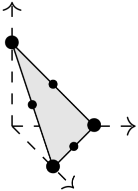
Tropical Geometry of Probabilistic Programming Languages

Davide Barbarossa, Paolo Pistone

db2437@bath.ac.uk

https://davidebarbarossa12.github.io/



POPL26, Rennes

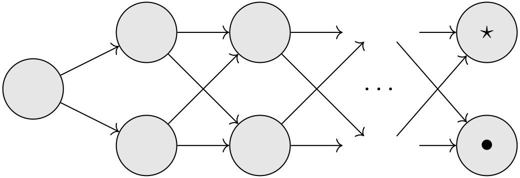
16 January, 2026

Outline

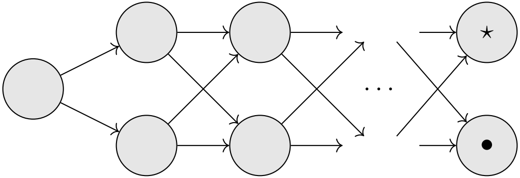
Outline

Probabilistic Models:

Probabilistic Models: compact representation of *very large* event spaces (e.g. Bayesian networks, Hidden Markov models)

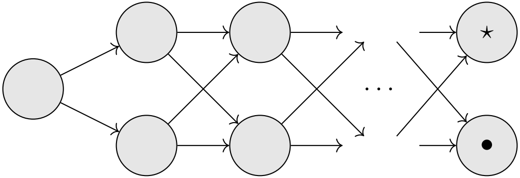


Probabilistic Models: compact representation of *very large* event spaces (e.g. Bayesian networks, Hidden Markov models)



2^n possible trajectories $\alpha_1\alpha_2\ldots\alpha_n \in \{0,1\}^n$

Probabilistic Models: compact representation of *very large* event spaces (e.g. Bayesian networks, Hidden Markov models)



2^n possible trajectories $\alpha_1\alpha_2\ldots\alpha_n \in \{0,1\}^n$

Statistical inference:
Probability of reaching $\star$?
Trajectories with highest probability?

Probabilistic Languages:

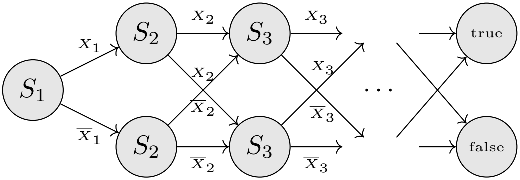
Probabilistic Languages:
compact and user-friendly specification of models via *programs*

Probabilistic Languages:
compact and user-friendly specification of models via *programs*

$$M := (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x) \dots (\lambda x.x \oplus_{X_{n-1}} x)(\text{true} \oplus_{X_n} \text{false})$$

Probabilistic Languages:
compact and user-friendly specification of models via *programs*

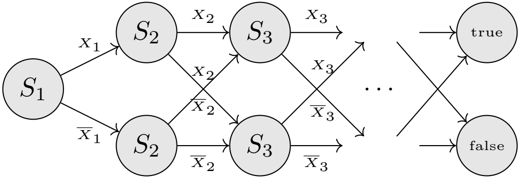
$$M := (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\ldots(\lambda x.x \oplus_{X_{n-1}} x)(\text{true} \oplus_{X_n} \text{false})$$



Probabilistic Languages:

compact and user-friendly specification of models via *programs*

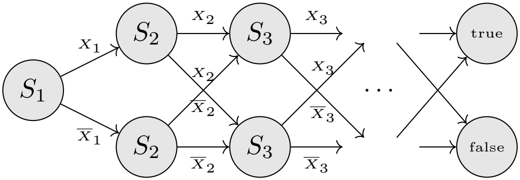
$$M := (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x) \dots (\lambda x.x \oplus_{X_{n-1}} x)(\text{true} \oplus_{X_n} \text{false})$$



Statistical inference → computational behavior of M:

Probabilistic Languages:
compact and user-friendly specification of models via *programs*

$$M := (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\ldots(\lambda x.x \oplus_{X_{n-1}} x)(\text{true} \oplus_{X_n} \text{false})$$



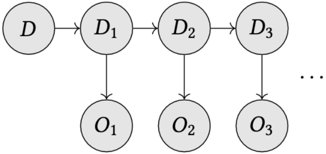
Statistical inference → computational behavior of M:

Probability that $M \twoheadrightarrow$ true?
Reductions $M \twoheadrightarrow$ true of highest probability?

Probabilistic Languages:
compact and user-friendly specification of models via *programs*

$$M := (\text{fix}\, B)(N D) : \mathit{Bool}, \text{ where } B \text{ is}$$

$$\lambda f x.\, \text{ifz}(x, \text{ifz}(O0, f(N0), 1), \text{ifz}(O1, f(N1), 1)) : (\mathit{Bool} \to \mathit{Bool}) \to \mathit{Bool} \to \mathit{Bool}$$



Statistical inference $\to$ computational behavior of M:

Probability that $M \twoheadrightarrow$ true?
Reductions $M \twoheadrightarrow$ true of highest probability?

Outline

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x) \dots (\lambda x.x \oplus_{X_n} x)\text{true}$$

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\dots(\lambda x.x \oplus_{X_n} x)\text{true}$$

Probability that $M \twoheadrightarrow$ true: sum over all reductions ($X^0 = X, X^1 = \overline{X}$):

$$\sum_{\alpha_1,\dots,\alpha_n \in \{0,1\}} X_1^{\alpha_1} \cdots X_n^{\alpha_n} = (X_1 + \overline{X}_1)(X_2 + \overline{X}_2)\cdots(X_n + \overline{X}_n) = 1$$

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x) \dots (\lambda x.x \oplus_{X_n} x)\text{true}$$

Probability that $M \twoheadrightarrow$ true: sum over all reductions ($X^0 = X, X^1 = \overline{X}$):

$$\sum_{\alpha_1,\dots,\alpha_n \in \{0,1\}} X_1^{\alpha_1} \cdots X_n^{\alpha_n} = (X_1 + \overline{X}_1)(X_2 + \overline{X}_2) \cdots (X_n + \overline{X}_n) = 1$$

Maximum a posteriori probability:

$$\max_{\alpha_1,\dots,\alpha_n \in \{0,1\}} \{X_1^{\alpha_1} \cdots X_n^{\alpha_1}\}$$

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\ldots(\lambda x.x \oplus_{X_n} x)\mathrm{true}$$

Probability that $M \to$ true: sum over all reductions ($X^0 = X, X^1 = \overline{X}$):

$$\sum_{\alpha_1,\ldots,\alpha_n \in \{0,1\}} X_1^{\alpha_1} \cdots X_n^{\alpha_n} = (X_1 + \overline{X}_1)(X_2 + \overline{X}_2)\cdots(X_n + \overline{X}_n) = 1$$

Maximum a posteriori probability:
equivalently, *minimum* a posteriori *negative log*-probability,

$$-\log\left(\max_{\alpha_1,\ldots,\alpha_n \in \{0,1\}} \{X_1^{\alpha_1} \cdots X_n^{\alpha_1}\}\right)$$

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\ldots(\lambda x.x \oplus_{X_n} x)\text{true}$$

Probability that $M \to$ true: sum over all reductions ($X^0 = X, X^1 = \overline{X}$):

$$\sum_{\alpha_1,\ldots,\alpha_n \in \{0,1\}} X_1^{\alpha_1} \cdots X_n^{\alpha_n} = (X_1 + \overline{X}_1)(X_2 + \overline{X}_2)\cdots(X_n + \overline{X}_n) = 1$$

Maximum a posteriori probability:
equivalently, *minimum* a posteriori *negative log*-probability,
($z_i = -\log X_i, \overline{z}_i = -\log \overline{X}_i$)

$$-\log\left(\max_{\alpha_1,\ldots,\alpha_n \in \{0,1\}} \{X_1^{\alpha_1} \cdots X_n^{\alpha_1}\}\right) = \min_{\alpha_1,\ldots,\alpha_n \in \{0,1\}} \{z_1^{\alpha_1} + \cdots + z_n^{\alpha_1}\}$$

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\dots(\lambda x.x \oplus_{X_n} x)\text{true}$$

Probability that $M \to$ true: sum over all reductions ($X^0 = X, X^1 = \overline{X}$):

$$\sum_{\alpha_1,\dots,\alpha_n \in \{0,1\}} X_1^{\alpha_1} \cdots X_n^{\alpha_n} = (X_1 + \overline{X}_1)(X_2 + \overline{X}_2)\cdots(X_n + \overline{X}_n) = 1$$

Maximum a posteriori probability:
equivalently, *minimum* a posteriori *negative log*-probability,
($z_i = -\log X_i, \overline{z}_i = -\log \overline{X}_i$)

$$-\log\left(\max_{\alpha_1,\dots,\alpha_n \in \{0,1\}} \{X_1^{\alpha_1} \cdots X_n^{\alpha_1}\}\right) = \min_{\alpha_1,\dots,\alpha_n \in \{0,1\}} \{z_1^{\alpha_1} + \cdots + z_n^{\alpha_1}\}$$
$$= \min\{z_1, \overline{z}_1\} + \cdots + \min\{z_n, \overline{z}_n\}$$

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\dots(\lambda x.x \oplus_{X_n} x)\text{true}$$

Probability that $M \twoheadrightarrow$ true: sum over all reductions ($X^0 = X, X^1 = \overline{X}$):

$$\sum_{\alpha_1,\dots,\alpha_n \in \{0,1\}} X_1^{\alpha_1}\cdots X_n^{\alpha_n} = (X_1 + \overline{X}_1)(X_2 + \overline{X}_2)\dots(X_n + \overline{X}_n) = 1$$

Maximum a posteriori probability:
equivalently, *minimum* a posteriori *negative log*-probability,
$(z_i = -\log X_i, \overline{z}_i = -\log \overline{X}_i)$

$$-\log\left(\max_{\alpha_1,\dots,\alpha_n\in\{0,1\}}\{X_1^{\alpha_1}\cdots X_n^{\alpha_1}\}\right) = \min_{\alpha_1,\dots,\alpha_n\in\{0,1\}}\{z_1^{\alpha_1}+\cdots+z_n^{\alpha_1}\}$$
$$= \min\{z_1,\overline{z}_1\}+\cdots+\min\{z_n,\overline{z}_n\}$$

$$M = (\lambda x.x \oplus_{X_1} x)(\lambda x.x \oplus_{X_2} x)\dots(\lambda x.x \oplus_{X_n} x)\mathrm{true}$$

Probability that $M \twoheadrightarrow$ true: sum over all reductions ($X^0 = X, X^1 = \overline{X}$):

$$\sum_{\alpha_1,\dots,\alpha_n \in \{0,1\}} X_1^{\alpha_1}\cdots X_n^{\alpha_n} = (X_1 + \overline{X}_1)(X_2 + \overline{X}_2)\dots(X_n + \overline{X}_n) = 1$$

Maximum a posteriori probability:
equivalently, *minimum* a posteriori *negative log*-probability,
($z_i = -\log X_i, \overline{z}_i = -\log \overline{X}_i$)

$$-\log\left(\max_{\alpha_1,\dots,\alpha_n\in\{0,1\}}\{X_1^{\alpha_1}\cdots X_n^{\alpha_1}\}\right) = \min_{\alpha_1,\dots,\alpha_n\in\{0,1\}}\{z_1^{\alpha_1}+\cdots+z_n^{\alpha_1}\}$$
$$= \min\{z_1,\overline{z}_1\}+\cdots+\min\{z_n,\overline{z}_n\}$$

Tropicalisation: move from $([0,1],+,\times)$ to $\mathbb{T} := ([0,+\infty],\min,+)$

A tropical curve needs *not* depend on all its monomials:

A tropical curve needs *not* depend on all its monomials:

$$X^4 + 4X^3\overline{X} + 6X^2\overline{X}^2 + 4X\overline{X}^3 + \overline{X}^4 \qquad X^4 + \overline{X}^4$$

A tropical curve needs *not* depend on all its monomials:

$$X^4 + 4X^3\overline{X} + 6X^2\overline{X}^2 + 4X\overline{X}^3 + \overline{X}^4 \qquad X^4 + \overline{X}^4$$

$$\min\{4X, 3X + \overline{X}, 2X + 2\overline{X}, X + 3\overline{X}, 4\overline{X}\} \qquad \min\{4X, 4\overline{X}\}$$

A tropical curve needs *not* depend on all its monomials:

$$X^4 + 4X^3\overline{X} + 6X^2\overline{X}^2 + 4X\overline{X}^3 + \overline{X}^4 \qquad X^4 + \overline{X}^4$$

$$\min\{4X, 3X + \overline{X}, 2X + 2\overline{X}, X + 3\overline{X}, 4\overline{X}\} = \min\{4X, 4\overline{X}\}$$

A tropical curve needs *not* depend on all its monomials:

$$\mathsf{t}^!\left(X^4 + 4X^3\overline{X} + 6X^2\overline{X}^2 + 4X\overline{X}^3 + \overline{X}^4\right) = \mathsf{t}^!\left(X^4 + \overline{X}^4\right)$$

$$\min\{4X, 3X + \overline{X}, 2X + 2\overline{X}, X + 3\overline{X}, 4\overline{X}\} \quad = \quad \min\{4X, 4\overline{X}\}$$

Tropicalisation

For s a formal power series in n variables, $\mathsf{t}^! s : \mathbb{T}^n \to \mathbb{T}$ is its evaluation within tropical numbers

A tropical curve needs *not* depend on all its monomials:

$$\mathsf{t}^!\left(X^4 + 4X^3\overline{X} + 6X^2\overline{X}^2 + 4X\overline{X}^3 + \overline{X}^4\right) = \mathsf{t}^!\left(X^4 + \overline{X}^4\right)$$

$$\min\{4X, 3X + \overline{X}, 2X + 2\overline{X}, X + 3\overline{X}, 4\overline{X}\} \qquad = \qquad \min\{4X, 4\overline{X}\}$$

Tropicalisation

For s a formal power series in n variables, $\mathsf{t}^! s : \mathbb{T}^n \to \mathbb{T}$ is its evaluation within tropical numbers

Key idea: $\mathsf{t}^! s$ actually depends on many less monomials than s.

The s above expresses the reductions of the program

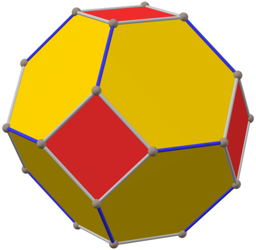
$$(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\text{true}$$

Only 2 monomials out of 5
⇒ maximum probabilities achieved by "all left" and "all right" strategies

Newton polytope: for a polynomial $s = \sum_{(i_1,\dots,i_n)} s_{i_1\dots i_n} x_1^{i_1}\dots x_n^{i_n}$,

$$NP(s) = \text{Convex Hull}(\{(i_1,\dots,i_n) \mid s_{i_1\dots i_n} \neq 0\}) \subset \mathbb{R}^n$$

Call s_{NP} the restriction of s to the vertices of $NP(s)$.

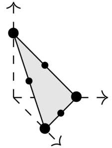


Newton polytope: for a polynomial $s = \sum_{(i_1,\ldots,i_n)} s_{i_1\ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$,

$$NP(s) = \text{Convex Hull}(\{(i_1, \ldots, i_n) \mid s_{i_1\ldots i_n} \neq 0\}) \subset \mathbb{R}^n$$

Call s_{NP} the restriction of s to the vertices of $NP(s)$.

$$\begin{aligned} NP((x_1+x_2+x_3)^2) &= NP(x_1^2 + 2x_1x_2 + 2x_1x_3 + x_2^2 + 2x_2x_3 + x_3^2) \\ &= \text{CH}(\{(2,0,0),(1,1,0),(1,0,1),(0,2,0),(0,1,1),(0,0,2)\}) \\ &= \text{CH}(\{(2,0,0),(0,2,0),(0,0,2)\}) \end{aligned}$$



$$s_{NP} = x_1^2 + x_2^2 + x_3^2$$

Newton polytope: for a polynomial $s = \sum_{(i_1,\ldots,i_n)} s_{i_1\ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$,

$$NP(s) = \text{Convex Hull}(\{(i_1, \ldots, i_n) \mid s_{i_1\ldots i_n} \neq 0\}) \subset \mathbb{R}^n$$

Call s_{NP} the restriction of s to the vertices of $NP(s)$.

Our idea: Minimal Newton polytope:

$$NP_{\min}(s) = \text{Convex Hull}\{\text{minimal vertices of } NP(s)\} \subset \mathbb{R}^n$$

Call $s_{\min}$ the restriction of s to the vertices of $NP_{\min}(s)$

Newton polytope: for a polynomial $s = \sum_{(i_1,\ldots,i_n)} s_{i_1\ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$,

$$NP(s) = \text{Convex Hull}(\{(i_1,\ldots,i_n) \mid s_{i_1\ldots i_n} \neq 0\}) \subset \mathbb{R}^n$$

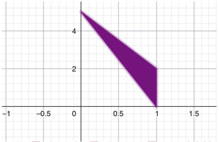
Call s_{NP} the restriction of s to the vertices of $NP(s)$.

Our idea: Minimal Newton polytope:

$$NP_{\min}(s) = \text{Convex Hull}\{\text{minimal vertices of } NP(s)\} \subset \mathbb{R}^n$$

Call $s_{\min}$ the restriction of s to the vertices of $NP_{\min}(s)$

- $NP(3x_1 + 2x_1x_2^2 + x_2^5) = \text{CH}\{(1,0),(1,2),(0,5)\}$
  $s_{\min} = 3x_1 + 2x_1x_2^2 + x_2^5$



Newton polytope: for a polynomial $s = \sum_{(i_1,\dots,i_n)} s_{i_1 \dots i_n} x_1^{i_1} \dots x_n^{i_n}$,

$$NP(s) = \text{Convex Hull}(\{(i_1, \dots, i_n) \mid s_{i_1 \dots i_n} \neq 0\}) \subset \mathbb{R}^n$$

Call s_{NP} the restriction of s to the vertices of $NP(s)$.

Our idea: Minimal Newton polytope:

$$NP_{\min}(s) = \text{Convex Hull}\{\text{minimal vertices of } NP(s)\} \subset \mathbb{R}^n$$

Call $s_{\min}$ the restriction of s to the vertices of $NP_{\min}(s)$

- $NP_{\min}(3x_1 + \cancel{2x_1x_2^2} + x_2^5) = \text{CH}\{(1,0), \cancel{(1,2)}, (0,5)\}$

  $s_{\min} = 3x_1 + \cancel{2x_1x_2^2} + x_2^5$

Tropical Geometry[1] is useful for Statistical Inference:
⇒ efficient methods to study *most likely* trajectories in Bayesian networks

[1] [Pachter, Sturmfeld, 2004] Tropical Geometry of Statistical Models, PNAS

Tropical Geometry[1] is useful for Statistical Inference:
⇒ efficient methods to study *most likely* trajectories in Bayesian networks

Theorem 2. (Pachter, Sturmfeld, 2004)
For a polynomial s in n variables, $NP(s), NP_{\min}(s)$ have $\mathcal{O}((\deg p)^{2n-1})$ vertices that can be computed efficiently from s.

[1] [Pachter, Sturmfeld, 2004] Tropical Geometry of Statistical Models, PNAS

Tropical Geometry[1] is useful for Statistical Inference:
$\Rightarrow$ efficient methods to study *most likely* trajectories in Bayesian networks

Theorem 2. (Pachter, Sturmfeld, 2004)
For a polynomial s in n variables, $NP(s), NP_{\min}(s)$ have $\mathcal{O}((\deg p)^{2n-1})$ vertices that can be computed efficiently from s.

By restricting to the vertices of $NP_{\min}(s)$ we *reduce the space of possible solutions* to a **inference** optimisation problems like:

Given a probabilistic model expressed via some (large) polynomial s on parameters $x_1, \ldots, x_n$, compute a "small" set S of trajectories such that:

- for every observation $\vec{\sigma}$ and assignment $\vec{x} \mapsto \vec{p}$, the most likely explanation of $\vec{\sigma}$ wrt to $\vec{p}$ is in S.

[1] [Pachter, Sturmfeld, 2004] Tropical Geometry of Statistical Models, PNAS

Tropical Geometry[1] is useful for Statistical Inference:
$\Rightarrow$ efficient methods to study *most likely* trajectories in Bayesian networks

Theorem 2. (Pachter, Sturmfeld, 2004)
For a polynomial s in n variables, $NP(s), NP_{\min}(s)$ have $\mathcal{O}((\deg p)^{2n-1})$ vertices that can be computed efficiently from s.

By restricting to the vertices of $NP_{\min}(s)$ we *reduce the space of possible solutions* to a **inference** optimisation problems like:

Given a probabilistic model expressed via some (large) polynomial s on parameters $x_1, \dots, x_n$, compute a "small" set S of trajectories such that:

- for every observation $\vec{\sigma}$ and assignment $\vec{x} \mapsto \vec{p}$, the most likely explanation of $\vec{\sigma}$ wrt to $\vec{p}$ is in S.

Our idea: Import Tropical Geometry tools inside PLT:
$\Rightarrow$ PL-oriented methods to study *most likely* reductions *directly on programs*

[1] [Pachter, Sturmfeld, 2004] Tropical Geometry of Statistical Models, PNAS

Tropical Geometry[1] is useful for Statistical Inference:
$\Rightarrow$ efficient methods to study *most likely* trajectories in Bayesian networks

Theorem 2. (Pachter, Sturmfeld, 2004)
For a polynomial s in n variables, $NP(s), NP_{\min}(s)$ have $\mathcal{O}((\deg p)^{2n-1})$ vertices that can be computed efficiently from s.

By restricting to the vertices of $NP_{\min}(s)$ we *reduce the space of possible solutions* to a **inference** optimisation problems like:

Given a probabilistic model expressed via some (large) polynomial s on parameters $x_1, \dots, x_n$, compute a "small" set S of trajectories such that:

- for every observation $\vec{\sigma}$ and assignment $\vec{x} \mapsto \vec{p}$, the most likely explanation of $\vec{\sigma}$ wrt to $\vec{p}$ is in S.

Our idea: Import Tropical Geometry tools inside PLT:
$\Rightarrow$ PL-oriented methods to study *most likely* reductions *directly on programs*

1. Define $NP_{\min}(M)$ with PLT compositional tools
2. Estimate $NP_{\min}(M)$ in a compositional PL-oriented way

[1] [Pachter, Sturmfeld, 2004] Tropical Geometry of Statistical Models, PNAS

Outline

Language $PCF\langle X_1, \ldots, X_n\rangle$: PCF + parametric coin flip $M \oplus_{X_i} N$

Semantics: We adapt the well-known Weighted Relational Model $Q\mathrm{Rel}_!$ of PCF

type ↦ set of indeces

term ↦ family of formal power series with coefficients in Q

$$[\![\mathrm{Bool}]\!] = \{true, false\} \qquad [\![\mathrm{Nat}]\!] = \mathbb{N} \qquad [\![A \to B]\!] = ![\![A]\!] \times [\![B]\!]$$

Language $PCF\langle X_1, \ldots, X_n \rangle$: PCF + parametric coin flip $M \oplus_{X_i} N$

Semantics: We adapt the well-known Weighted Relational Model $Q\mathrm{Rel}_!$ of PCF

type $\mapsto$ set of indeces

term $\mapsto$ family of formal power series with coefficients in Q

$$[\![\mathrm{Bool}]\!] = \{true, false\} \qquad [\![\mathrm{Nat}]\!] = \mathbb{N} \qquad [\![A \to B]\!] = ![\![A]\!] \times [\![B]\!]$$

Parametric interpretation of PCF$\langle X_1, \ldots, X_n \rangle$

Take Q = semiring of formal power series in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X_n}$ and coefficients in $\mathbb{N}^\infty$. For each $b \in [\![B]\!]$, we can define a formal power series

$$[\![x : A \vdash M : B]\!]_b$$

in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X_n}, x_{[\![A]\!]}$ and coefficients in $\mathbb{N}^\infty$.

A program M : Bool is interpreted as the two fps $[\![M]\!]_{true}$, $[\![M]\!]_{false}$ below:

$$[\![M]\!]_b = \sum_{\vec{i},\vec{j}\in\mathbb{N}^n} \sharp_{\vec{i},\vec{j}} X_1^{i_1}\overline{X}_1^{\,j_1}\cdots X_n^{i_n}\overline{X}_n^{\,j_n}$$

with $\sharp_{\vec{i},\vec{j}} \in \mathbb{N}^\infty$ the number of reductions to b of weight $X_1^{i_1}\overline{X}_1^{\,j_1}\ldots X_n^{i_n}\overline{X}_n^{\,j_n}$.

A program M : Bool is interpreted as the two fps $[\![M]\!]_{true}$, $[\![M]\!]_{false}$ below:

$$[\![M]\!]_b = \sum_{\vec{i},\vec{j}\in\mathbb{N}^n} \sharp_{\vec{i},\vec{j}} X_1^{i_1}\overline{X}_1^{j_1}\cdots X_n^{i_n}\overline{X}_n^{j_n}$$

with $\sharp_{\vec{i},\vec{j}} \in \mathbb{N}^\infty$ the number of reductions to b of weight $X_1^{i_1}\overline{X}_1^{j_1}\ldots X_n^{i_n}\overline{X}_n^{j_n}$.

$$[\![(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\text{true}]\!]_{true} = (X + \overline{X})^4$$

A program M : Bool is interpreted as the two fps $[\![M]\!]_{true}$, $[\![M]\!]_{false}$ below:

$$[\![M]\!]_b = \sum_{\vec{i},\vec{j}\in\mathbb{N}^n} \sharp_{\vec{i},\vec{j}} X_1^{i_1}\overline{X}_1^{\,j_1} \cdots X_n^{i_n}\overline{X}_n^{\,j_n}$$

with $\sharp_{\vec{i},\vec{j}} \in \mathbb{N}^\infty$ the number of reductions to b of weight $X_1^{i_1}\overline{X}_1^{\,j_1} \ldots X_n^{i_n}\overline{X}_n^{\,j_n}$.

$$[\![(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\mathrm{true}]\!]_{true} = (X + \overline{X})^4$$

$$[\![\mathrm{fix}(\lambda x.x \oplus_X \mathrm{true})]\!]_{true} = \sum_{n=0}^{\infty} X^n \overline{X}$$

Our idea: **A Finitary Claim about PCF$\langle\vec{X}\rangle$**

for all PCF$\langle\vec{X}\rangle$ programs M : Bool there is a finite set of reductions S such that, for all assignments $\vec{X} \mapsto \vec{p}$, S contains a most likely reduction;

Our idea: **A Finitary Claim about PCF$\langle\vec{X}\rangle$**

for all PCF$\langle\vec{X}\rangle$ programs M : Bool there is a finite set of reductions S such that, for all assignments $\vec{X} \mapsto \vec{p}$, S contains a most likely reduction;

equivalently, there is $d \in \mathbb{N}$ such that, for all assignments $\vec{X} \mapsto \vec{p}$, a reduction of highest probability is always found among those with at most d choices.

For s formal power series in n variables and coefficients in $\mathbb{N}^{\infty}$, $\mathsf{t}^{!}s : \mathbb{T}^{n} \to \mathbb{T}$

For s formal power series in n variables and coefficients in $\mathbb{N}^\infty$, $\mathsf{t}^! s : \mathbb{T}^n \to \mathbb{T}$

Theorem 3. (all discrete fps collapse)
For any s formal power series with coefficients in $\mathbb{N}^\infty$ there exists a *polynomial* s_0 such that $\mathsf{t}^! s = \mathsf{t}^! s_0$.

For s formal power series in n variables and coefficients in $\mathbb{N}^{\infty}$, $\mathtt{t}^{!}s : \mathbb{T}^{n} \to \mathbb{T}$

Theorem 3. (all discrete fps collapse)
For any s formal power series with coefficients in $\mathbb{N}^{\infty}$ there exists a *polynomial* s_0 such that $\mathtt{t}^{!}s = \mathtt{t}^{!}s_0$.

Definition.
The tropical degree $\mathfrak{d}_b(M)$ (for $b = true, false$) of a program M : Bool is the *smallest* degree of a polynomial s_0 such that $\mathtt{t}^{!}[\![M]\!]_b = \mathtt{t}^{!}s_0$.
The (minimal) Newton Polytope of M is $NP_{\min}(s_0)$.

For s formal power series in n variables and coefficients in $\mathbb{N}^\infty$, $\mathtt{t}^! s : \mathbb{T}^n \to \mathbb{T}$

Theorem 3. (all discrete fps collapse)
For any s formal power series with coefficients in $\mathbb{N}^\infty$ there exists a *polynomial* s_0 such that $\mathtt{t}^! s = \mathtt{t}^! s_0$.

Definition.
The tropical degree $\mathfrak{d}_b(M)$ (for $b = \mathit{true}, \mathit{false}$) of a program M : Bool is the *smallest* degree of a polynomial s_0 such that $\mathtt{t}^! [\![M]\!]_b = \mathtt{t}^! s_0$.
The (minimal) Newton Polytope of M is $NP_{\min}(s_0)$.

Theorem 4. (tropical degree is undecidable)
Both problems "is $\mathfrak{d}_b(M) < d$?" and "is $\mathfrak{d}_b(M) = d$?" are Π^0_1-hard.

Outline

We adapt non-idempotent intersection types[2]: For $\Gamma \vdash M : A$, $\gamma \in !\llbracket \Gamma \rrbracket, a \in \llbracket A \rrbracket$,

$$M : \langle \gamma \vdash^{s} a \rangle$$

with s a polynomial in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n$ and coefficients in $\mathbb{N}$

Each monomial of s records a reduction that the derivation has explored

[2] [Ehrhard, Pagani Tasson] PCS are fully abstract for Probabilistic PCF, POPL 2014.

We adapt non-idempotent intersection types[2]: For $\Gamma \vdash M : A$, $\gamma \in !\llbracket\Gamma\rrbracket, a \in \llbracket A \rrbracket$,

$$M : \langle \gamma \vdash^{s} a \rangle$$

with s a polynomial in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n$ and coefficients in $\mathbb{N}$

Each monomial of s records a reduction that the derivation has explored

$\Rightarrow$ Our idea: record in s as few reductions (of M to a given γ) as possible
& many enough to hope finding a most likely reduction

$\Rightarrow$ Need to compare different reductions
$\Rightarrow$ many typing derivations at once

$$M : \left\langle \gamma_j \vdash^{s_j} a_j \right\rangle_{j \in J}$$

[2] [Ehrhard, Pagani Tasson] PCS are fully abstract for Probabilistic PCF, POPL 2014.

We adapt non-idempotent intersection types[2]: For $\Gamma \vdash M : A$, $\gamma \in ![\![\Gamma]\!], a \in [\![A]\!]$,

$$M : \langle \gamma \vdash^s a \rangle$$

with s a polynomial in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n$ and coefficients in $\mathbb{N}$

Each monomial of s records a reduction that the derivation has explored

$\Rightarrow$ Our idea: record in s as few reductions (of M to a given γ) as possible & many enough to hope finding a most likely reduction

$\Rightarrow$ Need to compare different reductions

$$\frac{\gamma \vdash M : [a_1] \multimap \cdots \multimap [a_k] \multimap b \qquad \delta_1 \vdash N_1 : a_1 \quad \ldots \quad \delta_k \vdash N_k : a_k}{\gamma + \sum_i \delta_i \vdash M N_1 \ldots N_k : b}$$

[2] [Ehrhard, Pagani Tasson] PCS are fully abstract for Probabilistic PCF, POPL 2014.

We adapt non-idempotent intersection types[2]: For $\Gamma \vdash M : A$, $\gamma \in ![\![\Gamma]\!], a \in [\![A]\!]$,

$$M : \langle \gamma \vdash^s a \rangle$$

with s a polynomial in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n$ and coefficients in $\mathbb{N}$

Each monomial of s records a reduction that the derivation has explored

$\Rightarrow$ Our idea: record in s as few reductions (of M to a given γ) as possible & many enough to hope finding a most likely reduction

$\Rightarrow$ Need to compare different reductions

$$\frac{\gamma \vdash^{s_0} M : [a_1] \multimap \cdots \multimap [a_k] \multimap b \qquad \delta_1 \vdash^{s_1} N_1 : a_1 \quad \ldots \quad \delta_k \vdash^{s_k} N_k : a_k}{\gamma + \sum_i \delta_i \vdash^{s_0 s_1 \ldots s_k} M N_1 \ldots N_k : b}$$

The polynomial $s_0 s_1 \cdots s_k$ may have $\mathcal{O}((2n+k)^k)$ monomials!
Problem: It might grow too large in k

[2] [Ehrhard, Pagani Tasson] PCS are fully abstract for Probabilistic PCF, POPL 2014.

We adapt non-idempotent intersection types[2]: For $\Gamma \vdash M : A$, $\gamma \in ![\![\Gamma]\!], a \in [\![A]\!]$,

$$M : \langle \gamma \vdash^{s} a \rangle$$

with s a polynomial in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n$ and coefficients in $\mathbb{N}$

Each monomial of s records a reduction that the derivation has explored

$\Rightarrow$ Our idea: record in s as few reductions (of M to a given γ) as possible & many enough to hope finding a most likely reduction

$\Rightarrow$ Need to compare different reductions

$$\frac{\gamma \vdash^{s_0} M : [a_1] \multimap \cdots \multimap [a_k] \multimap b \qquad \delta_1 \vdash^{s_1} N_1 : a_1 \quad \ldots \quad \delta_k \vdash^{s_k} N_k : a_k}{\gamma + \sum_i \delta_i \vdash^{s_0 s_1 \ldots s_k} M N_1 \ldots N_k : b}$$

The polynomial $s_0 s_1 \cdots s_k$ may have $\mathcal{O}((2n+k)^k)$ monomials!
Problem: It might grow too large in k

Solution: Given *minimal* polynomials $s_0, \ldots, s_k$, a **"Viterbi-Newton" algorithm** $\mathbf{VN}(s_0, \ldots, s_k)$ computes the *minimal* polynomial $\left(\prod_{i=0}^{k} s_i\right)_{\min}$.

[2] [Ehrhard, Pagani Tasson] PCS are fully abstract for Probabilistic PCF, POPL 2014.

We adapt non-idempotent intersection types[2]: For $\Gamma \vdash M : A$, $\gamma \in !\llbracket\Gamma\rrbracket, a \in \llbracket A \rrbracket$,

$$M : \langle \gamma \vdash^{s} a \rangle$$

with s a polynomial in variables $X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n$ and coefficients in $\mathbb{N}$

Each monomial of s records a reduction that the derivation has explored

$\Rightarrow$ Our idea: record in s as few reductions (of M to a given γ) as possible & many enough to hope finding a most likely reduction

$\Rightarrow$ Need to compare different reductions

$$\frac{\gamma \vdash^{s_0} M : [a_1] \multimap \cdots \multimap [a_k] \multimap b \qquad \delta_1 \vdash^{s_1} N_1 : a_1 \quad \ldots \quad \delta_k \vdash^{s_k} N_k : a_k}{\gamma + \sum_i \delta_i \vdash^{\mathbf{VN}(s_0, s_1, \ldots, s_k)} M N_1 \ldots N_k : b}$$

The polynomial $s_0 s_1 \cdots s_k$ may have ~~$\mathcal{O}((2n+k)^k)$~~ $\mathcal{O}((k+1)^{4n-1})$ monomials!

Solution: Given *minimal* polynomials $s_0, \ldots, s_k$, a **"Viterbi-Newton" algorithm** $\mathbf{VN}(s_0, \ldots, s_k)$ computes the *minimal* polynomial $\left(\prod_{i=0}^{k} s_i\right)_{\min}$.

[2][Ehrhard, Pagani Tasson] PCS are fully abstract for Probabilistic PCF, POPL 2014.

$$(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\text{true} : \left\langle \vdash^{X^3+\overline{X}^3} true \right\rangle$$

$$(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\mathrm{true} : \left\langle \vdash^{X^3+\overline{X}^3} \mathit{true} \right\rangle$$

$$\dfrac{\dfrac{\dfrac{\overline{x : \langle x : [[a] \multimap a] \vdash^1 [a] \multimap a \rangle} \quad \overline{x : \langle x : [[a] \multimap a] \vdash^1 [a] \multimap a \rangle}}{x \oplus_X x : \langle x : [[a] \multimap a] \vdash^{X+\overline{X}} [a] \multimap a \rangle}}{\lambda x.x \oplus_X x : \left\langle \vdash^{X+\overline{X}} [[a] \multimap a] \multimap [a] \multimap a \right\rangle} \quad \dfrac{\dfrac{\overline{x : \langle x : [[1] \multimap 1] \vdash^1 [1] \multimap 1 \rangle} \quad \overline{x : \langle x : [[1] \multimap 1] \vdash^1 [1] \multimap 1 \rangle}}{x \oplus_X x : \langle x : [[1] \multimap 1] \vdash^{X+\overline{X}} [1] \multimap 1 \rangle}}{\lambda x.x \oplus_X x : \left\langle \vdash^{X+\overline{X}} [[1] \multimap 1] \multimap [1] \multimap 1 \right\rangle}}{(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \left\langle \vdash^{X^2+\overline{X}^2} [1] \multimap 1 \right\rangle}$$

$$\dfrac{\dfrac{\overline{x : \langle x : [1] \vdash^1 1 \rangle} \quad \overline{x : \langle x : [1] \vdash^1 1 \rangle}}{x \oplus_X x : \langle x : [1] \vdash^{X+\overline{X}} 1 \rangle}}{\lambda x.x \oplus_X x : \left\langle \vdash^{X+\overline{X}} [1] \multimap 1 \right\rangle}$$

$$\dfrac{\dfrac{(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \left\langle \vdash^{X^2+\overline{X}^2} [1] \multimap 1 \right\rangle \quad \lambda x.x \oplus_X x : \left\langle \vdash^{X+\overline{X}} [1] \multimap 1 \right\rangle}{(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \left\langle \vdash^{X^3+\overline{X}^3} [1] \multimap 1 \right\rangle} \quad \overline{1 : \langle \vdash^1 1 \rangle}}{(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)1 : \left\langle \vdash^{X^3+\overline{X}^3} 1 \right\rangle}$$

$$(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\text{true} : \left\langle\ \vdash^{X^3+\overline{X}^3} \mathit{true} \right\rangle$$

$$\dfrac{\dfrac{}{x : \langle x : [[a] \multimap a] \vdash^1 [a] \multimap a\rangle} \qquad \dfrac{}{x : \langle x : [[a] \multimap a] \vdash^1 [a] \multimap a\rangle}}{x \oplus_X x : \langle x : [[a] \multimap a] \vdash^{X+\overline{X}} [a] \multimap a\rangle}$$

$$\dfrac{x \oplus_X x : \langle x : [[a] \multimap a] \vdash^{X+\overline{X}} [a] \multimap a\rangle}{\lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [[a] \multimap a] \multimap [a] \multimap a\rangle}$$

$$\dfrac{\dfrac{}{x : \langle x : [[1] \multimap 1] \vdash^1 [1] \multimap 1\rangle} \qquad \dfrac{}{x : \langle x : [[1] \multimap 1] \vdash^1 [1] \multimap 1\rangle}}{x \oplus_X x : \langle x : [[1] \multimap 1] \vdash^{X+\overline{X}} [1] \multimap 1\rangle}$$

$$\dfrac{x \oplus_X x : \langle x : [[1] \multimap 1] \vdash^{X+\overline{X}} [1] \multimap 1\rangle}{\lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [[1] \multimap 1] \multimap [1] \multimap 1\rangle}$$

$$\dfrac{\lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [[a] \multimap a] \multimap [a] \multimap a\rangle \qquad \lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [[1] \multimap 1] \multimap [1] \multimap 1\rangle}{(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \langle \vdash^{X^2+\overline{X}^2} [1] \multimap 1\rangle}$$

$$\dfrac{\dfrac{\dfrac{}{x : \langle x : [1] \vdash^1 1\rangle} \qquad \dfrac{}{x : \langle x : [1] \vdash^1 1\rangle}}{x \oplus_X x : \langle x : [1] \vdash^{X+\overline{X}} 1\rangle}}{\lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [1] \multimap 1\rangle}$$

$$\dfrac{(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \langle \vdash^{X^2+\overline{X}^2} [1] \multimap 1\rangle \qquad \lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [1] \multimap 1\rangle}{(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \langle \vdash^{X^3+\overline{X}^3} [1] \multimap 1\rangle}$$

$$\dfrac{(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \langle \vdash^{X^3+\overline{X}^3} [1] \multimap 1\rangle \qquad \dfrac{}{1 : \langle \vdash^1 1\rangle}}{(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)1 : \langle \vdash^{X^3+\overline{X}^3} 1\rangle}$$

$$\dfrac{M : \left\langle\ \vdash^{X^2+\overline{X}^2} [a] \multimap b \right\rangle \qquad N : \left\langle\ \vdash^{X+\overline{X}} a \right\rangle}{MN : \left\langle\ \vdash^{X^3+\overline{X}^3} b \right\rangle}$$

$$(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\mathrm{true} : \left\langle \vdash^{X^3+\overline{X}^3} \mathit{true} \right\rangle$$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{x : \langle x : [[a] \multimap a] \vdash^1 [a] \multimap a \rangle} \quad \dfrac{}{x : \langle x : [[a] \multimap a] \vdash^1 [a] \multimap a \rangle}}{x \oplus_X x : \langle x : [[a] \multimap a] \vdash^{X+\overline{X}} [a] \multimap a \rangle}}{\lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [[a] \multimap a] \multimap [a] \multimap a \rangle} \quad \dfrac{\dfrac{\dfrac{}{x : \langle x : [[1] \multimap 1] \vdash^1 [1] \multimap 1 \rangle} \quad \dfrac{}{x : \langle x : [[1] \multimap 1] \vdash^1 [1] \multimap 1 \rangle}}{x \oplus_X x : \langle x : [[1] \multimap 1] \vdash^{X+\overline{X}} [1] \multimap 1 \rangle}}{\lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [[1] \multimap 1] \multimap [1] \multimap 1 \rangle}}{(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \langle \vdash^{X^2+\overline{X}^2} [1] \multimap 1 \rangle}$$

$$\dfrac{\dfrac{\dfrac{}{x : \langle x : [1] \vdash^1 1 \rangle} \quad \dfrac{}{x : \langle x : [1] \vdash^1 1 \rangle}}{x \oplus_X x : \langle x : [1] \vdash^{X+\overline{X}} 1 \rangle}}{\lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [1] \multimap 1 \rangle}$$

$$\dfrac{\dfrac{(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \langle \vdash^{X^2+\overline{X}^2} [1] \multimap 1 \rangle \quad \lambda x.x \oplus_X x : \langle \vdash^{X+\overline{X}} [1] \multimap 1 \rangle}{(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)\lambda x.x \oplus_X x : \langle \vdash^{X^3+\overline{X}^3} [1] \multimap 1 \rangle} \quad \dfrac{}{1 : \langle \vdash^1 1 \rangle}}{(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)(\lambda x.x \oplus_X x)1 : \langle \vdash^{X^3+\overline{X}^3} 1 \rangle}$$

$$\frac{M : \left\langle \vdash^{X^2+\overline{X}^2} [a] \multimap b \right\rangle \qquad N : \left\langle \vdash^{X+\overline{X}} a \right\rangle}{MN : \left\langle \vdash^{X^3+\overline{X}^3} b \right\rangle}$$

$$\mathbf{VN}(X^2 + \overline{X}^2, X + \overline{X}) = \Big((X^2 + \overline{X}^2)(X + \overline{X})\Big)_{\min} = X^3 + \overline{X}^3$$

Theorem 5. For all M : Bool there exists a derivation π of $M : \left\langle \vdash^s : true \right\rangle$ such that $s = NP_{\min}(M)$ and so $\deg s = \mathfrak{d}_{true}(M)$.

Theorem 5. For all M : Bool there exists a derivation π of $M : \left\langle \vdash^{s} : true \right\rangle$ such that $s = NP_{\min}(M)$ and so $\deg s = \mathfrak{d}_{true}(M)$.

Solution to Inference tasks

- from $NP_{\min}(M)$, one can trace back an associated maximal reduction
- from each vertex μ of $NP_{\min}(s)$ one can find (via standard linear optimisation techniques) an assignment $\vec{X}, \overline{\vec{X}} \mapsto \vec{p}, (1 - \vec{p})$ that makes the reduction μ a most likely explanation.

Theorem 5. For all M : Bool there exists a derivation π of $M : \left\langle \vdash^{s} : true \right\rangle$ such that $s = NP_{\min}(M)$ and so $\deg s = \mathfrak{d}_{true}(M)$.

Partial Solution to Inference tasks

- from $NP_{\min}(M)$, one can trace back an associated maximal reduction
- from each vertex μ of $NP_{\min}(s)$ one can find (via standard linear optimisation techniques) an assignment $\vec{X}, \overline{\vec{X}} \mapsto \vec{p}, (1 - \vec{p})$ that makes the reduction μ a most likely explanation.

Limitation: $\mathfrak{d}_{true}(M)$ is undecidable!

Theorem 5. For all M : Bool there exists a derivation π of $M : \left\langle \vdash^s : true \right\rangle$ such that $s = NP_{\min}(M)$ and so $\deg s = \mathfrak{d}_{true}(M)$.

Partial Solution to Inference tasks

- from $NP_{\min}(M)$, one can trace back an associated maximal reduction
- from each vertex μ of $NP_{\min}(s)$ one can find (via standard linear optimisation techniques) an assignment $\vec{X}, \overline{\vec{X}} \mapsto \vec{p}, (1 - \vec{p})$ that makes the reduction μ a most likely explanation.

Limitation: $\mathfrak{d}_{true}(M)$ is undecidable!

⇒ Our idea

The user can still perform proof-search within our typing system and *stop* when increasing the proof does not add new monomials!

Perspectives:

Perspectives:

- Implementation and study of effectiveness *in practice*

Perspectives:

- Implementation and study of effectiveness *in practice*
- Differential Privacy (via a different tropicalization)?

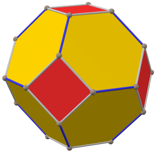
Perspectives:

- Implementation and study of effectiveness *in practice*
- Differential Privacy (via a different tropicalization)?
- Computational look at $Q\mathrm{Rel}_!$ via formal power series?

Perspectives:

- Implementation and study of effectiveness *in practice*
- Differential Privacy (via a different tropicalization)?
- Computational look at $Q\mathrm{Rel}_!$ via formal power series?

Thank you!



$$\frac{}{M : \emptyset}\ \emptyset \qquad \frac{}{x : \left\langle x : [a_i] \vdash^{1} a_i \right\rangle_{i \in I}}\ \text{id} \qquad \frac{}{\mathsf{n} : \left\langle \vdash^{1} n \right\rangle_{\{\star\}}}\ \mathsf{n} \qquad \frac{M : \left\langle \gamma_i \vdash^{s_i} n_i \right\rangle_{i \in I}}{\mathsf{succ}\ M : \left\langle \gamma_i \vdash^{s_i} n_i + 1 \right\rangle_{i \in I}}\ \mathsf{S} \qquad \frac{M : \left\langle \gamma_i \vdash^{s_i} n_i \right\rangle_{i \in I}}{\mathsf{pred}\ M : \left\langle \gamma_i \vdash^{s_i} n_i \dot{-} 1 \right\rangle_{i \in I}}\ \mathsf{P}$$

$$\frac{M : \left\langle \gamma_0 \vdash^{s_0} 0 \,\middle|\, \gamma_{i+1} \vdash^{s_{i+1}} i+1 \right\rangle_{i \in I \subset \mathbb{N}} \qquad N : \left\langle \delta_j \vdash^{t_j} a_j \right\rangle_{j \in J_0} \qquad P : \left\langle \delta'_j \vdash^{t'_j} a'_j \right\rangle_{j \in J_1}}{\mathsf{ifz}(M, N, P) : \text{merge} \left\langle \gamma_0 + \delta_j \vdash^{\mathbf{VN}(s_0, t_j)} a_j \,\middle|\, \gamma_{i+1} + \delta'_j \vdash^{\mathbf{VN}(s_{i+1}, t'_j)} a'_j \right\rangle_{i \in I, j \in J_0 + J_1}}\ \text{ifz}$$

$$\frac{M : \left\langle \gamma_i \vdash^{s_i} a_i \right\rangle_{i \in I} \qquad N : \left\langle \gamma_j \vdash^{s'_j} a_j \right\rangle_{j \in J}}{M \oplus_X N : \text{merge} \left\langle \gamma_i \vdash^{s_i \cdot X} a_i \,\middle|\, \gamma_j \vdash^{s'_j \cdot \overline{X}} a_j \right\rangle_{i \in I, j \in J}}\ \oplus \qquad \frac{M : \left\langle \gamma_i, x : m_i \vdash^{s_i} b_i \right\rangle_{i \in I}}{\lambda x.M : \left\langle \gamma_i \vdash^{s_i} m_i \multimap b_i \right\rangle_{i \in I}}\ \lambda$$

$$\frac{M : \left\langle \gamma_i \vdash^{s_i} m_i \multimap b_i \right\rangle_{i \in I} \qquad N : \left\langle \left\langle \delta_{ij} \vdash^{s'_{ij}} m_{ij} \right\rangle_{j \in J_i} \right\rangle_{i \in I}}{MN : \text{merge} \left\langle \gamma_i + \sum_j \delta_{ij} \vdash^{\mathbf{VN}(s_i, s'_{i1}, \ldots, s'_{ip_i})} b_i \right\rangle_{i \in I}}\ @ \qquad \frac{M : \left\langle \gamma_i \vdash^{s_i} m_i \multimap b_i \right\rangle_{i \in I} \qquad \mathsf{Y}M : \left\langle \left\langle \delta_{ij} \vdash^{s'_{ij}} m_{ij} \right\rangle_{j \in J_i} \right\rangle_{i \in I}}{\mathsf{Y}M : \text{merge} \left\langle \gamma_i + \sum_j \delta_{ij} \vdash^{\mathbf{VN}(s_i, s'_{i1}, \ldots, s'_{ip_i})} b_i \right\rangle_{i \in I}}\ \mathsf{Y}$$

$$\frac{}{M : \emptyset}\ \emptyset \qquad \frac{}{x : \left\langle x : [a_i] \vdash^1 a_i \right\rangle_{i \in I}}\ \text{id} \qquad \frac{}{\mathsf{n} : \left\langle \vdash^1 n \right\rangle_{\{\star\}}}\ \text{n} \qquad \frac{M : \left\langle \gamma_i \vdash^{s_i} n_i \right\rangle_{i \in I}}{\mathsf{succ}\ M : \left\langle \gamma_i \vdash^{s_i} n_i + 1 \right\rangle_{i \in I}}\ \text{S} \qquad \frac{M : \left\langle \gamma_i \vdash^{s_i} n_i \right\rangle_{i \in I}}{\mathsf{pred}\ M : \left\langle \gamma_i \vdash^{s_i} n_i \dot{-} 1 \right\rangle_{i \in I}}\ \text{P}$$

$$\frac{M : \left\langle \gamma_0 \vdash^{s_0} 0 \,\middle|\, \gamma_{i+1} \vdash^{s_{i+1}} i+1 \right\rangle_{i \in I \subset \mathbb{N}} \qquad N : \left\langle \delta_j \vdash^{t_j} a_j \right\rangle_{j \in J_0} \qquad P : \left\langle \delta'_j \vdash^{t'_j} a'_j \right\rangle_{j \in J_1}}{\mathsf{ifz}(M, N, P) : \mathsf{merge} \left\langle \gamma_0 + \delta_j \vdash^{\mathbf{VN}(s_0, t_j)} a_j \,\middle|\, \gamma_{i+1} + \delta'_j \vdash^{\mathbf{VN}(s_{i+1}, t'_j)} a'_j \right\rangle_{i \in I, j \in J_0 + J_1}}\ \text{ifz}$$

$$\frac{M : \left\langle \gamma_i \vdash^{s_i} a_i \right\rangle_{i \in I} \qquad N : \left\langle \gamma_j \vdash^{s'_j} a_j \right\rangle_{j \in J}}{M \oplus_X N : \mathsf{merge} \left\langle \gamma_i \vdash^{s_i \cdot X} a_i \,\middle|\, \gamma_j \vdash^{s'_j \cdot \overline{X}} a_j \right\rangle_{i \in I, j \in J}}\ \oplus \qquad \frac{M : \left\langle \gamma_i, x : m_i \vdash^{s_i} b_i \right\rangle_{i \in I}}{\lambda x.M : \left\langle \gamma_i \vdash^{s_i} m_i \multimap b_i \right\rangle_{i \in I}}\ \lambda$$

$$\frac{M : \left\langle \gamma_i \vdash^{s_i} m_i \multimap b_i \right\rangle_{i \in I} \qquad N : \left\langle \left\langle \delta_{ij} \vdash^{s'_{ij}} m_{ij} \right\rangle_{j \in J_i} \right\rangle_{i \in I}}{MN : \mathsf{merge} \left\langle \gamma_i + \sum_j \delta_{ij} \vdash^{\mathbf{VN}(s_i, s'_{i1}, \ldots, s'_{ip_i})} b_i \right\rangle_{i \in I}}\ @ \qquad \frac{M : \left\langle \gamma_i \vdash^{s_i} m_i \multimap b_i \right\rangle_{i \in I} \qquad \mathsf{Y}M : \left\langle \left\langle \delta_{ij} \vdash^{s'_{ij}} m_{ij} \right\rangle_{j \in J_i} \right\rangle_{i \in I}}{\mathsf{Y}M : \mathsf{merge} \left\langle \gamma_i + \sum_j \delta_{ij} \vdash^{\mathbf{VN}(s_i, s'_{i1}, \ldots, s'_{ip_i})} b_i \right\rangle_{i \in I}}\ \mathsf{Y}$$

In all derivable $M : \left\langle \gamma \vdash^s a \right\rangle$, s is minimal and $|s| \in \mathcal{O}(|\pi|^{2n-1})$

$$\frac{}{M:\emptyset}\ \emptyset \qquad \frac{}{x:\left\langle x:[a_i]\vdash^1 a_i\right\rangle_{i\in I}}\ \text{id} \qquad \frac{}{\mathsf{n}:\left\langle \vdash^1 n\right\rangle_{\{\star\}}}\ \text{n} \qquad \frac{M:\left\langle \gamma_i\vdash^{s_i} n_i\right\rangle_{i\in I}}{\mathsf{succ}\ M:\left\langle \gamma_i\vdash^{s_i} n_i+1\right\rangle_{i\in I}}\ \text{S} \qquad \frac{M:\left\langle \gamma_i\vdash^{s_i} n_i\right\rangle_{i\in I}}{\mathsf{pred}\ M:\left\langle \gamma_i\vdash^{s_i} n_i\dot{-}1\right\rangle_{i\in I}}\ \text{P}$$

$$\frac{M:\left\langle \gamma_0\vdash^{s_0} 0 \;\middle|\; \gamma_{i+1}\vdash^{s_{i+1}} i+1\right\rangle_{i\in I\subset\mathbb{N}} \qquad N:\left\langle \delta_j\vdash^{t_j} a_j\right\rangle_{j\in J_0} \qquad P:\left\langle \delta'_j\vdash^{t'_j} a'_j\right\rangle_{j\in J_1}}{\mathsf{ifz}(M,N,P):\mathsf{merge}\left\langle \gamma_0+\delta_j\vdash^{\mathrm{VN}(s_0,t_j)} a_j \;\middle|\; \gamma_{i+1}+\delta'_j\vdash^{\mathrm{VN}(s_{i+1},t'_j)} a'_j\right\rangle_{i\in I, j\in J_0+J_1}}\ \text{ifz}$$

$$\frac{M:\left\langle \gamma_i\vdash^{s_i} a_i\right\rangle_{i\in I} \qquad N:\left\langle \gamma_j\vdash^{s'_j} a_j\right\rangle_{j\in J}}{M\oplus_X N:\mathsf{merge}\left\langle \gamma_i\vdash^{s_i\cdot X} a_i \;\middle|\; \gamma_j\vdash^{s'_j\cdot\overline{X}} a_j\right\rangle_{i\in I, j\in J}}\ \oplus \qquad \frac{M:\left\langle \gamma_i, x:m_i\vdash^{s_i} b_i\right\rangle_{i\in I}}{\lambda x.M:\left\langle \gamma_i\vdash^{s_i} m_i\multimap b_i\right\rangle_{i\in I}}\ \lambda$$

$$\frac{M:\left\langle \gamma_i\vdash^{s_i} m_i\multimap b_i\right\rangle_{i\in I} \qquad N:\left\langle\left\langle \delta_{ij}\vdash^{s'_{ij}} m_{ij}\right\rangle_{j\in J_i}\right\rangle_{i\in I}}{MN:\mathsf{merge}\left\langle \gamma_i+\sum_j \delta_{ij}\vdash^{\mathrm{VN}(s_i,s'_{i1},\ldots,s'_{ip_i})} b_i\right\rangle_{i\in I}}\ @ \qquad \frac{M:\left\langle \gamma_i\vdash^{s_i} m_i\multimap b_i\right\rangle_{i\in I} \qquad \mathsf{Y}M:\left\langle\left\langle \delta_{ij}\vdash^{s'_{ij}} m_{ij}\right\rangle_{j\in J_i}\right\rangle_{i\in I}}{\mathsf{Y}M:\mathsf{merge}\left\langle \gamma_i+\sum_j \delta_{ij}\vdash^{\mathrm{VN}(s_i,s'_{i1},\ldots,s'_{ip_i})} b_i\right\rangle_{i\in I}}\ \mathsf{Y}$$

In all derivable $M : \left\langle \gamma \vdash^s a \right\rangle$, s is minimal and $|s| \in \mathcal{O}(|\pi|^{2n-1})$

If M is affine, stable derivation efficient $|\pi| \in \mathcal{O}(\mathfrak{d}_{true}(M)|M|3^t)$, while the number of reductions is (at least) in $\mathcal{O}(2^{|M|})$.