A(\overline{s}), \Gamma \vdash_a^b \Delta}{(\forall \overline{x}) A(\overline{x}), \Gamma \vdash_a^b \Delta} \forall : \mathsf{l} \qquad \frac{\Gamma \vdash_a^b \Delta, A(\overline{y})}{\Gamma \vdash_a^b \Delta, (\forall \overline{x}) A(\overline{x})} \forall : \mathsf{r}$$

$$\frac{A(\overline{y}), \Gamma \vdash_a^b \Delta}{(\exists \overline{x}) A(\overline{x}), \Gamma \vdash_a^b \Delta} \exists : \mathsf{l} \qquad \frac{\Gamma \vdash_a^b \Delta, A(\overline{s})}{\Gamma \vdash_a^b \Delta, (\exists \overline{x}) A(\overline{x})} \exists : \mathsf{r}$$

Inference Rules of **Gp**

- *The structural rules*

$$\frac{A, A, \Gamma \vdash_a^b \Delta}{A, \Gamma \vdash_a^b \Delta}\text{c: l} \qquad \frac{\Gamma \vdash_a^b \Delta, A, A}{\Gamma \vdash_a^b \Delta, A}\text{c: r}$$

$$\frac{\Gamma \vdash_a^b \Delta \quad \vdash_c^d A}{A, \Gamma \vdash_{\mathtt{max}(a,1-d)}^{\mathtt{min}(1,b+1-c)} \Delta}\text{w: l} \qquad \frac{\Gamma \vdash_a^b \Delta \quad \vdash_c^d A}{\Gamma \vdash_{\mathtt{max}(a,c)}^{\mathtt{min}(1,b+d)} \Delta, A}\text{w: r}$$

$$\frac{\Gamma \vdash_a^b \Delta}{\Gamma \vdash_c^d \Delta}\text{M}\uparrow \qquad \frac{\Gamma \vdash_a^b \Delta \quad \Gamma \vdash_c^d \Delta}{\Gamma \vdash_{\mathtt{max}(a,c)}^{\mathtt{min}(b,d)} \Delta}\text{M}\downarrow$$

$$\frac{A, B \vdash_1^1 \quad \vdash_a^b A \quad \vdash_c^d B}{\vdash_{\mathtt{max}(1,a+c)}^{\mathtt{min}(1,b+d)} A, B}\text{Add} \qquad \frac{\Pi \vdash^{\emptyset} \Lambda}{\Gamma \vdash^{\emptyset} \Delta}\bot$$

Example

- Assume an ML algorithm learns $P_{[0.92,0.92]}(p^u(a) \to p^u(b,c))$ and $P_{[0.87,0.95]}\forall\overline{x}\exists y(p^u(\overline{x}) \to q^r(y))$ and decides for $\exists y(p^u(a) \to q^r(y))$.
- The explanation can be:

$$\dfrac{\vdash^{0.95}_{0.87} A \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p^u(a) \vdash^{0.92}_{0.92} p^u(b,c) \qquad q^r(v) \vdash^{1}_{1} q^r(v)}{p^u(a), p^u(b,c) \to q^r(v) \vdash^{0.92}_{0.92} q^r(v)} \to: \mathsf{l}}{p^u(b,c) \to q^r(v) \vdash^{0.92}_{0.92} p^u(a) \to q^r(v)} \to: \mathsf{r}}{p^u(b,c) \to q^r(v) \vdash^{0.92}_{0.92} \exists y(p^u(a) \to q^r(y))} \exists^{\mathsf{i}}: \mathsf{r}}{\exists y(p^u(b,c) \to q^r(y)) \vdash^{0.92}_{0.92} \exists y(p^u(a) \to q^r(y))} \exists^{\mathsf{i}}: \mathsf{l}}{\forall\overline{x}\exists y(p^u(\overline{x}) \to q^r(y)) \vdash^{0.92}_{0.92} \exists y(p^u(a) \to q^r(y))} \forall: \mathsf{l}}{\vdash^{1}_{0.79} \exists y(p^u(a) \to q^r(y))} \mathsf{cut}$$

Properties

Theorem (Correspondence)

A sequent $\Gamma \vdash_1^1 \Delta$ *is provable in* ***Gp*** *iff* $\Gamma \vdash \Delta$ *is provable in* ***G***.

Theorem (Soundness & Completeness)

A sequent $\Gamma \vdash_a^b \Delta$ *is provable in* ***Gp*** *iff* $P_{[a,b]}(\bigwedge \Gamma \rightarrow \bigvee \Delta)$ *is valid.*

Summary

- The inference system **G** (by T.Kutsia and B.Buchberger) is an extension of the standard first-order **LK** calculus with the additional $\forall$ and $\exists$ quantifier rules over sequence variables.
- We probabilize the system **G**, in a similar way, as M.Boričić probabilized classical propositional sequent calculus.
- New system keeps properties like soundness and completeness.
- Future work: find real-world example; study cut-elimination.

Acknowledgement

- This work was supported by the Shota Rustaveli National Science Foundation of Georgia under the project №FR-22-4254.

References

- Kutsia, T., Buchberger, B.:
  Predicate logic with sequence variables and sequence function symbols.
  International Conference on Mathematical Knowledge Management, pp.205-219, Springer, 2004.

- Boričić, M.:
  Sequent calculus for classical logic probabilized.
  Archive for Mathematical Logic 58, pp.119-136, 2019.

- Ognjanovic, Z., Raškovic, M.:
  Some first-order probability logics.
  Theoretical Computer Science, 247(1-2), pp.191-212, 2000.

Questions?