Basing Markov Transition Systems on the Giry Monad

Ernst-Erich Doberkat
Chair for Software Technology
Technische Universität Dortmund
http://ls10-www.cs.uni-dortmund.de

Overview

Motivation

You are interested in the behavior of coalitions, not in the behavior of individual objects (Economics, Biology). You model the system through a stochastic model (e. g., a Markov transition system). Then it is sometimes preferable not to compare the states of a modelling system but rather distributions over these states.

Some questions come up

- How do you model morphisms?
- How is equivalent behavior modelled?

This is what will be done:

- A brief introduction to the Giry monad as the mechanism underlying Markov transition systems.
- A discussion of moving equivalence relations around.
- A look into morphisms and congruences, including factoring.

Measurable Space

A measurable space $(X, \mathcal{A})$ is a set X with a σ-algebra $\mathcal{A}$ — a Boolean algebra of subsets of X which is closed under countable unions.

Measurable Map

Let $(Y, \mathcal{B})$ be another measurable space. A map $f : X \to Y$ is $\mathcal{A}$-$\mathcal{B}$-measurable iff $f^{-1}[B] \in \mathcal{A}$ for all $B \in \mathcal{B}$.

Measurable spaces with measurable maps form a category **Meas**.

Subprobability

A subprobability μ on $\mathcal{A}$ is a map $\mu : \mathcal{A} \to [0, 1]$ with $\mu(\emptyset) = 0$ such that $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} A_n$, provided $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ is mutually disjoint.

$\mathbb{S}(X, \mathcal{A})$

The set $\mathbb{S}(X, \mathcal{A})$ of all subprobabilities is made into a measurable space: take as a σ-algebra $\mathcal{A}^\bullet$ the smallest σ-algebra which makes the evaluations $ev_A : \mu \mapsto \mu(A)$ measurable.

$\mathbb{S}$: Meas $\rightarrow$ Meas

What about morphisms? Let $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ be measurable. Then $\mathbb{S}(f)(\mu) : \mathcal{B} \ni B \mapsto \mu(f^{-1}[B])$.

Easy

$\mathbb{S}(f) : \mathbb{S}(X, \mathcal{A}) \rightarrow \mathbb{S}(Y, \mathcal{B})$ is $\mathcal{A}^{\bullet}$-$\mathcal{B}^{\bullet}$-measurable.

Proposition

$\mathbb{S}$ is an endofunctor on the category of measurable spaces.

Remark

$\mathbb{S}$ is also an endofunctor also on the subcategory of Polish spaces or the subcategory of analytic spaces.

Well ...

But I promised you a monad.

RECALL

A Kleisli triple $(\mathbb{F}, \eta, -^*)$ on a category $\mathbf{C}$ has

1. a map $\mathbb{F} : Obj(\mathbf{C}) \to Obj(\mathbf{C})$,
2. for each object A a map $\eta_A : A \to \mathbb{F}(A)$,
3. for each morphism $f : A \to \mathbb{F}(B)$ a morphism $f^* : \mathbb{F}(A) \to \mathbb{F}(B)$

that satisfy these laws

$$\eta_A^* = id_{\mathbb{F}(A)}, f^* \circ \eta_A = f, g^* \circ f^* = (g^* \circ f)^*$$

$(\mathbb{F}, \mathfrak{e}, \mathfrak{m})$

Put $\mathfrak{m}_A := id^*_{\mathbb{F}(A)}$ and $\mathfrak{e}_A := \eta_A$. One shows that $\mathfrak{m} : \mathbb{F}^2 \xrightarrow{\bullet} \mathbb{F}$ and $\mathfrak{e} : 1\!\!1_{\mathbf{C}} \xrightarrow{\bullet} \mathbb{F}$ satisfy the usual laws for a monad.

$\mathbb{S}$?

Put for measurable $K : (X, \mathcal{A}) \to \mathbb{S}(Y, \mathcal{B})$

$$K^*(\mu)(B) := \int_X K(x)(B)\ \mu(dx).$$

and

$$\eta_A(x)(D) := \begin{cases} 1, & x \in D, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition

$(\mathbb{S}, \eta, -^*)$ is a Kleisli tripel. The corresponding monad is the Giry monad.

Kleisli Morphisms

A stochastic relation $K : (X, \mathcal{A}) \rightsquigarrow (Y, \mathcal{B})$ is a $\mathcal{A}$-$\mathcal{B}^\bullet$-measurable map. Stochastic relations are just the morphisms in the Kleisli category associated with this monad. Kleisli composition $L{*}K$ for $K : (X, \mathcal{A}) \rightsquigarrow (Y, \mathcal{B})$ and $L : (Y, \mathcal{B}) \rightsquigarrow (Z, \mathcal{C})$ is given through

$$\big(L{*}K\big)(x)(D) = \int_Y L(y)(D)\ K(x)(dy)$$

Remark on Language

In Probability Theory stochastic relations are referred to as sub Markov kernels; the Kleisli product of stochastic relations is called there the convolution of kernels.

From now on

X is assumed to be a Polish space with the Borel sets $\mathcal{B}(X)$ as the σ-algebra; $\mathcal{B}(X)$ is the smallest σ-algebra which contains the open sets of X.

Countably Generated Equivalence

An equivalence relation ρ on X is called smooth iff there is a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{B}(X)$ of Borel sets with

$$x \ \rho \ x' \Leftrightarrow \left[\forall n \in \mathbb{N} : x \in A_n \Leftrightarrow x' \in A_n\right]$$

iff $\rho = \ker(f)$ for some measurable $f : X \to \mathbb{R}$.

Example

Assume X is the state space for a Markov transition system $\mathcal{K}$ that interprets a modal logic with a countable number of formulas. Put

$$x \ \rho_{\mathcal{K}} \ x' \text{ iff } \forall \varphi : \mathcal{K}, x \models \varphi \Leftrightarrow \mathcal{K}, x' \models \varphi.$$

Then $\rho_{\mathcal{K}}$ is smooth. This is used when looking at behavioral equivalence of Markov transition systems.

Invariant Sets

A Borel subset $B \subseteq X$ is called ρ-invariant iff $B = \bigcup\{[x]_\rho \mid x \in B\}$, thus iff $x \in B$ and $x \; \rho \; x'$ implies $x' \in B$.
The invariant Borel sets form a σ-algebra $\mathcal{INV}\left(\mathcal{B}(X), \rho\right)$ on X.

Lift a Relation $X \hookrightarrow \mathbb{S}(X)$

Define the lifting $\overline{\rho}$ of the smooth equivalence relation ρ from X to $\mathbb{S}(X)$ through

$$\mu \; \overline{\rho} \; \mu' \text{ iff } \forall B \in \mathcal{INV}\left(\mathcal{B}(X), \rho\right) : \mu(B) = \mu'(B).$$

Example

Let $\rho_{\mathcal{K}}$ be defined through a modal logic. Then

$$\mu \; \overline{\rho_{\mathcal{K}}} \; \mu' \text{ iff } \forall \varphi : \mu(\llbracket \varphi \rrbracket_{\mathcal{K}}) = \mu'(\llbracket \varphi \rrbracket_{\mathcal{K}}).$$

This is used for looking at distributional equivalence of Markov transition systems.

$\mathbb{S}(X) \hookrightarrow X$

If ξ is a smooth equivalence relation on $\mathbb{S}(X)$, then

$$x \lfloor \xi \rfloor x' \text{ iff } \eta_X(x) \; \xi \; \eta_X(x')$$

defines a smooth equivalence relation on X (η_X comes from the monad).

Of course

If ρ is smooth on X, $\rho = \lfloor \overline{\rho} \rfloor$.

Grounded

$$\xi = \overline{\lfloor \xi \rfloor}$$

Near Grounded

$$\xi \subseteq \overline{\lfloor \xi \rfloor}$$

Characterization

The equivalence relation ξ in $\mathbb{S}(X)$ is grounded iff $\xi = \ker(H)$ for a point-affine, surjective and continuous map $H : \mathbb{S}(X) \to \mathbb{S}(\mathbb{R})$.

Congruence

Let $K : X \rightsquigarrow X$ be a stochastic relation. The smooth equivalence relation ρ on X is a congruence for K iff $K(x)(D) = K(x')(D)$, whenever $x \ \rho \ x'$ and $D \subseteq X$ is a ρ-invariant Borel set.

If ρ cannot distinguish x and x', and if it cannot distinguish the elements of D, then the probabilities must be equal.

Proposition: Factoring

ρ is a congruence for K iff there exists a stochastic relation $K_\rho : X/\rho \rightsquigarrow \mathbb{S}(X/\rho)$ such that this diagram commutes

$$\begin{tikzcd}
X \arrow[r, "\varepsilon_\rho"] \arrow[d, "K"'] & X/\rho \arrow[d, "K_\rho"] \\
\mathbb{S}(X) \arrow[r, "\mathbb{S}(\varepsilon_\rho)"'] & \mathbb{S}(X/\rho)
\end{tikzcd}$$

(ε_ρ is a morphism $K \to K_\rho$)

RANDOMIZED CONGRUENCE

The smooth equivalence relation ξ on $\mathbb{S}(X)$ is a randomized congruence for the stochastic relation $K : X \rightsquigarrow X$ iff

- ξ is near grounded,
- $\mu \; \xi \; \mu'$ implies $K^*(\mu) \; \xi \; K^*(\mu')$.

Factoring for the randomized case?

MORPHISM

If $K : X \rightsquigarrow X$ and $L : Y \rightsquigarrow Y$ are stochastic relations, then $f : X \to Y$ measurable is called a morphism $K \to L$ iff $L \circ f = \mathbb{S}(f) \circ K$.

RANDOMIZED MORPHISM

If $K : X \rightsquigarrow X$ and $L : Y \rightsquigarrow Y$ are stochastic relations, then $\Phi : X \rightsquigarrow Y$ is called a randomized morphism $K \Rightarrow L$ iff $L * \Phi = \Phi * K$.

MORPHISM

If $f : K \to L$ is a morphism, then $\ker(f)$ is a congruence for K.

RANDOMIZED MORPHISM

If $\Phi : K \Rightarrow L$ is a randomized morphism, then $\ker(\Phi^*)$ is a randomized congruence for K.

The randomized morphism $\Phi : K \Rightarrow L$ is called near-grounded iff $\ker(\Phi^*)$ is near grounded.

For morphism $f : K \to L$ there exists a unique morphism $g : K/\ker(f) \to L$ with

$$\begin{tikzcd} K \arrow[r, "f"] \arrow[d, "\varepsilon_{\mathbf{ker}(f)}"'] & L \\ K/\ker(f) \arrow[ur, "g"'] & \end{tikzcd}$$

If $\Phi : K \Rightarrow L$ is near-grounded, then there exists a unique randomized morphism $\Gamma : K/\ker(\Phi) \Rightarrow L$ with

$$\begin{tikzcd} K \arrow[r, Rightarrow, "\Phi"] \arrow[d, Rightarrow, "\varepsilon_{\mathbf{ker}(\Phi)}"'] & L \\ K/\ker(\Phi) \arrow[ur, Rightarrow, "\Gamma"'] & \end{tikzcd}$$

Some Remarks & Consequences

Factoring is done along the lines of Universal Algebra. Polish (or analytic) base spaces seem to be essential, at least for the proofs.

Using these results, various forms of comparing the behavior of Markov transition systems can be investigated (behavioral equivalence. logical equivalence, bisimulations etc.)

In particular, these results are tools for investigating bisimulations for weak morphisms.

It seems that in general some work could be done to investigate properties of Kleisli categories for not-set based functors.

Some References

M. Giry.
A categorical approach to probability theory.
In *Categorical Aspects of Topology and Analysis*, number 915 in Lect. Notes Math., pages 68 – 85, Berlin, 1981. Springer-Verlag.

P. Panangaden.
Probabilistic relations.
In C. Baier, M. Huth, M. Kwiatkowska, and M. Ryan, editors, *Proc. PROBMIV*, pages 59 – 74, 1998.

E.-E. Doberkat.
Stochastic Relations. Foundations for Markov Transition Systems.
Chapman & Hall/CRC Press, Boca Raton, New York, 2007.

E.-E. Doberkat.
Kleisli morphisms and randomized congruences for the Giry monad.
J. Pure Appl. Alg., 211:638–664, 2007.

E.-E. Doberkat.
Stochastic coalgebraic logic: Bisimilarity and behavioral equivalence.
Ann. Pure Appl. Logic, 155:46 – 68, 2008.