Math. Struct. in Comp. Science (2002), *vol.* 12, *pp.* 481–512. © 2002 Cambridge University Press
DOI: 10.1017/S0960129502003614 Printed in the United Kingdom

Universal aspects of probabilistic automata

LUTZ SCHRÖDER[†] and PAULO MATEUS[‡]

[†] *Department of Mathematics and Computer Science, University of Bremen, Germany.*
[‡] *CMA, Department of Mathematics, IST, Technical University of Lisbon, Portugal.*

Received 15 October 2000; revised 25 July 2001

Frequently, mathematical structures of a certain type and their morphisms fail to form a category for lack of composability of the morphisms; one example of this problem is the class of probabilistic automata when equipped with morphisms that allow restriction as well as relabelling. The proper mathematical framework for this situation is provided by a generalisation of category theory in the shape of the so-called precategories, which are introduced and studied in this paper. In particular, notions of adjointness, weak adjointness and partial adjointness for precategories are presented and justified in detail.
This makes it possible to use universal properties as characterisations of well-known basic constructions in the theory of (generative) probabilistic automata: we show that accessible automata and decision trees, respectively, form coreflective subprecategories of the precategory of probabilistic automata. Moreover, the aggregation of two automata is identified as a partial product, whereas restriction and interconnection of automata are recognised as Cartesian lifts.

1. Introduction

More frequently than one would initially suspect, reasonable notions of morphisms between structured sets (or more general objects) fail to be stable under composition. A typical example is formed by the generalised morphisms between probability spaces and probabilistic automata, respectively, as introduced in Mateus *et al.* (1999); further examples are discussed in Section 2. Thus, the need for a 'weak category theory' arises; the objects of this theory will have to be structures made up of objects, morphisms and a composition operation in the usual sense, but with rather weaker axiomatic requirements than usually imposed on categories. In particular, the composability of adjacent morphisms will, of course, have to be removed from the list of axioms. As a consequence, the precise form of the identity and associativity laws, if any, is *a priori* unclear, since these equations admit a host of different interpretations in the presence of a properly partial composition operation. Further problems of the same nature appear in the attempt to transfer categorical concepts such as natural transformations or adjointness to this more general setting.

Until recently, such structures have been largely ignored in the literature, with the exception of Ehresmann's work on multiplicative graphs (Ehresmann 1965), which was continued, for example, in Coppey (1976) and Coppey (1980). The slightly more general

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

notion of composition graph was introduced in Schröder (1999), partly motivated by the study of quotient categories. This notion specialises to the concept of semicategory, which has its main applications in the construction of extensions of categories (Schröder 1999; Schröder and Herrlich 2000a; Schröder and Herrlich 2000b), as well as to that of precategory, which was introduced, in parallel, in Mateus *et al.* (1999). The latter concept provides the proper setting for the study of structured sets with non-composable morphisms and hence forms the base for the applications presented here; as indicated above, these applications deal with the precategories of probability spaces and probabilistic automata, respectively, in the sense of Mateus *et al.* (1999).

In particular, a notion of adjointness for functors between precategories (or, in fact, multiplicative graphs) has been introduced and justified by suitable characterisation theorems in Schröder (1999). We give a brief account of this theory and present examples in the shape of full coreflective subprecategories of the precategory of probabilistic automata formed by the accessible probabilistic automata and the decision trees, respectively.

We then move on to present a weakened notion of adjointness, which naturally appears in certain other situations such as the Eilenberg–Moore construction for precategories (Schröder 2001) and coincides with the original definition on the level of categories. In order to deal with an additional problem that arises in the applications presented here, namely that independent products of probability spaces (or aggregations of probabilistic automata) admit factorisations of independent pairs of maps, but not of arbitrary pairs, we finally generalise the adjointness definition to accommodate this type of 'weak universal property'. This additional complication is independent of the fact that we consider generalised morphisms of probability spaces, but rather appears in the case of classical measure preserving transformations as well. Our definition of weak adjointness turns out to be naturally adaptable to this problem on the theoretical side, and the definition of partial adjointness thus obtained does indeed cover the said examples.

Finally, we show how restriction and interconnection of probabilistic automata can be interpreted as the formation of subobjects (in fact, as Cartesian lifts of inclusions) in the precategory of probabilistic automata as suggested in Mateus *et al.* (1999); while the former concept is independent of the previous results, the latter builds on the definition of aggregation and hence on the adjointness concepts outlined above. Since the classical concept of a measure preserving transformation does not allow for proper subobjects, this result illustrates the necessity of generalised morphisms as well as that of the theory of partial adjointness: it is the combination of these two concepts that makes a theoretically sound definition of interconnection possible, thus implementing the dogmatics of the Categorical Manifesto (Goguen 1991).

2. Precategories

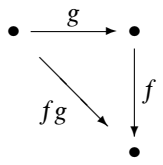
We recall the basic notions of the theory of precategories as developed in Schröder (1999) (where precategories are treated as a special case of the more general notion of composition graph); categorical terminology generally follows Adámek *et al.* (1990).

A *precategory* is a structure consisting of a (possibly large) class of *objects*, denoted A, B, and so on, *morphisms* between these objects, denoted $f : A \to B$, and so on, in

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

particular, an *identity* $id_A : A \to A$ for each object A (we will, in fact, not distinguish between A and id_A), and a partial binary operation on morphisms, called *composition* and denoted $(f, g) \mapsto fg$, subject to the following axioms:

(i) Whenever fg is defined, f, g and fg form a commutative triangle in the usual sense:



(ii) For each morphism $f : A \to B$, we have $id_B f$ and $f id_A$ are defined and equal to f.

(iii) Composition is *weakly associative*: if fg and gh are defined, the equation

$$f(gh) = (fg)h$$

holds strongly (that is, whenever one side is defined, so is the other, and both sides are equal).

Whether or not the classes $hom(A, B)$ of morphisms between two given objects A and B are required to form sets is irrelevant here. Precategories are denoted by bold capital letters **A**, **B**, and so on. The classes of objects and morphisms are written $\mathrm{Ob}\,\mathbf{A}$ and $\mathrm{Mor}\,\mathbf{A}$, respectively.

The duality principle of category theory (Adámek *et al.* 1990) extends to precategories. As usual, dual concepts will implicitly be named by adding the prefix 'co-', unless otherwise stated.

A *graph morphism* $F : \mathbf{A} \to \mathbf{B}$ between precategories is a map F on objects and an assignment that associates to each morphism $f : A \to B$ in **A** a morphism $Ff : FA \to FB$ in **B** such that identities are preserved (that is, $Fid_A = id_{FA}$ for each object A). F is called a *functor* if, moreover, $FfFg$ is defined and equal to $F(fg)$ whenever fg is defined. A precategory **B** is a *subprecategory* of a precategory **A** if $\mathrm{Ob}\,\mathbf{B} \subset \mathrm{Ob}\,\mathbf{A}$, $\mathrm{Mor}\,\mathbf{B} \subset \mathrm{Mor}\,\mathbf{A}$, and the inclusion $\mathbf{B} \hookrightarrow \mathbf{A}$ is a functor.

Remark 2.1. Obviously, every category is a precategory. Formally, the only difference between categories and precategories is that for categories, strong equality is replaced by existential equality (that is, both sides are defined and equal) in Condition (iii) above. (Taking g as an identity then produces the usual definedness requirement for composites.) Removing Condition (iii) altogether leads to C. Ehresmann's *multiplicative graphs* or *neocategories* (Ehresmann 1965); *composition graphs* (Schröder 1999; Schröder and Herrlich 2000b) are defined by Condition (i) alone. Applications of composition graphs, in particular, the more restrictive *semicategories* (which are defined by the existential version of Condition (iii) and a weakened version of Condition (ii)), and more details can be found in Schröder and Herrlich (2000b).

The seemingly slight axiomatic gap between categories and precategories should not be misconstrued as an indication that precategories are 'almost categories'. In the concept of precategory, equational partiality is given up and replaced by proper partiality, so that the difference between categories and precategories is essentially the same as that between algebra and partial algebra. (For example, arbitrary subsets of monoids equipped with

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

the induced partial operation are one-object precategories.) The weak associative law in the definition of precategory is, in most cases, a convenient simplification rather than a necessity; in fact, all of the results on precategories presented below can be adapted to the setting of mere multiplicative graphs (at the sole cost of making them harder to read). Moreover, it should be noted that precategories obey a suitably generalised weak associative law for terms of length 4, but not for longer terms (Schröder 1999), and that, even in short terms, the weak associative law does not warrant the general omission of brackets.

The proper generalisation of concrete categories is, then, given by the concept of a *concrete precategory* over a precategory **B**, that is, a precategory **A**, equipped with a *forgetful functor* $F : \mathbf{A} \to \mathbf{B}$ that is *initial* in the sense that, whenever $g : A \to B$, $f : B \to C$ and $h : A \to C$ are morphisms in **A** such that $FfFg$ is defined and equals Fh, we have fg is defined and equals h ('F reflects commutatation of triangles'). A subprecategory **B** of a precategory **A** is called *initial* if the inclusion functor $\mathbf{B} \hookrightarrow \mathbf{A}$ is initial, that is, if fg is defined in **B** whenever fg is defined in **A** and belongs to **B**. A *concrete subprecategory* of a concrete precategory is an initial subprecategory with the forgetful functor inherited by restriction.

Remark 2.2. It is easily seen that initial functors in the sense just defined are faithful. Properly understood, initial functors are initial (or Cartesian) morphisms in the concrete quasicategory (Adámek *et al.* 1990) of precategories and functors over the quasicategory of large directed graphs and graph morphisms. A functor between *categories* is initial iff it is faithful.

Initial functors can, of course, be defined for arbitrary composition graphs (in the sense of Remark 2.1) in literally the same way as for precategories. Note that, if **A** is a composition graph and $F : \mathbf{A} \to \mathbf{B}$ is an initial functor into a category (or even just a precategory) **B**, then **A** is automatically a precategory. By virtue of this mechanism, all the concrete examples presented in this paper, most of which consist of structured sets (at worst: pairs of sets) with some choice of 'structure preserving maps', are indeed precategories.

We conclude this section with a number of examples.

Example 2.3. Let Σ be a signature. A homomorphism $f : A \to B$ of (partial) Σ-algebras is called *full* (Grätzer 1979) if, for any operation σ of arity n in Σ, the definedness of

$$\sigma(f(x_1), \ldots, f(x_n)) = f(x)$$

in B implies the existence of $x'_1, \ldots, x'_n$ and x' in A such that $\sigma(x'_1, \ldots, x'_n) = x'$ is defined, $f(x') = f(x)$, and $f(x'_i) = f(x_i)$ for $i = 1, \ldots, n$.

Full homomorphisms are not stable under composition: for example, let Σ consist of a single unary operation σ. Let A be the algebra with carrier set $\{0\}$ in which $\sigma(0)$ is undefined, let B be the algebra over $\{0, 1\}$ in which $\sigma(1) = 1$ and $\sigma(0)$ is undefined, and let C be the algebra over $\{0\}$ in which $\sigma(0) = 0$. Moreover, let $f : A \hookrightarrow B$ and $g : B \to C$ denote the obvious maps. Then f and g are full homomorphisms, but gf is not.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Example 2.4. In a suitable multisorted setting with a unit sort, relation morphisms in the usual sense can be regarded as a special case of homomorphisms of partial algebras; in the special case of unary relations, we thus obtain a precategory **fRel(1)** of unary relations and full relation morphisms. Explicitly, the objects of **fRel(1)** are pairs (X, A), where $A \subset X$, and a map $f : (X, A) \to (Y, B)$ is a morphism iff $f[A] \subset B$ (that is, f is a relation morphism) and for each $x \in X$ such that $f(x) \in B$, there exists $x' \in A$ such that $f(x') = f(x)$. Using essentially the same counterexample as above (replacing operations by their domains of definition), one easily shows that **fRel(1)** is not a category.

Example 2.5. Full binary relation morphisms (essentially) appear in Knauer and Böttcher (1992) under the name 'halfstrong graph morphisms'; in the same paper, two other types of graph morphisms (quasi-strong and locally strong graph morphisms) are discussed that also fail to compose.

Example 2.6. It is easily checked that a relation morphism is full if and only if it is (extremal epi, extremal mono)-factorisable in the category of relation morphisms. The observation that morphisms with this factorisation property in general fail to compose produces numerous other examples. Starting, for example, from the category of topological spaces and continuous maps, this leads to the precategory of topological spaces and continuous maps $f : X \to Y$ with the additional property that whenever $f^{-1}[A]$ is open for some $A \subset Y$, there exists an open set $U \subset Y$ such that $U \cap f[X] = A \cap f[X]$.

Example 2.7. In a similar spirit, groups and embeddings of normal subgroups form a precategory.

Example 2.8. Sets and non-constant maps form a precategory $\mathbf{Set_{nc}}$. In fact, $\mathbf{Set_{nc}}$ satisfies an even stronger associative law: the equation $(fg)h = f(gh)$ holds strongly without further preconditions. Such things might be called *strong precategories*. However, it is easily seen that the strong precategories are precisely the precategories of non-zero morphisms in pointed categories (equivalently: the precategories of non-constant morphisms in arbitrary categories) and are, thus, of rather limited interest as a separate concept (Schröder 1999).

Example 2.9. The composite of two topologically-algebraic functors need not be topologically-algebraic (Herrlich *et al.* 1980). Thus, (large) categories and topologically-algebraic functors form a 'quasiprecategory' (that is, a structure that behaves like a precategory except for the fact that its collections of objects and morphisms, respectively, form illegitimate conglomerates (Adámek *et al.* 1990)).

Example 2.10. Taking sets X with a distinguished subset $A \subset X$, written in the form (X, A), as objects, and as morphisms $f : (X, A) \to (Y, B)$ partial maps $f : X \to Y$ that are *sufficiently defined*, that is, defined at least on A, one obtains a precategory **Suf**. (It is easily seen that sufficiently defined maps fail to compose in general.) We will see below that this is a particularly 'bad' example.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

3. Probability spaces

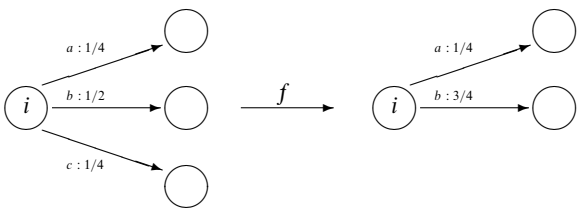
Recall that a *probability space* in all its generality is a triple $(\Omega, \mathfrak{B}, P)$, where Ω is a non-empty set (the *outcome space*), $\mathfrak{B}$ is a σ-algebra over Ω, and P is a probability measure on $\mathfrak{B}$ (Halmos 1950). However, for the purpose of studying finite or countably infinite automata, it suffices to work with the following notion of a (countable) probability presentation.

Definition 3.1. A *probability presentation* is a pair (Ω, p) where Ω is a countable non-empty set and $p : \Omega \to \mathbb{R}_0^+$ is a map such that $\sum_{\omega\in\Omega} p(\omega) = 1$. For a subset B of Ω, we denote the value $\sum_{\omega\in B} p(\omega)$ by $P(B)$; where not otherwise stated, we will use the letters p and P without distinction for all probability presentations.

Obviously, every such probability presentation (Ω, p) induces a probability space with σ-algebra $\mathfrak{P}\Omega$; conversely, a probability presentation can be specified by the data (Ω, P).

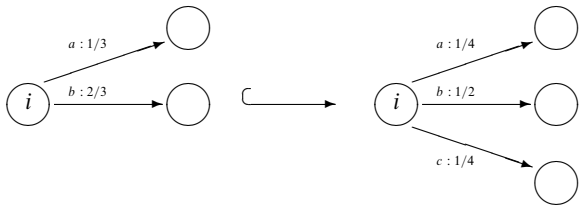
In this context, probability spaces will serve mainly as underlying structures of probabilistic automata. Now, there are certain standard operations that are commonly performed on such automata (whatever may be chosen as their precise definition). Among these operations are relabelling and restriction (of actions). For example, these two operations are explicitly incorporated in the specification language for probabilistic processes PCCS (Giacalone *et al.* 1990); three different types of probabilistic automata are discussed in van Glabbeek *et al.* (1995) as models for this language.

Now, it is a natural requirement that models of any language should be equipped with morphisms that reflect the structure of the language; in particular, there should be morphisms between probabilistic automata that represent relabelling and restriction. On the probabilistic side, these operations correspond to the formation of image distributions and conditional distributions, respectively; *cf.* van Glabbeek *et al.* (1995). For example, if a (generative) automaton m has actions a, b, and c that occur with a probability of $1/4$, $1/2$, and $1/4$, respectively, resulting in a transition from an initial state i to distinct states, then a relabelling f that renames c to b and leaves the rest unchanged would produce the following result:



On the other hand, applying a restriction to m that excludes action c would lead to an inclusion of automata:

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press



Here, the subautomaton is equipped with the conditional probability distribution. (This style of association of morphisms to relabelling and restriction has been suggested in Winskel and Nielsen (1995) for the non-probabilistic case.)

In summary, it appears that two types of morphisms between probability spaces are required: measure-preserving (surjective) maps (Halmos 1950), that is, maps f with the property

$$P(B) = P(f^{-1}[B])$$

for each subset B of the target space (note that any map with this property is *probabilistically surjective*, that is, its image has probability 1), and embeddings of subsets A of the outcome space Ω, equipped with the conditional probability distribution (that is, the probability of a subset $B \subset A$ in the subspace A is given by $P(B|A)$).

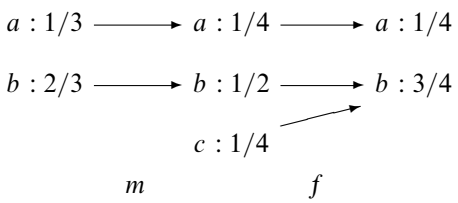
The natural way to unify these concepts into a single definition of morphism consists of generating morphisms in the same fashion as one would reconstruct the morphisms of a category from a factorisation system: if $(\mathscr{E}, \mathscr{M})$ is a factorisation system on a category (Adámek *et al.* 1990), then, by definition, all morphisms of the category are of the form me, where $m \in \mathscr{M}$ and $e \in \mathscr{E}$; by analogy, we shall define morphisms between probability spaces in such a way that each morphism f factors as $f = me$, where e is a measure preserving surjective map and m is a subspace embedding as defined above. It is easily seen that this leads to the following notion.

Definition 3.2. A *probability presentation morphism* (or just a *morphism*) $f : (\Omega_1, p_1) \to (\Omega_2, p_2)$ is a map $f : \Omega_1 \to \Omega_2$ such that $P(f[\Omega_1]) > 0$ and

$$P_2(B|f[\Omega_1]) = P_1(f^{-1}[B])$$

for each $B \subset \Omega_2$.

The main defect of these morphisms is that they are not stable under composition; for example, composing an embedding and a surjection 'in the wrong order' may fail to yield a morphism:



https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Although both m and f are probability presentation morphisms, fm is not. However, given the previous discussion, this difficulty cannot be avoided; in fact, it is easily seen that arbitrary maps between finite probability spaces can be decomposed as in the above example, so forming the compositive closure of the probability presentation morphisms as just defined is not a feasible option.

Thus, one has to accommodate the fact that probability spaces and probability presentation morphisms form a proper precategory **Prob**. The obvious forgetful functor (the 'outcome space functor') is denoted by $Out : \mathbf{Prob} \to \mathbf{Set}$.

Since we only consider countable outcome spaces, it is sufficient to check the morphism condition pointwise; that is, we have the following criterion.

Lemma 3.3. Let $(\Omega_1, p_1), (\Omega_2, p_2)$ be probability presentations, and let $f : \Omega_1 \to \Omega_2$ be a map. Then $f : (\Omega_1, p_1) \to (\Omega_2, p_2)$ is a morphism iff $P(f[\Omega_1]) > 0$ and $p_2(\omega)/P_2(f[\Omega_1]) = P_1(f^{-1}[\{\omega\}])$ for each $\omega \in f(\Omega_1)$.

Remark 3.4. In order to study the interconnection of continuous space and continuous time probabilistic automata, a more general class of probability spaces (with morphisms defined essentially as above) has been considered in Mateus (2000). This class includes probability spaces over Euclidean spaces, thus covering all cases of practical interest. The restriction to the countable case in the above definitions has been chosen largely in the interests of readability.

4. Generative discrete probabilistic automata

Reflecting the role played by randomisation in concurrent and distributed systems, several new models of probabilistic processes and automata have recently emerged, including generative probabilistic automata (van Glabbeek *et al.* 1995), probabilistic automata with non-determinism (Condon *et al.* 1998) and probabilistic I/O automata (Wu *et al.* 1997). We only treat the case of generative probabilistic automata here, although most results should be adaptable to other settings. As discussed in van Glabbeek *et al.* (1995), generative probabilistic automata constitute a more general and useful model than the classical reactive model (Rabin 1963), since, in addition to the relative probabilities of transitions under the same action, they contain information about the relative probabilities of transitions under different actions. For further details and comparisons of these models, see van Glabbeek *et al.* (1995), where generative probabilistic automata are used as a semantic model of probabilistic concurrent programs. In the more restrictive form of (a simplified type of) decision trees, generative probabilistic automata appear in machine learning and knowledge representation (Pinto *et al.* 2000).

From the point of view of concurrency theory, the study of composition and interaction of automata is fundamental, whereas from the point of view of classical Control Theory, realisation and minimisation are the main points of attention. During the 70's, substantial work was undertaken to express realisation of automata within a categorical framework; the motto in those days was 'minimisation is universal'. However, as discussed in Arbib and Manes (1974), the analogous task has turned out to be extremely hard in the probabilistic case.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

In more recent research, especially in concurrency (Winskel and Nielsen 1995), the motto is 'combination is universal'; this is, in fact, part of what can be achieved by means of precategorical methods.

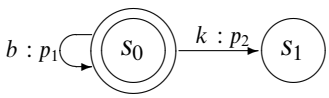
We explicitly recall the following definitions.

Definition 4.1. A generative probabilistic automaton is a tuple $M = (A, S, \delta)$ where A is a countable set with distinguished element $\perp$, S is a countable set with distinguished element i, and $\delta = (\delta^s)_{s \in S} = ((T^s, p^s))_{s \in S}$ is a family of probability presentations, where $T^s \subset A \times S$ for all s, such that

(i) $(\perp, t) \in T^s$ iff $t = s$;
(ii) $p^s(a, t) > 0$ for all $(a, t) \in T^s$.

The elements of A are called *actions*, the distinguished element the *idle action*. The elements of S are called *states*; the distinguished element is the *initial* state. T^s is the set of *possible transitions* in s. Finally, δ^s is the *transition probability distribution* in s.

Example 4.2. A *monotonous broadcaster* is a generative probabilistic automaton that constantly *broadcasts a message* (b) with probability p_1 until it *kaputs* (k) with probability p_2. Between broadcasts and if the system does not fail, the machine does nothing ($\perp$), with probability $p_3 = (1 - p_1 - p_2) > 0$:



Now that a notion of morphism between probability spaces has been established, the classical definition of a morphism between (non-probabilistic) automata can be adapted naturally.

Definition 4.3. A morphism $f : M_1 \to M_2$ between probabilistic automata $M_1 = (A_1, S_1, \delta_1)$ and $M_2 = (A_2, S_2, \delta_2)$ is a pair $(\overline{f}, \underline{f})$, where $\overline{f} : A_1 \to A_2$ and $\underline{f} : S_1 \to S_2$ are maps preserving distinguished elements, and $(\overline{f} \times \underline{f})$ restricts to a probability presentation morphism $\delta_1^s \to \delta_2^{\underline{f}(s)}$ for each $s \in S_1$.

By the remarks at the beginning of Section 2, probabilistic automata and their morphisms form a (proper) precategory, which we denote by **PA**.

We adapt several classical notions to the setting of probabilistic automata in the following definitions.

Definition 4.4. Let M be a probabilistic automaton. We write $s \to^a t$ if $(a, t) \in T^s$. For $w \in A^*$, we inductively define $\to_*^w$ in the obvious way. A state t is called *accessible from* s if $s \to_*^w t$ for some $w \in A^*$. A state is called *accessible* if it is accessible from the initial state.

Definition 4.5. Let M be a probabilistic automaton. We call M

(i) *accessible* if every state of M is accessible;
(ii) *acyclic* if for each $s \in S$, $s \to^w s$ implies $w \in \{\perp\}^*$;

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

(iii) *without merges* if, for $s_1, s_2, t \in S$ and $a, b \in A$, we have $s_1 \to^a t$ and $s_2 \to^b t$ imply $s_1 = s_2$ and $a = b$.

A *decision tree* is an accessible and acyclic probabilistic automaton without merges.

The full subprecategories of **PA** spanned by all accessible automata and all decision trees are denoted by **AcPA** and **DT**, respectively.

Decision trees 'remember' the history of the automaton in each state; that is, there is an injection from the paths to the states preserving the transition relation. Automata of this kind are relevant, for instance, in the semantics of the probabilistic situation calculus (Pinto *et al.* 2000).

We now proceed to develop the theory of precategories.

5. Functor graphs

It is well known that the category **Cat** of small categories and functors is Cartesian closed (Adámek *et al.* 1990); the associated power objects are the functor categories. Similarly, the category **Precat** of small precategories and functors is Cartesian closed, and the power objects in this category do indeed generalise functor categories; thus, we obtain the correct definition of a natural transformation in the context of precategories. For the more general setting of multiplicative graphs (*cf.* Remark 2.1), a description of power objects can already be found in Coppey (1976), although the issue seems to have caused a certain amount of confusion (compare, for example, the various definitions in Coppey (1976) and Ehresmann (1965)). The even more general case of composition graphs is discussed in Schröder (1999).

Definition 5.1. Let $F, G : \mathbf{A} \to \mathbf{B}$ be functors between precategories. A *natural transformation* from F to G is a family $\nu = (\nu_f)$, indexed over all morphisms f in $\mathbf{A}$, where $\nu_f : FA \to GB$ is a morphism in $\mathbf{B}$ for each $f : A \to B$, such that, for each $f : A \to B$, $Gf\nu_A$ and $\nu_B Ff$ are defined in $\mathbf{B}$ and equal to ν_f. The *functor precategory* $[\mathbf{A}, \mathbf{B}]$ is defined as the precategory with functors from $\mathbf{A}$ to $\mathbf{B}$ as objects, and natural transformations between these as morphisms. The composite $\mu\nu$ of two natural transformations $\mu : G \to H$ and $\nu : F \to G$ is defined in $[\mathbf{A}, \mathbf{B}]$ iff $\mu_f \nu_A$ (equivalently: $\mu_B \nu_f$) is defined in $\mathbf{B}$ for each $f : A \to B$ in $\mathbf{A}$; in this case, $(\mu\nu)_f = \mu_f \nu_A = \mu_B \nu_f$.

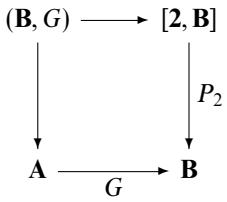
Thus defined, $[\mathbf{A}, \mathbf{B}]$ is in fact a precategory. We omit the verification that the evaluation map $ev : [\mathbf{A}, \mathbf{B}] \times \mathbf{A} \to \mathbf{B}$ that sends (ν, f) to ν_f does indeed constitute a $_ \times \mathbf{A}$-couniversal arrow. Note, however, that $[\mathbf{A}, \mathbf{B}]$ is essentially uniquely determined by this property.

The comma construction

The notion of a functor precategory now leads unambiguously to a generalisation of Lawvere's comma construction to precategories. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories. Then the *comma precategory* $(\mathbf{B}, G)$ is defined by the following pullback

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

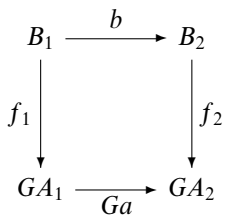
diagram:



Here, **2** denotes the category $1 \to 2$, so that $[\mathbf{2}, \mathbf{B}]$ is the precategory analogue of the arrow category; that is, the objects of $[\mathbf{2}, \mathbf{B}]$ are the morphisms of **B**. The arrow $P_2 : [\mathbf{2}, \mathbf{B}] \to \mathbf{B}$ in the diagram is the projection that sends $f : A \to B$ (that is, an object of $[\mathbf{2}, \mathbf{B}]$) to B.

We will now give a more explicit description.

The objects of $(\mathbf{B}, G)$ are *G-structured morphisms* consisting of an object A in **A** and a morphism $f : B \to GA$ in **B**; as usual, such an object is denoted somewhat inaccurately as $f : B \to GA$. Given two such G-structured morphisms $f_i : B_i \to GA_i$, $i = 1, 2$, a morphism from $f_1 : B_1 \to GA_1$ to $f_2 : B_2 \to GA_2$ can be regarded as a pair $v = (b, a)$, where $b : B_1 \to B_2$ and $a : A_1 \to A_2$ are morphisms in **B** and **A**, respectively, such that $f_2 b$ and $(Ga)f_1$ are defined and equal:

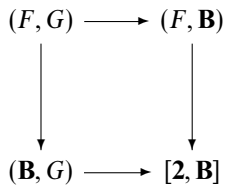


In particular, there is a faithful functor $(\mathbf{B}, G) \to \mathbf{B} \times \mathbf{A}$ called the *underlying functor*.

If $\mu = (d, c)$ is a further morphism from f_2 to $f_3 : B_3 \to GA_3$, then μv is defined in $(\mathbf{B}, G)$ iff $(d, c)(b, a)$ is defined in $\mathbf{B} \times \mathbf{A}$ and (db, ca) is a morphism from f_1 to f_3 in $(\mathbf{B}, G)$ (for the latter, it suffices that one of the terms $f_3(db)$ and $G(ca)f_1$ be defined); that is, the underlying functor of $(\mathbf{B}, G)$ is initial. In particular, any *concrete* graph morphism between two comma precategories (of either of the two forms just given) over $\mathbf{A} \times \mathbf{B}$, that is, any graph morphism that commutes with the underlying functors, is already an initial functor.

Dually, we have a comma precategory $(G, \mathbf{B})$ of G-costructured morphisms $f : GA \to B$.

Remark 5.2. More generally, one defines the comma precategory (F, G), where $G : \mathbf{A} \to \mathbf{B}$ and $F : \mathbf{C} \to \mathbf{B}$ are functors between precategories, by the pullback diagram



where the right-hand and bottom arrows denote the projections from the defining pullback diagrams for $(F, \mathbf{B})$ and $(\mathbf{B}, G)$, respectively. The objects of (F, G) are morphisms $f : FC \to GB$ that are simultaneously G-structured and F-costructured.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

6. Adjunctions

The generalisation of the concept of adjunction to the setting of precategories is not quite as unambiguous as the situation with functor precategories, and indeed there are at least two reasonable notions of adjointness for precategories that coincide with the usual definition in the case of categories. All of the classical definitions pose severe difficulties in this context.

The definitions via universal arrows and adjoint situations (Adámek *et al.* 1990), respectively, are not easily translated, since it is not clear in what sense the involved equations should be required to hold (weakly, strongly, and so on). Moreover, it would be desirable to have an equivalence between these two definitions; this apparently cannot be obtained with any of the naive generalisations, since the respective proofs use compositions that do not occur in the said equations.

On the other hand, a definition via natural isomorphy of hom-'functors' (which in the more general setting turn into mere graph morphisms into the category of sets and partial maps) does lead to a very strong concept, from which universal arrows as well as adjoint situations can be derived. This approach, however, has the crucial disadvantage that the existence of a thus defined adjoint functor between two composition graphs implies a rather strong form of associativity, which is not in general available in precategories.

An alternative definition that is strong enough to produce universal arrows, adjoint situations, and even an equivalence between these concepts (when properly defined), and which at the same time is general enough to be applied to arbitrary precategories (in fact, multiplicative graphs), is proposed in Schröder (1999). It is based on the lesser-known observation that two functors $G : \mathbf{A} \to \mathbf{B}$ and $F : \mathbf{B} \to \mathbf{A}$ between categories $\mathbf{A}$ and $\mathbf{B}$ form an adjoint pair iff there exists a concrete isomorphism between $(F, \mathbf{A})$ and $(\mathbf{B}, G)$ (Herrlich and Strecker 1979). This is in fact visually so similar to the definition by hom-functors that one might be tempted to replace the old motto 'do everything via hom-functors' by 'do everything via the comma construction'.

We state the definition explicitly.

Definition 6.1. A functor $G : \mathbf{A} \to \mathbf{B}$ between precategories is called *adjoint* if there exist a functor $F : \mathbf{B} \to \mathbf{A}$ and a concrete isomorphism $(\mathbf{B}, G) \cong (F, \mathbf{A})$. In that case, F is called a *left adjoint* of G.

In order to state the characterisation theorem, we will need a certain amount of terminology.

Definition 6.2. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories. A G-structured morphism $f : B \to GA$ is called *universal* if each G-structured morphism $k : B \to GC$ admits a unique extension $k^{\#} : A \to C$ such that $k = (Gk^{\#})f$, and *left G-composable* if $(Gg)f$ is defined for each morphism $g : A \to C$ in $\mathbf{A}$. $f : B \to GA$ *G-induces composability* if, whenever $g : A \to C$ and $h : C \to D$ are morphisms in $\mathbf{A}$ such that $(Gh)((Gg)f)$ is defined, then hg is defined (in other words: $(Gh)((Gg)f)$ and $(G(hg))f$ are strongly equal).

Dually, we have, for example, the notion of a *right composable* costructured arrow. We will implicitly regard G-structured morphisms as being $id_{\mathbf{B}}$-costructured. If G is an identity

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

functor, it will be left out in the notation; for example, we write 'left composable' instead of 'left $id_{\mathbf{A}}$-composable'. If $\mu : F \to G$ is a natural transformation, $\mu_A : FA \to GA$ is left G-composable and right F-composable for each object A.

Example 6.3. Let (Ω_1, p_1) and (Ω_2, p_2) be probability presentations. Then probabilistically surjective (that is, measure preserving) morphisms $f : (\Omega_1, p_1) \to (\Omega_2, p_2)$ are left composable in **Prob**: if $g : (\Omega_2, p_2) \to (\Omega_3, p_3)$ is a further (arbitrary) morphism, $g^{-1}[gf[\Omega_1]] \supset f[\Omega_1]$ implies $P(g^{-1}[gf[\Omega_1]]) = 1$; since $gf[\Omega_1] \subset g[\Omega_2]$, we thus have $P(gf[\Omega_1]) = P(gf[\Omega_1]|g[\Omega_2])P(g[\Omega_2]) = P(g^{-1}[gf[\Omega_1]])P(g[\Omega_2]) = P(g[\Omega_2])$. Hence, $P_1((gf)^{-1}[\{\omega\}]) = P_1(f^{-1}[g^{-1}[\{\omega\}]]) = P_2(g^{-1}[\{\omega\}]) = p_3(\omega)/P_3(g[\Omega_2]) = p_3(\omega)/P_3(gf[\Omega_1])$ for $\omega \in gf[\Omega_1]$, so that gf is a morphism by Lemma 3.3.

Now the announced equivalent formulations of adjointness read as follows.

Theorem 6.4. (Schröder 1999) Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories. Equivalent are

(i) G is adjoint.
(ii) For each object B in $\mathbf{B}$, there exists a right composable and left G-composable universal arrow $\eta_B : B \to GA_B$ that G-induces composability.
(iii) There exists a functor $F : \mathbf{B} \to \mathbf{A}$ and natural transformations $\eta : id_{\mathbf{B}} \to GF$ and $\varepsilon : FG \to id_{\mathbf{A}}$ (called the *unit* and the *co-unit*, respectively) such that for each B, we have $\eta_B : B \to GFB$ is left G-composable and G-induces composability, and such that $(G\varepsilon)\eta_G = id_G$ and $\varepsilon_F F\eta = id_F$ are defined.

Proof. (*i*) $\Rightarrow$ (*iii*): Let $\Phi : (\mathbf{B}, G) \to (F, \mathbf{A})$ be a concrete isomorphism. For $f : B \to C$ in $\mathbf{B}$, let $\eta_f : B \to GFC$ be the image of the F-costructured arrow $Ff : FB \to FC$ under Φ^{-1}. Then the η_f form a natural transformation $\eta : id_{\mathbf{B}} \to GF$: for $f : B \to C$ in $\mathbf{B}$, we have a morphism (f, id_{FC}) from $Ff : FB \to FC$ to $id_{FC} : FC \to FC$ in $(F, \mathbf{A})$, which (by concreteness of Φ^{-1}) is mapped to a morphism (f, id_{FC}) from $\eta_f : B \to GFC$ to $\eta_C : C \to GFC$ in $(\mathbf{B}, G)$; thus, $\eta_C f = \eta_f$, as required.

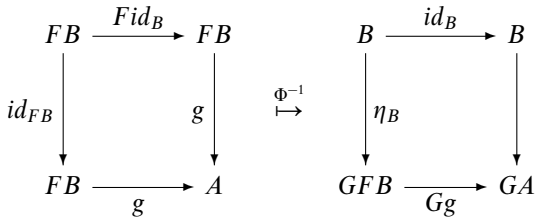
$$\begin{array}{ccc} FB & \xrightarrow{Ff} & FC \\ {\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle id_{FC}} \\ FC & \xrightarrow[id_{FC}]{} & FC \end{array} \quad \overset{\Phi^{-1}}{\mapsto} \quad \begin{array}{ccc} B & \xrightarrow{f} & C \\ {\scriptstyle \eta_f}\downarrow & & \downarrow{\scriptstyle \eta_C} \\ GFC & \xrightarrow[Gid_{FC}]{} & GFC \end{array}$$

It can be seen analogously that $GFf\eta_B = \eta_f$.

Moreover, for $g : FB \to A$ in $\mathbf{A}$, the morphism (id_B, g) from $id_{FB} : FB \to FB$ to $g : FB \to A$ in $(F, \mathbf{A})$ is mapped to a morphism (id_B, g) with domain $\eta_B : B \to GFB$ in

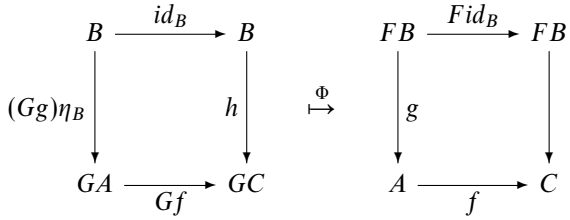
https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

$(\mathbf{B}, G)$ under Φ^{-1}:



Thus, $Gg\eta_B : B \to GA$ is defined (and its image under Φ is $g : FB \to A$); that is, η_B is left G-composable.

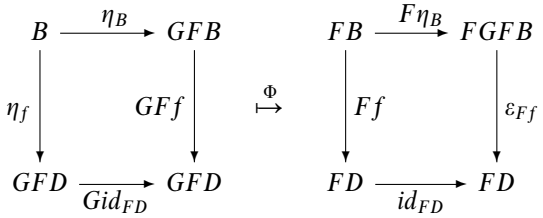
Finally, if $g : FB \to A$ and $f : A \to C$ are morphisms in $\mathbf{A}$ such that $(Gf)((Gg)\eta_B) =: h$ is defined, then (id_B, f) is a morphism from $(Gg)\eta_B : B \to GA$ to $h : B \to GC$ in $(\mathbf{B}, G)$. By the above, its image under Φ is a morphism (id_B, f) with domain $g : FB \to A$:



In particular, fg is defined; that is, η_B G-induces composability.

Analogously, a natural transformation $\varepsilon : FG \to id_{\mathbf{A}}$ is defined by taking $\varepsilon_f : FGA \to C$ as the image under Φ of $Gf : GA \to GC$ for each morphism $f : A \to C$ in $\mathbf{A}$.

To show that $\varepsilon_F F\eta = id_F$, let $f : B \to D$ be a morphism in $\mathbf{B}$. By naturality of η, we have a morphism (η_B, id_{FD}) from $\eta_f : B \to GFD$ to $GFf : GFB \to GFD$ in $(\mathbf{B}, G)$, which is mapped under Φ to a morphism (η_B, id_{FD}) from Ff to ε_{Ff} in $(\mathbf{A}, F)$:



That is, $\varepsilon_{Ff} F\eta_B = Ff$, as required. The equation $(G\varepsilon)\eta_G = id_G$ is shown analogously.

(*iii*) $\Rightarrow$ (*ii*): All that remains to be shown is that $\eta : B \to GFB$ is G-universal for each object B in $\mathbf{B}$. To this end, let $f : B \to GA$ be a G-structured morphism. Then we have strong equations

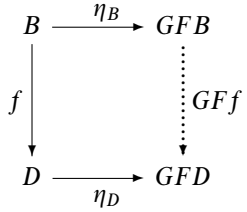
$$(G(\varepsilon_A Ff))\eta_B = (G\varepsilon_A)((GFf)\eta_B) = (G\varepsilon_A)(\eta_{GA} f) = ((G\varepsilon_A)\eta_{GA})f = f.$$

Thus, $\varepsilon_A Ff$ is defined and extends f. Similarly, $\varepsilon_A F((G\bar{f})\eta_B) = \bar{f}$ is defined for each F-costructured morphism $\bar{f} : FB \to A$; hence $(G\bar{f})\eta_B = f$ implies $\bar{f} = \varepsilon_A Ff$, that is, the factorisation is unique.

(*ii*) $\Rightarrow$ (*i*): Define F on objects by $FB = A_B$ for each B, and given $f : B \to D$ in $\mathbf{B}$, let

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

$Ff : FB \to FD$ be the unique factorisation of $\eta_D f$ (which is defined by assumption on η):



It is obvious that F preserves identities. If gf is defined for morphisms $f : B \to D$ and $g : D \to E$ in $\mathbf{B}$, then $\eta_E(gf)$ is defined and hence so also are the terms

$$\begin{aligned} \eta_E(gf) = (\eta_E g)f &= ((GFg)\eta_D)f \\ &= (GFg)(\eta_D f) \\ &= (GFg)((GFf)\eta_B) = (G((Fg)Ff))\eta_B. \end{aligned}$$

Thus, $(Fg)Ff = F(gf)$, that is, $F : \mathbf{B} \to \mathbf{A}$ is indeed a functor.

Now let (the object part of) $\Phi : (\mathbf{B}, G) \to (F, \mathbf{A})$ assign to each G-structured morphism $f : B \to GA$ its unique factorising morphism $f^\# : FB \to A$; this is bijective by assumption. We show that Φ 'is' a full concrete graph morphism. As explained in Section 5, this implies that Φ is a concrete functor; full concrete functors that are bijective on objects are concrete isomorphisms.

Explicitly, we have to show that (b, a) is a morphism from $f : B \to GA$ to $h : D \to GC$ in $(\mathbf{B}, G)$ iff (b, a) is a morphism from $f^\# : FB \to A$ to $h^\# : FD \to C$ in $(F, \mathbf{A})$, that is, that definedness of $(Ga)f = hb$ is equivalent to definedness of $af^\# = h^\# Fb$. Now the latter is equivalent to definedness of $(Ga)((Gf^\#)\eta_B) = (Gh^\#)((GFb)\eta_B)$ by virtue of the fact that η_B G-induces composability and by unicity of the factorisation. This, in turn, is equivalent to the definedness of $(Ga)f = hb$, since we have strong equations $(Ga)((Gf^\#)\eta_B) = (Ga)f$ and $(Gh^\#)((GFb)\eta_B) = (Gh^\#)(\eta_D b) = ((Gh^\#)\eta_D)b = hb$. □

Remark 6.5. In the notation of Theorem 6.4, the $\varepsilon_A : FGA \to A$ are (by symmetry) left composable and right F-composable and F-co-induce composability.

Remark 6.6. Using the construction of comma precategories via pullbacks, it is easily seen that, for any functors K and H with suitable domains and codomains, the comma precategories (K, GH) and (FK, H) (*cf.* Remark 5.2) are concretely isomorphic if G is adjoint to F. This allows, in particular, an easy proof of the fact that adjoint functors compose: if $G_1 : \mathbf{A} \to \mathbf{B}$ is adjoint to F_1 and $G_2 : \mathbf{B} \to \mathbf{C}$ is adjoint to F_2, then $G_2 G_1$ is adjoint to $F_1 F_2$ since

$$(\mathbf{C}, G_2 G_1) \cong (F_2, F_1) \cong (F_1 F_2, \mathbf{A}).$$

Example 6.7. In Schröder (2001), the theory of monads is extended to precategories; in particular, it is shown that the Kleisli construction can be performed in this setting and indeed gives rise to a smallest realisation of a monad by an adjunction. For example, one has a power set monad $(X, A) \mapsto (\mathfrak{P}X, \mathfrak{P}A)$ on **fRel**(**1**). The associated Kleisli precategory can be described as the following precategory **frRel**(**1**): Objects are as in **fRel**(**1**), and

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

morphisms $\rho : (X, A) \to (Y, B)$ are *full relations*, that is, relations $\rho \subset X \times Y$ (as usual, we write $x\rho y$ for $(x, y) \in \rho$) such that

(i) $x \in A$ and $x\rho y$ imply $y \in B$;
(ii) whenever $x\rho y$ implies $y \in B$ for some $x \in X$, there exists $x' \in A$ such that, for each $y \in Y$, $x\rho y$ iff $x'\rho y$.

(The composite of two full relations is defined iff the composite relation is full.) The details of the theory of monads are beyond the scope of this paper. However, it is comparatively easy to check that the inclusion **fRel(1)** $\hookrightarrow$ **frRel(1)** is in fact co-adjoint. More specifically, given an object (X, A) in **frRel(1)**,

$$\tau : (\mathfrak{P}X, \mathfrak{P}A) \to (X, A),$$

where $B\tau x \Leftrightarrow x \in B$, is a co-universal arrow for (X, A) with the required properties.

Example 6.8. A very simple example of a coreflective subprecategory (that is, a co-adjoint embedding) can be found in the precategory **Prob** of countable probability spaces as defined in Section 3: let $\mathbf{Prob}_{>\mathbf{0}}$ denote the full subprecategory of **Prob** spanned by the spaces with the property that each element has a positive probability. Then, given a space (Ω, p), the embedding of the subspace with carrier set $\{\omega \in \Omega \mid p(\omega) > 0\}$ is a $\mathbf{Prob}_{>\mathbf{0}}$-coreflection of (Ω, p).

Example 6.9. Now that a notion of adjointness for precategories is available, a reasonable definition of a precategory **A** *with products* might require that the diagonal functors $\Delta : \mathbf{A} \to [\mathbf{I}, \mathbf{A}]$ (obtained by applying the co-universal property of the functor precategory to the projection $\mathbf{A} \times \mathbf{I} \to \mathbf{A}$; more explicitly, ΔA is the constant functor associated to an object A in **A**) are co-adjoint for each small discrete category **I** (Mac Lane 1972); by Theorem 6.4, this is equivalent to the existence of product objects $\Pi_I A_i$ and projections $\pi_i : \Pi_I A_i \to A_i$, for each family $(A_i)_I$ of **A** objects, such that the π_i are left and right composable, such that definedness of $(\pi_i f)g$ for all i implies definedness of fg, and such that the usual universal property holds.

Given this definition of products (we shall shortly see that a weaker version is appropriate in many cases), one can construct a free product completion $\Pi\mathbf{A}$ of a precategory **A** in literally the same way as for categories (Kock 1995): The objects of **A** are set-indexed families $(A_i)_I$ of **A**-objects, and the morphisms $(A_i)_I \to (B_j)_J$ are pairs $(\phi, (f_j)_J)$, where $\phi : J \to I$ is a map and $f_j : A_{\phi(j)} \to B_j$ is an **A**-morphism for each j; composition is defined in the obvious way. $\Pi\mathbf{A}$ indeed has products in the sense just defined; equivalently: the obvious embedding $\Pi\mathbf{A} \hookrightarrow \Pi\Pi\mathbf{A}$ is co-adjoint.

Natural isomorphisms

In order to use universal properties, that is, adjunctions, as definitions for constructions, we need at least the essential uniqueness of the left adjoint as in category theory. Since 'essential' means 'up to natural isomorphy' here, this requires a definition of the notion of natural isomorphism. By the philosophy outlined above, the natural definition will make use of comma graphs.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Definition 6.10. Let $F, G : \mathbf{A} \to \mathbf{B}$ be functors between precategories. F and G are called *naturally isomorphic* if there exists a concrete isomorphism $(\mathbf{B}, F) \cong (\mathbf{B}, G)$.

We will see below that natural isomorphisms can be equivalently described in terms of mutually inverse natural transformations with additional requirements similar to those in Theorem 6.4. Given this definition, it is immediate that left adjoints are unique up to *natural co-isomorphy*; more precisely, we have the following proposition.

Proposition 6.11. Let $G : \mathbf{A} \to \mathbf{B}$ be an adjoint functor between precategories with left adjoint F. Then a functor $F' : \mathbf{B} \to \mathbf{A}$ is a left adjoint of G iff F' is naturally co-isomorphic to F.

The somewhat strange notion of co-isomorphism is indeed necessary here, since the concept of natural isomorphism fails to be self-dual in the general case. For the sake of simplicity, we shall restrict the further exposition to 'reasonable' precategories. In this setting, natural isomorphisms will turn out to admit an alternative description via mutually inverse natural transformations without additional properties; by virtue of that description, natural isomorphy is in many concrete cases seen to be self-dual.

Definition 6.12. A morphism f in a precategory is called *invertible* if it has an *inverse*, that is, a morphism g such that $gf = id$ and $fg = id$ are defined. A precategory $\mathbf{A}$ is called *admissible* if invertible morphisms in $\mathbf{A}$ are left composable.

(It is easily seen that inverses are unique.)

In admissible precategories, invertibility of natural transformations can be checked pointwise (as in the case of categories).

Proposition 6.13. Let $\mathbf{A}$ and $\mathbf{B}$ be precategories, with $\mathbf{B}$ admissible. Then $[\mathbf{A}, \mathbf{B}]$ is admissible, and a natural transformation μ in $[\mathbf{A}, \mathbf{B}]$ is invertible iff μ is object-wise invertible, that is, if μ_A is invertible for each object A in $\mathbf{A}$.

Proof. It is clear that invertibility of μ implies invertibility and, hence, left composability of the μ_A. By the definition of composition in functor precategories, object-wise left composable natural transformations are left composable; hence, $[\mathbf{A}, \mathbf{B}]$ is admissible.

Conversely, if $\mu : F \to G$ is a natural transformation and each μ_A has an inverse ν_A, then we can define $\nu : G \to F$ by $\nu_f := F f \nu_A$ for $f : A \to C$ in $\mathbf{A}$ (which is defined because ν_A is invertible). ν is natural: by repeated application of the weak associative law, $\nu_C G f = \nu_C((G f \mu_A)\nu_A) = \nu_C((\mu_C F f)\nu_A) = F f \nu_A = \nu_f$. By object-wise left composability, both μ and ν are left composable; thus, $\mu\nu$ and $\nu\mu$ are defined. $\mu\nu = id_G$ and $\nu\mu = id_F$ follow from objectwise equality. □

As indicated above, we have the following equivalent characterisation of natural isomorphy.

Theorem 6.14. Let $F, G : \mathbf{A} \to \mathbf{B}$ be functors between precategories, and let $\mathbf{B}$ be co-admissible. Then F and G are naturally isomorphic iff there exists an invertible natural transformation $\mu : F \to G$.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Thus, in the situation of the Theorem, an invertible natural transformation $\mu : F \to G$ will be called a *natural isomorphism.*

Proof. Let $\Phi : (\mathbf{B}, F) \to (\mathbf{B}, G)$ be a concrete isomorphism (so that F and G are naturally isomorphic). For $f : A \to C$ in $\mathbf{A}$, let $\mu_f : FA \to GC$ be the image of $Ff : FA \to FC$ under Φ. This defines a natural transformation $\mu : F \to G$; the proof of the naturality equations for μ is analogous to that for η in the proof of Theorem 6.4.

To show that μ is invertible, it suffices that μ is objectwise invertible (Proposition 6.13). Let $v_A : GA \to FA$ be the image of $id_{GA} : GA \to GA$ under Φ^{-1}. Application of Φ^{-1} to the morphism (μ_A, id_A) from $\mu_A : FA \to GA$ to $id_{GA} : GA \to GA$ in $(\mathbf{B}, G)$ yields $v_A \mu_A = id_{FA}$:

$$\begin{array}{ccc} FA & \xrightarrow{\mu_A} & GA \\ {\scriptstyle \mu_A}\downarrow & & \downarrow{\scriptstyle id_{GA}} \\ GA & \xrightarrow[Gid_A]{} & GA \end{array} \quad \overset{\Phi^{-1}}{\mapsto} \quad \begin{array}{ccc} FA & \xrightarrow{\mu_A} & GA \\ {\scriptstyle id_{FA}}\downarrow & & \downarrow{\scriptstyle v_A} \\ FA & \xrightarrow[Fid_A]{} & FA \end{array}$$

Analogously, it can be shown that $\mu_A v_A = id_{GA}$.

Conversely, let $\mu : F \to G$ be invertible with inverse v. We define a concrete functor $\Phi : (\mathbf{B}, F) \to (\mathbf{B}, G)$ on objects as mapping F-structured morphisms $f : B \to FA$ to $\mu_A f : B \to GA$ (which is defined because $\mathbf{B}$ is co-admissible). As explained in Section 5, it suffices to show that this mapping extends to a concrete graph morphism. To this end, let (b, a) be a morphism from $f : B \to FA$ to $g : D \to FC$ in $(\mathbf{B}, F)$, that is, $(Fa)f = gb$. Then $\mu_C(gb)$ is defined and equal to $(\mu_C g)b$ on the one hand and to $\mu_C((Fa)f) = (\mu_C Fa)f = ((Ga)\mu_A)f = (Ga)(\mu_A f)$ on the other hand, so (b, a) is a morphism from $\mu_A f : B \to GA$ to $\mu_C g : D \to GC$ in $(\mathbf{B}, G)$, as required.

Analogously, one can construct a concrete functor $\Psi : (\mathbf{B}, G) \to (\mathbf{B}, F)$ using v; it is clear that Φ and Ψ are mutually inverse. □

Remark 6.15. In the case of arbitrary precategories, the characterisation of natural isomorphisms becomes slightly more complicated: F and G are naturally isomorphic if and only if there exists an invertible natural transformation $\mu : F \to G$ such that μ and its inverse are objectwise right composable.

The notion of isomorphy of objects can be defined via natural isomorphy: it is sensible to call two objects in a precategory $\mathbf{A}$ *isomorphic* if the associated functors $1 \to \mathbf{A}$ (where $\mathbf{1}$ denotes the terminal category) are naturally isomorphic. Informally speaking, two objects A and B are isomorphic (co-isomorphic) in this sense if the morphisms with codomain (domain) A are 'the same' as the morphisms with codomain (domain) B.

All of the examples of precategories presented so far are admissible and co-admissible (so that natural isomorphisms and co-isomorphisms coincide), except one, as follows.

Example 6.16. Consider the 'bad' precategory **Suf** from Example 2.10. A morphism $f : (X, A) \to (Y, B)$ in **Suf** is left composable iff $f[A] \subset B$, and right composable iff f is total. Thus, invertible morphisms in $\mathbf{A}$ are right composable, but need not be left composable; that is, **Suf** is co-admissible, but not admissible. In fact, it is easily

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

verified that the objects $(\{0\},\{0\})$ and $(\{0\},\varnothing)$ are naturally isomorphic, but not naturally co-isomorphic in the sense of Remark 6.15.

7. Coreflections of probabilistic automata

To establish a relationship between probabilistic automata and decision trees, we first show that the reduction of an automaton to its accessible part is universal. Since inaccessible states are usually redundant, this reduction procedure is of obvious pragmatic importance; for instance, it is commonly performed as a preliminary step in minimisation or state exploration.

In the theoretical framework built up so far, this construction can be formally captured in the following statement.

Proposition 7.1. **AcPA** is a coreflective subprecategory of **PA**.

The proof needs the following observation.

Lemma 7.2. Let $f : M_1 \to M_2$ be a morphism between probabilistic automata. If s is accessible in M_1, then $\underline{f}(s)$ is accessible in M_2.

Proof of Proposition 7.1. Let $E : \mathbf{AcPA} \hookrightarrow \mathbf{PA}$ denote the inclusion functor; we have to show that E is co-adjoint.

To this end, let $M = (A, S, \delta)$ be a probabilistic automaton. Let S_a be the set of accessible states of M, and denote the inclusion $S_a \hookrightarrow S$ by m. Then $\varepsilon = (id_A, m) : M_a = (A, S_a, \delta) \to M$ is a couniversal arrow with the required properties: M_a is accessible by construction. Since the restriction ε^s of $id_A \times m$ to $\delta^s \to \delta^{m(s)}$ is an identity for each s, ε is left composable and right E-composable and E-co-induces composability. Finally, if M_1 is an accessible automaton and $f : M_1 \to M$ is a morphism, then, by the above lemma, f restricts to a pair of maps $M_1 \to M_a$, which is a morphism since the ε^s are identities. □

In a slightly more complicated fashion, we can now (co-)universally construct a decision tree from an accessible probabilistic automaton.

Definition 7.3. Let M be a probabilistic automaton. A *proper transition* in M is defined as a triple $(s,a,t) \in S \times (A \setminus \{\perp\}) \times S$ such that $s \to^a t$. A *path* in M is a sequence $w = (s_1,a_1,t_1)\ldots(s_n,a_n,t_n)$ of proper transitions such that if $n > 0$, then $s_1 = i$ and $t_k = s_{k+1}$, $k = 1\ldots n-1$. We denote the set of all paths in M by pth_M. A morphism $f : M_1 \to M_2$ in **PA** induces a map $f_{pth} : pth_{M_1} \to pth_{M_2}$ defined inductively by

$$\begin{aligned} f_{pth}(\varepsilon) &= \varepsilon; \\ f_{pth}(w(s,a,t)) &= \begin{cases} f_{pth}(w) & \text{if } \overline{f}(a) = \perp \\ f_{pth}(w)(\underline{f}(s), \overline{f}(a), \underline{f}(t)) & \text{otherwise} \end{cases}. \end{aligned}$$

Finally, we define $lst_M : pth_M \to S$ as follows:

$$\begin{aligned} lst_M(\varepsilon) &= i; \\ lst_M(w(s,a,t)) &= t. \end{aligned}$$

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Remark 7.4. Note that lst_M is surjective iff M is accessible and injective iff the accessible coreflexion of M is acyclic and without merges. Therefore, lst_M is bijective iff M is a decision tree.

Theorem 7.5. **DT** is a coreflective subprecategory of **AcPA**.

Proof. Let $E : \mathbf{DT} \hookrightarrow \mathbf{AcPA}$ denote the inclusion functor, and let $M = (A, S, \delta)$ be an accessible automaton. Then $M_{dt} = (A, pth_M, \delta_{pth})$, where ε is the distinguished element of pth_M,

$$T^w_{pth} = \{(\bot, w)\} \cup \{(a, w(lst(w), a, t)) \mid (lst(w), a, t) \text{ is a proper transition}\},$$

and

$$p^w_{pth}(a, v) = p^{lst(w)}(a, lst(v))$$

is a decision tree, and

$$\varepsilon_M = (id_A, lst_M) : M_{dt} \to M$$

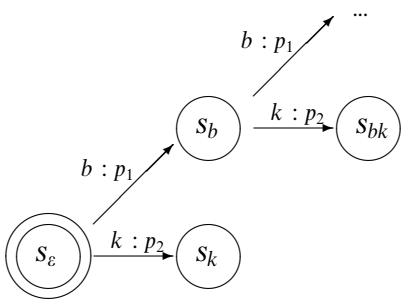
is an E-couniversal arrow with the required additional properties: the latter follow from the fact that the restrictions ε^w_M of $id_A \times lst_M$ to $\delta^w_{pth} \to \delta^{lst(w)}$ are, by construction, isomorphisms. Now let D be a decision tree, and let $f : D \to M$ be a morphism. By Remark 7.4, we have a pair of maps

$$\tilde{f} = (\overline{f}, f_{pth} \circ lst_D^{-1}) : D \to M_{dt},$$

which, again by virtue of the fact that the ε^w_M are isomorphisms, is a morphism; since, obviously, $\varepsilon_M \tilde{f} = f$, we have the required factorisation. Uniqueness is shown inductively. □

Remark 7.6. By Remark 6.6, it follows that **DT** is also a coreflective subprecategory of **PA**.

Example 7.7. Consider the monotonous broadcaster of Example 4.2, and suppose that we require to know at each state how many broadcasts have been emitted so far. This is achieved by forming the **DT**-coreflection as defined above:



(The coreflective arrow to the original monotonous broadcaster maps s_ε, s_b, s_{bb}, and so on, to s_0 and s_k, s_{bk}, s_{bbk}, and so on, to s_1.)

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

8. Weak and partial adjunctions

The symmetry in the definition of adjointness is appealing for both aesthetic and technical reasons. However, in practice one often encounters functors that are 'almost adjoint' in the sense that the universal arrows fail to satisfy the left composability requirement with respect to the functor in question (*cf.* Theorem 6.4). In such situations, the symmetry is missing; for example, the co-units fail to be co-universal.

Thus, there is a need for a weakened notion of adjointness, which, given the theory presented so far, will preferably be defined in terms of comma precategories. Translated into this language, the phenomenon corresponds to an embedding rather than an isomorphism of comma precategories. On such an embedding, two seemingly technical restrictions have to be imposed.

Definition 8.1. Let $\Phi : \mathbf{C} \to (\mathbf{B}, F)$ a functor, where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are precategories and $F : \mathbf{A} \to \mathbf{B}$ is a functor. Φ is called *covering* if every F-structured arrow of the form $Ff : FA_1 \to FA_2$ belongs to the image of Φ, and *semifull* if, whenever $\alpha : \Phi C_1 \to \Phi C_2$ is a morphism of the form $\alpha = (b, id)$ or $\alpha = (id, a)$, there exists $\bar{\alpha} : C_1 \to C_2$ such that $\Phi\bar{\alpha} = \alpha$. A subprecategory of $(\mathbf{B}, F)$ is called *semifull* or *covering* if the associated embedding has the respective property.

(Dually, one has *cocovering* functors $\Phi : \mathbf{C} \to (F, \mathbf{B})$.)

In this terminology, the weakened definition of adjointness reads as follows.

Definition 8.2. A functor $G : \mathbf{A} \to \mathbf{B}$ between precategories is called *weakly adjoint* if there exists a functor $F : \mathbf{B} \to \mathbf{A}$ and a semifull cocovering concrete embedding $(\mathbf{B}, G) \hookrightarrow (F, \mathbf{A})$. In this case, F is called a *weak left adjoint* of G.

The cocovering condition serves to ensure naturality of the unit; the purpose of the semifullness condition is illustrated by the proof of Theorem 6.4. Note that semifullness is quite close to fullness, since every morphism (b, a) in $(\mathbf{B}, F)$ has a factorisation of the form $(b, a) = (id, a)(b, id)$.

Now we have a variant of Theorem 6.4 for weakly adjoint functors.

Theorem 8.3. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories. Then the following are equivalent:

(i) G is weakly adjoint.
(ii) For each object B in $\mathbf{B}$, there exists a right composable G-universal arrow $\eta_B : B \to GA_B$ that G-induces composability.
(iii) There exists a functor $F : \mathbf{B} \to \mathbf{A}$ and natural transformations $\eta : id_{\mathbf{B}} \to GF$ (called the *weak unit*) and $\varepsilon : FG \to id_{\mathbf{A}}$ such that $\eta_B : B \to GFB$ G-induces composability for each B and such that $(G\varepsilon)\eta_G = id_G$ and $\varepsilon_F F\eta = id_F$ are defined.

The proof is analogous to (in fact, a subset of) that of Theorem 6.4; the image (on objects) of the embedding $(\mathbf{B}, G) \hookrightarrow (F, \mathbf{A})$ implicit in Condition (i) consists precisely of the factorisations of G-structured morphisms through the universal arrows of Condition (iii). Note that the notions of adjoint and weakly adjoint functor coincide if the codomain is a category.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Weak left adjoints are 'essentially' unique. In the general case, the meaning of 'essentially' is somewhat more intricate here than in the uniqueness statement for proper left adjoints (Proposition 6.11): in the notation of the above theorem, two semifull cocovering concrete embeddings $\Phi_i : (\mathbf{B}, G) \hookrightarrow (F_i, \mathbf{A})$, $i = 1, 2$, give rise to a concrete isomorphism between semifull cocovering subprecategories of the $(F_i, \mathbf{A})$ rather than between the full comma precategories; on top of that, these subprecategories depend on the given embeddings.

These difficulties disappear under suitable admissibility conditions, as follows.

Theorem 8.4. Let $\mathbf{A}$ and $\mathbf{B}$ be precategories, with $\mathbf{A}$ admissible, and let $G : \mathbf{A} \to \mathbf{B}$ be a weakly adjoint functor with weak left adjoint F and weak unit η. Then a functor $F' : \mathbf{B} \to \mathbf{A}$ is a weak left adjoint of G iff there exists a natural co-isomorphism (that is, an invertible natural transformation) $\mu : F \to F'$ such that $(G\mu)\eta$ is defined.

The definedness condition can, of course, be omitted if $\mathbf{B}$ is co-admissible.

Proof. If F' is a weak left adjoint of G, one has a concrete isomorphism between semifull cocovering subprecategories $\mathbf{K}$ and $\mathbf{L}$ of $(F, \mathbf{A})$ and $(F', \mathbf{A})$, respectively. It is easily checked that such an isomorphism suffices to perform all relevant steps in the (dualised) proof of Theorem 6.14, so we obtain an invertible natural transformation $\mu : F \to F'$, where $\mu_f : FB \to F'C$ is, by construction, a $\mathbf{K}$-object for each $f : B \to C$ in $\mathbf{B}$. Since $\mathrm{Ob}\,\mathbf{K}$ consists of the factorisations of G-structured morphisms, this implies that $(G\mu_f)\eta_B$ is always defined; that is, $(G\mu)\eta$ is defined.

For the reverse implication, all that remains to be shown is that the composition of the given embedding $\Phi : (\mathbf{B}, G) \hookrightarrow (F, \mathbf{A})$ with the concrete isomorphism $\Psi : (F, \mathbf{A}) \cong (F', \mathbf{A})$ induced by μ is cocovering. This follows from the fact that the F-costructured morphisms $\mu_f : FB \to F'C$ (where $f : B \to C$ in $\mathbf{B}$), which are mapped to $F'f : F'B \to F'C$ under Ψ (*cf.* the proof of Theorem 6.14), belong to the image of Φ by the definedness assumption. □

In several standard situations that are associated with adjunctions in the case of categories, the generalisation to precategories leads to weak adjunctions.

Example 8.5. The Eilenberg–Moore realisation of a monad on a precategory in general produces a weak adjunction (Schröder 2001).

Example 8.6. There are good indications that a generalisation of Kock–Zöberlein doctrines (Kock 1995) to precategories will involve *weakly* co-adjoint units in the associated algebras (Schröder 1999). Since categories with products can be described as algebras of the product-doctrine on the 2-quasicategory of all categories (the object part of which is the free product completion as outlined in Example 6.9), this would lead to an alternative (weaker) definition of products in which the diagonal functors $\Delta : \mathbf{A} \to [\mathbf{I}, \mathbf{A}]$ ($\mathbf{I}$ discrete) are merely required to be weakly co-adjoint. This means explicitly that products are defined as in Example 6.9 except that the projections are no longer required to be right composable. Rather than pursuing this rather involved point here, we present two examples of such 'weak products'.

In the 'strong' precategory $\mathbf{Set_{nc}}$ of sets and non-constant maps (*cf.* Example 2.8), the usual Cartesian products of non-empty families $(Y_i)_I$ are proper 'weak products'

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

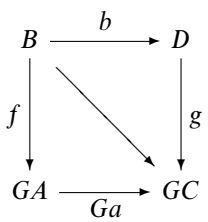
in the sense just indicated: it is easily verified that the relevant properties are satisfied (in particular, if $(f_i : X \to Y_i)_I$ is a family of non-constant maps, the induced map $f : X \to \prod_I Y_i$ is non-constant), and the projections in general fail to be right composable (that is, composites $\pi_i f : X \to Y_i$, where $f : X \to \prod_I Y_i$ is non-constant, may be constant).

Second, in the precategory **fRel(1)** of unary relations and full relation morphisms (*cf.* Example 2.4), one has 'weak coproducts': given a family $((X_i, A_i))_I$ of unary relations, the 'weak coproduct' is just the usual coproduct of relations, that is, the family of injections $\mu_i : (X_i, A_i) \hookrightarrow (\coprod_I X_i, \coprod_I A_i)$, where $\coprod$ denotes the disjoint union. It is easily seen that the μ_i are right composable and that, given a family of full relation morphisms $f_i : (X_i, A_i) \to (Y, B)$, the extension $(\coprod_I X_i, \coprod_I A_i) \to (Y, B)$ is full. However, the μ_i in general fail to be left composable: let $A = (\{0\}, \{0\})$ and $B = (\{0\}, \varnothing)$, let μ_A and μ_B denote the injections into the coproduct $A \amalg B$, and let f be the unique map $A \amalg B \to A$. Then f is a full relation morphism, but $f\mu_B$ is not.

Having said all this, it is only natural to generalise further towards a concrete isomorphism between *subprecategories* of the comma precategories $(\mathbf{B}, G)$ and $(F, \mathbf{A})$ on *both* sides. The result will be an adjunction where the universal property of the unit holds only for a restricted class of structured arrows. Indeed, an example of this nature is encountered in the *category* (sic) of probability spaces and measure preserving maps, where a pair (f_1, f_2) of morphisms with common domain factors (uniquely) through the independent product of the codomains iff f_1 and f_2 are independent (the definitions are recalled in Section 9).

In such 'partial' adjunctions, co-units need not exist. For example, it is easily seen that in the example just given, the 'would-be co-unit', that is, the diagonal map into the independent product of a space with itself, in general fails to be measure-preserving. Thus, the envisaged notion of partial adjointness will not be symmetric; on the level of comma categories, this will mean that the relevant subprecategory of $(\mathbf{B}, G)$ cannot be required to be covering. In order to perform arguments as in the proof of Theorem 6.4, it will, however, have to be semifull. (It is clear that an embedding of, say, a discrete subprecategory will not lead to sensible results, so certain restrictions of this nature are inevitable.) Moreover, to obtain a tractable notion, we will have to insist on a closure property, as follows.

Definition 8.7. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories, and let $\mathbf{K}$ be a subprecategory of $(\mathbf{B}, G)$. A $(\mathbf{B}, G)$-morphism (b, a) between $\mathbf{K}$-objects $f : B \to GA$ and $g : D \to GC$ is called $\mathbf{K}$-*firm* if the diagonal $gb = Gaf : B \to GC$ is a $\mathbf{K}$-object:



$\mathbf{K}$ is called *stable* if all $\mathbf{K}$-morphisms are $\mathbf{K}$-firm.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

If **K** is semifull, firmness of a **K**-morphism (b, a) is equivalent to the statement that (b, a) factors as $(b, a) = (b, id)(id, a)$ in **K**.

Thus we arrive at the following definition.

Definition 8.8. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories, and let **K** be a stable semifull subprecategory of $(\mathbf{B}, G)$. G is called **K**-*partially adjoint* if there exists a functor $F : \mathbf{B} \to \mathbf{A}$ and a semifull cocovering concrete embedding $\mathbf{K} \hookrightarrow (F, \mathbf{A})$. In this case, F is called a **K**-*partial left adjoint* of G.

The restriction that **K** be stable is not a serious one. For example, given any subprecategory **K** of $(\mathbf{B}, G)$, the subprecategory of $(\mathbf{B}, G)$ consisting of the **K**-objects and the **K**-firm morphisms is stable and semifull. Thus, arbitrary collections of structured morphisms can be captured by this definition.

As indicated above, there is, in general, no equivalent formulation of partial adjointness in terms of adjoint situations; however, one still has a description via (partial) universal arrows as in the motivation. This requires suitably modified versions of the various conditions on structured arrows, as follows.

Definition 8.9. Let **K** be a subprecategory of $(\mathbf{B}, G)$, where $G : \mathbf{A} \to \mathbf{B}$ is a functor between precategories. A G-structured morphism $f : B \to GA$ is called **K**-*partially universal* if, for each **K**-object $g : B \to GC$, there exists a unique morphism $g^{\#} : A \to C$ such that $(Gg^{\#})f = g$, and **K**-*strictly right composable* if $fg : C \to GA$ is defined and a **K**-object for each $g : C \to B$ in **B**. f **K**-*partially G-induces composability* if, whenever $g : A \to C$ and $h : C \to D$ are morphisms in **A** such that $(Gg)f : B \to GC$ and $(Gh)((Gg)f) : B \to GD$ are defined and in $\mathrm{Ob}\,\mathbf{K}$, we have hg (and, hence, $(G(hg))f$) is defined.

(Note that **K**-strictly right composable structured morphisms are **K**-objects.) This allows the following straightforward, if rather more modest, adaptation of Theorem 8.3.

Theorem 8.10. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories, and let **K** be a stable semifull subprecategory of $(\mathbf{B}, G)$. The following are equivalent:

(i) G is **K**-partially adjoint.
(ii) For each object B in **B**, there exists a **K**-strictly right composable **K**-partially G-universal arrow $\eta_B : B \to GA_B$ that **K**-partially G-induces composability.

Proof. $(i) \Rightarrow (ii)$: Let $F : \mathbf{B} \to \mathbf{A}$ be a functor, and let $\Phi : \mathbf{K} \hookrightarrow (F, \mathbf{A})$ be a semifull cocovering embedding. The construction of the $\eta_B : B \to GA_B = GFB$ and the verification of the fact that the η_B are right composable and G-induce composability are analogous to the proof of Theorem 6.4, '$(i)\Rightarrow(iii)$'. To show that the η_B are **K**-partially universal, let $f : B \to GA$ be a **K**-object, and let $f^{\#} : FB \to A$ be its image under Φ. Then for $\bar{f} : FB \to A$, we have $G\bar{f}\eta_B = f$ iff $(id_B, \bar{f})$ is a morphism from $\eta_B : B \to GFB$ to $f : B \to GA$ in **K** iff $(id_B, \bar{f})$ is a morphism from $id_{FB} : FB \to FB$ to $f^{\#} : FB \to A$ in $(F, \mathbf{A})$ iff $\bar{f} = f^{\#}$.

$(ii) \Rightarrow (i)$: The partial left adjoint and the required embedding are constructed analogously as in the proof of Theorem 6.4, '$(ii)\Rightarrow(i)$'; the crucial point is that, by stability of **K**, the

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

verification of the fact that a semifull concrete graph morphism has been defined can be restricted to the special case of **K**-morphisms of the form (b, id) or (id, a). □

Note that, by construction, the partial universal arrows η_B form part of a natural transformation $\eta : id_{\mathbf{B}} \to GF$, called the **K**-*partial unit*.

Remark 8.11. Despite the fact that examples for partial adjointness that are of the same nature as independent products (*cf.* Section 9 for details) in the sense that proper subprecategories of the comma precategory appear on *both* sides are as yet unknown outside probability theory, examples of partial adjointness do abound: given any weakly adjoint functor (recall that standard constructions lead to weakly adjoint functors!), its weak left adjoint is a partially co-adjoint functor. Hence, the notion of partial adjointness actually arises necessarily as a complement to weak adjointness, thus partially restoring the lost symmetry.

Finally, we do, of course, need a uniqueness result for partial left adjoints. The situation is largely the same as in the case of weak adjointness, since nothing much has changed on the side of costructured morphisms. The main point to observe is that, given a **K**-partially adjoint functor G with partial left adjoint F and partial universal arrows η_B, definedness of $Gf\eta_B$ no longer implies that $f : FB \to A$ belongs to the image of the relevant embedding $\mathbf{K} \hookrightarrow (F, \mathbf{A})$. Thus, the adapted version of the unicity theorem (Theorem 8.4) reads as follows.

Theorem 8.12. Let **A** and **B** be precategories, with **A** admissible, and let $G : \mathbf{A} \to \mathbf{B}$ be a **K**-partially adjoint functor (that is, **K** is a stable semifull subprecategory of $(\mathbf{B}, G)$) with **K**-partial left adjoint F and **K**-partial unit η. Then a functor $F' : \mathbf{B} \to \mathbf{A}$ is a **K**-partial left adjoint of G iff there exists a natural co-isomorphism (that is, an invertible natural transformation) $\mu : F \to F'$ such that $(G\mu_B)\eta_B : B \to GF'B$ is defined and a **K**-object for each object B in **B**.

The proof is analogous to that of Theorem 8.4. In order that we can omit the additional condition on μ, we have to require not only that **B** be co-admissible, but also that **K** be *isomorphism-closed* in the sense that, whenever $f : B \to GA$ is a **K**-object and $g : A \to C$ is invertible, we have $(Gg)f : B \to GC$ is a **K**-object (this is the case in 'reasonable' applications, including the ones presented here).

9. Independent products

In order to model the synchronous product of probabilistic processes (van Glabbeek *et al.* 1995), we require the notion of aggregation of probabilistic automata. In the classical case, the aggregation of two (non-probabilistic) automata is just their categorical product (Winskel and Nielsen 1995). In the probabilistic case, a prerequisite for such a product would be a 'product' of probability spaces. Now, even when morphisms between probability spaces are restricted to the classical measure preserving transformations (which form a category), there is no product (in the strict categorical sense) available, since only

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

a restricted class of pairs of random quantities (namely, the independent pairs) factors through the standard (independent) distribution on the Cartesian product.

Now recall from Example 6.9 that a feasible definition of a precategory $\mathbf{A}$ with binary products requires the diagonal functor $\Delta : \mathbf{A} \to \mathbf{A}^2$ to be co-adjoint. Thus, the appropriate (pre)categorical concept to be used in this situation is that of partial (co-)adjointness as introduced in the previous section; this will provide a universal characterisation of independent products of probability spaces as well as of the aggregation of probabilistic automata.

Hence, we will have to single out a suitable subprecategory of the precategory of 2-sources of probability presentation morphisms (that is, of $(\Delta, \mathbf{Prob}^2)$). To begin, we will work with independent sources, defined as in the classical case (that is, for random quantities).

Definition 9.1. A 2-*source*, that is, a pair (f_1, f_2) of morphisms in **Prob** with common domain (in other words, a Δ-costructured morphism, where Δ denotes the diagonal functor $\mathbf{Prob} \to \mathbf{Prob}^2$) is called *independent* if

$$P(f_1^{-1}[B_1] \cap f_2^{-1}[B_2]) = P(f_1^{-1}[B_1])P(f_2^{-1}[B_2])$$

for all subsets B_1 and B_2 of the respective codomains. The full concrete subprecategory of $(\Delta, \mathbf{Prob}^2)$ spanned by all independent 2-sources is denoted **Ind**.

Proposition 9.2. Let $\mathscr{S} = (f_1, f_2)$ be an independent 2-source, and let $g_i f_i$ be a morphism, $i = 1, 2$. Then $(g_1 f_1, g_2 f_2)$ is an independent 2-source.

Proof. Let B_i be a subset of the codomain of g_i, $i = 1, 2$. By independence of $\mathscr{S}$, applied to the $g_i^{-1}[B_i]$, we obtain the required condition:

$$P(f_1^{-1}g_1^{-1}[B_1] \cap f_2^{-1}g_2^{-1}[B_2]) = P(f_1^{-1}g_1^{-1}[B_1])P(f_2^{-1}g_2^{-1}[B_2]). \qquad \square$$

Corollary 9.3. **Ind** is a stable subprecategory of $(\Delta, \mathbf{Prob}^2)$.

With respect to **Ind**, we do indeed obtain the desired partial co-adjunction, which, as expected, involves independent products, as follows.

Definition 9.4. Let (Ω_1, p_1) and (Ω_2, p_2) be probability presentations. The *independent product* of (Ω_1, p_1) and (Ω_2, p_2) is the probability presentation $(\Omega_1 \times \Omega_2, p_{12})$, where $p_{12}(\omega_1, \omega_2) = p_1(\omega_1)p_2(\omega_2)$.

Lemma 9.5. Let $f_i : (\Omega, p) \to (\Omega_i, p_i)$, $i = 1, 2$, form a 2-source in **Prob**. Then the factorising map $f = \langle f_1, f_2 \rangle$ is a morphism $f : (\Omega, p) \to (\Omega_1 \times \Omega_2, p_{12})$ iff there exists a constant c such that, for each pair $(\omega_1, \omega_2) \in f[\Omega]$,

$$P(f_1^{-1}[\{\omega_1\}] \cap f_2^{-1}[\{\omega_2\}]) = cp_1(\omega_1)p_2(\omega_2). \qquad (*)$$

Proof. f is a morphism iff, for each pair $(\omega_1, \omega_2) \in f[\Omega]$, $P(f^{-1}[\{(\omega_1, \omega_2)\}]) = p_{12}(\omega_1, \omega_2)/P_{12}(f[\Omega])$; equivalently, $P(f_1^{-1}[\{\omega_1\}] \cap f_2^{-1}[\{\omega_2\}]) = p_1(\omega_1)p_2(\omega_2)/P_{12}(f[\Omega])$. The last condition implies Equation $(*)$ for $c = 1/P_{12}(f[\Omega])$. Conversely, if there exists c as in the statement, then taking the sum of both sides of Equation $(*)$ over $f[\Omega]$ yields $1 = cP_{12}(f[\Omega])$, as required. $\square$

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Theorem 9.6. Independent products are **Ind**-partial Δ-couniversal arrows; that is, Δ is **Ind**-partially co-adjoint.

Proof. By Corollary 9.3, **Ind** is stable. Let (Ω_1, p_1) and (Ω_2, p_2) be probability presentations, and let $\mathscr{P} = (\pi_1, \pi_2)$ denote the pair of projections associated to their independent product.
$\mathscr{P}$ *is* **Ind**-*partially co-universal:* Let (f_1, f_2) be an independent 2-source. Then (f_1, f_2) fulfills the condition of Lemma 9.5, taking $c = P(f_1[\Omega])P(f_2[\Omega])$, and hence factors through the independent product, as required.
$\mathscr{P}$ *is* **Ind**-*strictly left composable:* Let g_i be a morphism with domain (Ω_i, p_i), $i = 1, 2$. Then the composites $g_i\pi_i$ are defined by Example 6.3, and by Lemma 9.2, $(g_1\pi_1, g_2\pi_2)$ is an independent 2-source.
$\mathscr{P}$ **Ind**-*partially* Δ-*co-induces composability:* Let g and h be morphisms in **Prob** such that $(\pi_1 g, \pi_2 g)$ and $\mathscr{S} = ((\pi_1 g)h, (\pi_2 g)h)$ are independent 2-sources. Then gh is the factorisation of $\mathscr{S}$ on the level of maps, and hence is a morphism by the above. □

Remark 9.7. Theorem 9.6 is easily generalised to 2-sources that fulfill the equivalent conditions of Lemma 9.5 (such sources might be called *proportionally independent*). In this case the associated subprecategory of $(\Delta, \mathbf{Prob}^2)$ is not full, since not all morphisms between proportionally independent sources are firm.

Remark 9.8. The morphisms into the independent product of two spaces (Ω_i, p_i), $i = 1, 2$, that arise as factorisations of independent sources are precisely the *probabilistically rectangular maps*, that is, maps $f : (\Omega, p) \to (\Omega_1 \times \Omega_2, p_{12})$ such that $P_{12}(f[\Omega]) = P_1(\pi_1 f[\Omega_1])P_2(\pi_2 f[\Omega_2])$ (where the π_i denote the projections).

Similarly, the factorisations of proportionally independent sources (that is, the morphisms into the independent product that compose with the projections) are precisely the *slice-constant* morphisms. Here, a morphism $f : (\Omega, p) \to (\Omega_1 \times \Omega_2, p_{12})$ is called slice-constant if the function that maps ω_1 to $P_2(\{\omega_2 \mid (\omega_1, \omega_2) \in f[\Omega]\})$ is constant on $\pi_1 f[\Omega]$ *and* the analogous condition holds for the second factor.

10. Combining probabilistic automata

We now turn our attention to three basic combination mechanisms for probabilistic automata: action restriction, aggregation and interconnection via action calling. These concepts extend the corresponding definitions for classical (non-probabilistic) automata or transition systems, as explored for example in Winskel and Nielsen (1995). The use of such constructions within probabilistic systems, without the universal characterisations, is already present in van Glabbeek *et al.* (1995).

Restriction

In the spirit of Winskel and Nielsen (1995), restrictions of probabilistic automata will require the formation of Cartesian lifts; for the time being, we stick to a definition carried over verbatim from category theory.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Definition 10.1. Let $F : \mathbf{A} \to \mathbf{B}$ be a functor between precategories. A morphism $f : A \to C$ in $\mathbf{A}$ is called *Cartesian* (or *initial*) with respect to F if, whenever $Ffh = Fg$ is defined for morphisms $g : D \to C$ and $h : FD \to FA$, there exists a unique morphism $\bar{h} : D \to A$ such that $F\bar{h} = h$ and $f\bar{h} = g$. A *Cartesian lift* of an F-structured morphism $f : B \to FA$ is a Cartesian morphism $\bar{f}$ (or its domain) with codomain A such that $F\bar{f} = f$.

It is easy to see that Cartesian lifts are unique up to identity-carried isomorphisms.

A prerequisite for the restriction of automata, then, is the formation of Cartesian lifts with respect to the forgetful functor $Out : \mathbf{Prob} \to \mathbf{Set}$.

Proposition 10.2. Let (Ω, p) be a probability presentation, and let $\Omega_0 \subset \Omega$ such that $P(\Omega_0) > 0$. Then (Ω_0, p_0), where $p_0(\omega) = p(\omega)/P(\Omega_0)$, is a Cartesian lift of $\Omega_0 \hookrightarrow Out(\Omega, p)$ with respect to Out.

Proof. $m : (\Omega_0, p_0) \hookrightarrow (\Omega, p)$ is a morphism by construction. Now let (Ω_1, p_1) be a probability presentation, and let $h : \Omega_1 \to \Omega_0$ be a map such that $mh : (\Omega_1, p_1) \to (\Omega, p)$ is a morphism; we have to verify that $h : (\Omega_1, p_1) \to (\Omega_0, p_0)$ is a morphism. Now for $\omega \in h[\Omega_1] \subset \Omega_0$, we have (recall Lemma 3.3)

$$\begin{aligned} P_1(h^{-1}[\{\omega\}]) &= P_1((mh)^{-1}[\{\omega\}]) \\ &= \frac{p(\omega)}{P(mh[\Omega])} \\ &= \frac{p_0(\omega)P(\Omega_0)}{P_0(h[\Omega_1])P(\Omega_0)} = \frac{p_0(\omega)}{P_0(h[\Omega_1])}. \end{aligned}$$

□

Since we will be dealing with restriction of *actions*, the relevant Cartesian lifts in **PA** will be formed with respect to the functor $Alph : \mathbf{PA} \to \mathbf{pSet}$ to the category of pointed sets that maps a probabilistic automaton $M = (A, S, \delta)$ to its alphabet A.

Proposition 10.3. Let $M = (A, S, \delta)$ be a probabilistic automaton, and let A_0 be a subset of A containing $\perp$. Then (A_0, S, δ_0), where $T_0^s = T^s \cap (A_0 \times S)$ and δ_0^s is the Cartesian lift of $T_0^s \hookrightarrow Out(\delta^s)$ with respect to Out, is a Cartesian lift of $A_0 \hookrightarrow Alph(M)$ with respect to $Alph$.

Proof. The proof is straightforward. □

Definition 10.4. Let $M = (A, S, \delta)$ be a probabilistic automaton, and let A_0 be a subset of A containing $\perp$. Then the *restriction* of M to A_0, denoted $M|_{A_0}$, is defined as the Cartesian lift of $A_0 \hookrightarrow Alph(M)$.

As we shall see, restriction is especially important in combination with aggregation.

Remark 10.5. Let $F : \mathbf{A} \to \mathbf{B}$ be a functor between precategories, and let $\mathbf{C}$ be a full coreflective subprecategory of $\mathbf{A}$ with a choice of coreflective arrows $c_A : C_A \to A$ such that Fc_A is an identity for each object A in $\mathbf{A}$. Then it is easy to see that, given a Cartesian morphism (with respect to F) $f : A \to C$ in $\mathbf{A}$, where C is an object in $\mathbf{C}$, we have $fc_A : C_A \to C$ is Cartesian with respect to the restriction $F_{\mathbf{C}} : \mathbf{C} \to \mathbf{B}$ of F. (Note, in

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

particular, that fc_A is indeed defined by virtue of the left composability requirement on co-universal arrows.)

Now the coreflections established in Section 7 do have identity-carried coreflective arrows with respect to *Alph*. Thus, Cartesian lifts with respect to the restrictions of *Alph* to **AcPA** and **DT**, respectively, are given as the coreflections of the corresponding lifts with respect to $Alph : \mathbf{PA} \to \mathbf{pSet}$ as constructed above. These Cartesian lifts then represent the correct notion of restriction for accessible automata and decision trees, respectively.

Aggregation and interconnection

It is now natural to define aggregation of two automata in terms of independent products of probability spaces, obtaining, in turn, a partial product of automata.

Definition 10.6. A 2-source (f_1, f_2) with domain (A, S, δ) in **PA** is called *independent* if the restriction of $(\overline{f_1} \times \underline{f_1}, \overline{f_2} \times \underline{f_2})$ to δ^s is independent in **Prob** for each $s \in S$. The full concrete subprecategory of $(\Delta, \mathbf{PA}^2)$ spanned by all independent 2-sources is denoted **IndA**.

Proposition 10.7. Δ is **IndA**-partially coadjoint; the **IndA**-partial Δ-couniversal arrow for a pair of automata $M_i = (A_i, S_i, \delta_i)$, $i = 1, 2$, is their *aggregation*, given by $(A_1 \times A_2, S_1 \times S_2, \delta)$, where $T^{s_1 s_2} = T_1^{s_1} \times T_2^{s_2}$ and $\delta^{s_1 s_2}$ is the independent product of $\delta_1^{s_1}$ and $\delta_2^{s_2}$ in **Prob**, together with the natural projections.

Proof. The proof is straightforward, using Theorem 9.6. □

The aggregation of M_1 and M_2 is denoted $M_1 \| M_2$. Intuitively, the aggregation corresponds to the two automata working synchronously, but without any interaction.

If we want to impose some interaction, for instance, a *calling interaction* where an action a_1 of M_1 is only allowed in connection with an action a_2 of M_2, we have to restrict the aggregation. In the restricted (product) alphabet, all pairs involving a_1 except (a_1, a_2) are dropped. Thus, we reach the following notion of interconnection by action calling.

Definition 10.8. Let $M_i = (A_i, S_i, \delta_i)$, $i = 1, 2$, be probabilistic automata, and let a_1 and a_2 be actions in M_1 and M_2, respectively. Then the *interconnection of* M_1 *and* M_2 *by* a_1 *calling* a_2 is

$$M_1 \|_{a_1 >> a_2} M_2 := (M_1 \| M_2)|_{A_0},$$

where

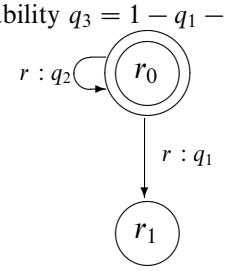
$$A_0 = ((A_1 \setminus \{a_1\}) \times A_2) \cup \{(a_1, a_2)\}.$$

In the case of non-probabilistic automata, this type of manipulation has been used as a fundamental operation (Costa *et al.* 1993), replacing symmetric action sharing as implemented, for example, in CSP (Hoare 1978). (Of course, symmetric sharing can be modelled by restricting the aggregation as well.)

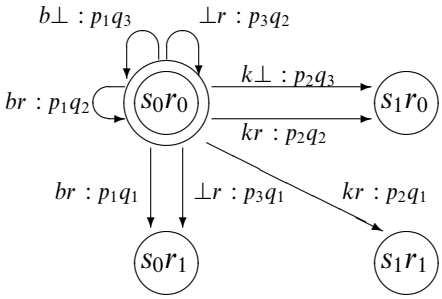
Remark 10.9. It is easily verified that **AcPA** and **DT** are closed under aggregation in **PA**. Recall, however, that Cartesian lifts with respect to *Alph* in these precategories are formed as coreflections of Cartesian lifts in **PA** (*cf.* Remark 10.5), which leads to correspondingly modified definitions of interconnection.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

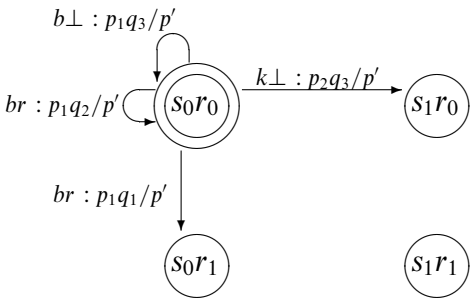
Example 10.10. Consider now a *receiver* that tries to *read* (r) what is broadcasted by some other automaton. At each unit of time, the receiver reads successfully with probability q_1 and unsuccessfully with probability q_2; if successful, the automaton ceases to attempt further readings. With probability $q_3 = 1 - q_1 - q_2 > 0$, the automaton does nothing.



If we aggregate the broadcaster and the receiver without interaction, we obtain the following automaton (only the probability transitions for the initial state are shown):



Now, in order to 'tune' the receiver to the broadcaster, we impose a calling interaction $r >> b$; that is, the receiver can only read if the broadcaster is broadcasting:



Here, $p' = 1 - p_2q_1 - p_2q_2 - p_3q_1 - p_3q_2$. (Again, the figure shows only the probability transitions for the initial state.)

It should once more be emphasised that interaction of automata depends heavily on a notion of morphism that allows for surjective maps (in this case, product projections) as well as for subobject embeddings (in this case, the restriction of the product). Moreover, the universal characterisation of interaction requires some variant of categorical products. Thus, the necessity of the precategorical approach chosen here is prominently exemplified by this construction.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

11. Conclusion

We have motivated and developed a theory of precategories that generalises parts of category theory, including, in particular, the theory of adjoint functors. Moreover, we have demonstrated how the standard operations of restriction and relabelling of probabilistic automata necessitate notions of morphisms between probability spaces and, subsequently, between probabilistic automata that fail to be stable under composition and hence require precategory theory for their treatment. Thanks to the above-mentioned (pre-existing) theory of adjointness for precategories, well-known 'improvements' of probabilistic automata such as the accessible modification or the formation of the associated decision tree are expressible in terms of (co-)universal constructions.

A more careful analysis then led to a universal characterisation of the aggregation and interconnection of probabilistic automata. For reasons that are inherent in the nature of probability spaces, even when equipped with the classical measure preserving transformations as morphisms, this required a generalisation of the theory of adjoint functors to partial adjointness, which turned out to fit naturally into the framework presented so far.

As laid out in Goguen (1991), this approach has various benefits:

— The theoretical basis for the mentioned operations on automata is broadened, resulting in an increase of trust in the 'correctness' of their definitions.
— Certain operations come for free thanks to the abstract theory.
— A guideline is provided for the extension of these and possibly other concepts to different settings, including existing or forthcoming alternative notions of probabilistic automata, possibly based on less restrictive types of probability spaces.

Acknowledgements

This work was partially supported by *Fundação para a Ciência e a Tecnologia*, the PRAXIS XXI Projects PRAXIS/P/MAT/10002/1998 ProbLog, PCEX/P/MAT/46/96 ACL and 2/2.1/TIT/1658/95 LogComp, as well as by the ESPRIT IV Working Groups 22704 ASPIRE and 23531 FIREworks. The diagrams were typeset using Paul Taylor's diagram package.

References

Adámek, J., Herrlich, H. and Strecker, G. E. (1990) *Abstract and concrete categories*, Wiley Interscience, New York.

Arbib, M. and Manes, E. (1974) Machines in a category: An expository introduction. *SIAM Review* **16** (2) 163–192.

Condon, A., Hellerstein, L., Pottle, S. and Wigderson, A. (1998) On the power of finite automata with both non-deterministic and probabilistic states. *SIAM Journal of Computation* **27** (3) 739–762.

Coppey, L. (1976) Décompositions directes des précategories. *Math. Arbeitspapiere der Universität Bremen* **7** 26–41.

Coppey, L. (1980) Quelques problèmes typiques concernant les graphes multiplicatifs. *Diagrammes* **3** C1–C46.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press

Costa, J., Sernadas, A. and Sernadas, C. (1993) Data encapsulation and modularity: Three views of inheritance. In: Borzyszkowski, A. and Sokolowski, S. (eds.) Mathematical Foundations of Computer Science '93. *Springer-Verlag Lecture Notes in Computer Science* **711** 382–391.

Ehresmann, C. (1965) *Catégories et structures*, Dunod, Paris.

Giacalone, A., Jou, C.-C. and Smolka, S. A. (1990) Algebraic reasoning for probabilistic concurrent systems. Proceedings of the 19th ICALP. *Springer-Verlag Lecture Notes in Computer Science* **623**.

Van Glabbeek, R., Smolka, S. A., Steffen, B. and Tofts, C. (1995) Reactive, generative and stratified models for probabilistic processes. *Information and Computation* **121** (1) 59–80.

Goguen, J. (1991) A categorical manifesto. *Mathematical Structures in Computer Science* **1** (1) 49–67.

Graetzer, G. (1979) *Universal Algebra*, second edition, Springer-Verlag.

Halmos, P. (1950) *Measure Theory*, Van Nostrand, New York.

Herrlich, H. and Strecker, G. E. (1979) *Category theory*, second edition, Heldermann Verlag, Berlin.

Herrlich, H., Nakagawa, R., Strecker, G. E. and Titcomb, T. (1980) Equivalence of topologically algebraic and semi-topological functors. *Canad. J. Math.* **32** 34–39.

Hoare, C. (1978) Communicating Sequential Processes. *Comm. ACM* **21** 666–677.

Kock, A (1995) Monads for which structures are adjoint to units. *J. Pure Appl. Algebra* **104** 41–59.

Knauer, U. and Böttcher, M. (1992) Endomorphism spectra of graphs. *Discr. Math.* **109** 45–57.

Mac Lane, S. (1972) *Categories for the Working Mathematician*, Springer-Verlag.

Mateus, P., Sernadas, A. and Sernadas, C. (1999) Precategories for combining probabilistic automata. Proceedings of CTCS 99. *Electron. Notes Theor. Comput. Sci.* **29**.

Mateus, P. (2000) *Interconnection of Probabilistic Systems*, Ph. D. thesis, IST, Universidade Técnica de Lisboa (Supervised by A. Sernadas and C. Sernadas).

Pinto, J., Sernadas, A., Sernadas, C. and Mateus, P. (2000) Non-determinism and uncertainty in the situation calculus. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems* **8** (2).

Rabin, M. (1963) Probabilistic automata. *Inform. and Control* **6** (3) 220–245.

Schröder, L. (1999) *Composition Graphs*, Ph. D. thesis (in German), University of Bremen. Also: Logos-Verlag, Berlin.

Schröder, L. (2001) Monads on composition graphs. *Appl. Cat. Struct.* (to appear).

Schröder, L. and Herrlich, H. (2000a) Abstract initiality. *Comm. Math. Univ. Carol.* **41** 575–583.

Schröder, L. and Herrlich, H. (2000b) Free adjunction of morphisms. *Appl. Cat Struct.* **8** 595–606.

Winskel, G. and Nielsen, M. (1995) Models of concurrency. In: Gabbay, D., Abramsky, S. and Maibaum, T. (eds.) *Handbook of Logic in Computer Science* **4**, Oxford Science Publications 1–148.

Wu S.-H., Smolka, S. and Stark, E. (1997) Composition and behaviors of probabilistic I/O automata. *Theor. Comput. Sci.* **176** (1) 1–38.

https://doi.org/10.1017/S0960129502003614 Published online by Cambridge University Press