Mixture Languages
Oliver Richardson and Jialu Bao

Syntax

\[ \mathcal{L}(\Phi) \equiv \varphi, \varphi' \ ::= \skip | \phi | \varphi; \varphi' | \varphi \oplus \varphi' | \varphi(s) \]

set of primitive commands

\[ \in \Phi \]

primitive

sequencing

mixture

partial execution

\[ s \in [0,1] \]

Operational Semantics

Inductive Construction

sequential execution

\[ \mathbb{X}[\varphi_1; \varphi_2](s) := \begin{cases} 2 \cdot \mathbb{X}[\varphi_1](2s) & s \leq \frac{1}{2} \\ 2 \cdot \mathbb{X}[\varphi_2](2s-1) & s > \frac{1}{2} \end{cases} \]

not associative, but resulting paths are equivalent:

\[ \varphi_1 \oplus \varphi_2 = \varphi_2 \oplus \varphi_1 \]

Based on this, we can derive:

partial execution ("clipping")

\[ \mathbb{X}[\varphi(c)](s) := \mathbb{X}[\varphi](cs) \cdot c \]

parallel execution ("mixture")

\[ \mathbb{X}[\varphi_1 \oplus \varphi_2](s) := \mathbb{X}[\varphi_1](s) + \mathbb{X}[\varphi_2](s) \]

no op

\[ \mathbb{X}[\text{skip}](s) := 0 \]

Key Example: Probabilistic Models

\[ D(\mu \parallel p) = \mathbb{E}_\mu \left[ \log \frac{\mu}{p} \right] \]

Relative entropy of belief \( \mu \) with respect to reality \( p \)

Observe \( p(Y|X) \)

Draw \( y \sim p \mid X \)

Probabilistic Dependency Graphs

(and hence Bayesian Networks and Factor Graphs)

are mixtures of observe commands!

Probabilistic programs are sequences of draw commands

(and deterministic observe commands).