Universal Aspects of Probabilistic Automata

LUTZ SCHRÖDER[1] and PAULO MATEUS[2]

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

Received 16 July 1999

For lack of composability of their morphisms, probability spaces, and hence probabilistic automata, fail to form categories; however, they fit into the more general framework of precategories, which are introduced and studied here. In particular, the notion of adjunction and weak adjunction for precategories is presented and justified in detail. As an immediate benefit, a concept of (weak) product for precategories is obtained. Thus, universal properties can be used for characterizing well-known basic constructions in the theory of probabilistic automata: The aggregation of two automata is shown to be a weak product, whereas restriction and interconnection of automata are recognized as Cartesian lifts. Finally, we establish that the precategory of decision trees is coreflexive in the precategory of probabilistic automata.

1. Introduction

Until recently, structures that are weaker than categories have been largely ignored in the literature, with the exception of Ehresmann's work on multiplicative graphs (see e.g. Ehresmann (1965)), which was continued e.g. in Coppey (1976) and Coppey (1980). In Schröder and Herrlich (1999) and Schröder (1999), the notions of composition graph and semicategory were used to construct extensions of categories. Parallel to this, precategories where introduced in Mateus et al. (1999) to provide universal characterizations of constructions in the theory of probabilistic automata; precategories can be regarded as composition graphs that satisfy additional axioms, concerning identities and the degree of associativity, which typically hold when the objects in question are structured sets.

In Schröder (1999), a notion of adjointness for functors between multiplicative graphs was introduced and justified by suitable characterization theorems. This theory is extended here to include weak adjunctions, i.e. cases where the universal property holds only for arrows that satisfy additional conditions; this level of generality is needed to deal with independent products in the precategories of probability spaces respectively probabilistic automata. Thus, the universal characterizations of aggregation, restriction and interconnection of probabilistic automata given in Mateus et al. (1999), and a characterization of several constructions that relate different probabilistic models such as probabilistic automata and decision trees, are put on a sound theoretical basis.

To begin, we introduce composition graphs and precategories. Using the couniversal property of the evaluation functor, we obtain a generalization of the notion of functor category, and hence of the notion of natural transformation, to the setting of precategories. This leads directly to a Lawvere-type comma construction for precategories, which in turn provides a satisfactory generalization of the notion of adjointness: Calling functors F and G adjoint iff the comma precategory of G-structured arrow is concretely isomorphic to the comma precategory of F-costructured arrows, one has characterizations of adjointness by universal arrows as well as by adjoint situations. Weak adjunctions are then defined by requiring concrete isomorphy of suitable subgraphs of the said comma precategories.

Having thus established the 'precategory-theoretical' background, we introduce the precategory of probability spaces; the independent product of two probability spaces provides a first example of a weak adjunction. An analogous result holds in the precategory **PA** of generative probabilistic automata, the definition of which of course depends on that of probability spaces; the corresponding independent product is the desired aggregation of two automata. Furthermore, we recognize restriction and interconnection of automata as Cartesian lifts. Finally, we show that accessible probabilistic automata and decision trees, respectively, form coreflexive subprecategories of **PA**.

2. Composition Graphs and Precategories

We recall the basic notions of the theory of composition graphs as developed in Schröder (1999); categorical terminology generally follows . In this context we regard a *graph* G as a (large) algebra with two unary operations d and c, called *domain* and *codomain*, such that $c^2 = dc = c$ and $d^2 = cd = d$. A *graph morphism* is a homomorphism between two such algebras. The nodes of G (i.e. the elements of $c[G] = d[G]$) are called *objects* or *identities* and are denoted by capital letters A, B etc., or in the form id_A, id_B etc.; the elements of G are called *morphisms*.

Definition 2.1. A *composition graph* is a graph $\mathbf{A}$ with a partial binary operation, called *composition* and denoted by $(f, g) \mapsto fg$, such that $d(fg) = dg$, $c(fg) = cf$, and $df = cg$, whenever fg is defined. $\mathbf{A}$ is called

(i) *strongly identitive* if, for each f in $\mathbf{A}$, fdf and cff are defined and equal to f;
(ii) *well-associative* if, whenever fg and gh are defined in $\mathbf{A}$, then the equation $f(gh) = (fg)h$ holds strongly (i.e. whenever one side is defined, then so is the other, and both sides are equal);
(iii) a *precategory* if $\mathbf{A}$ is both strongly identitive and well-associative.

A *functor* between composition graphs is a graph morphism F such that $FfFg$ is defined and equal to $F(fg)$ whenever fg is defined. Subobjects (in the most general sense) in the quasicategory of composition graphs and functors will be called *subgraphs*.

Obviously, every category is a precategory. Strongly identitive composition graphs were introduced by C. Ehresmann (cf. Ehresmann (1965)), who called them multiplicative graphs or neocategories. Applications of composition graphs, in particular the more restrictive *semicategories*, and some more details can be found in Schröder and Herrlich (1999).

We call a functor *initial* if it is an initial (or Cartesian) morphism in the concrete quasicategory of composition graphs and functors over the quasicategory of graphs. A functor F is initial in this sense iff for each triple (f, g, h) of morphisms such that $cf = ch$, $dg = dh$, and $cg = df$, fg is defined and equal to h if $FfFg$ is defined and equal to Fh ('F reflects commutative triangles'). Initial functors between strongly identitive composition graphs are automatically faithful. Note that, if $F : \mathbf{A} \to \mathbf{B}$ is an initial functor and $\mathbf{B}$ is a category, then $\mathbf{A}$ is a precategory.

A composition graph with an initial forgetful functor is called *concrete*; a *concrete subgraph* of a concrete composition graph is an initial subgraph with the forgetful functor inherited by restriction. We will exclusively be concerned here with composition graphs that are concrete over **Set**; hence we shall restrict the exposition to precategories. However, almost all results can be adapted to the setting of merely strongly identitive composition graphs.

3. Functor Graphs

It is well known that the category **Cat** of small categories and functors is Cartesian closed (cf. Adámek et al. (1990)); the associated power objects are the functor categories. Similarly, the category **siCgr** of small strongly identitive composition graphs and functors is Cartesian closed, and the power objects in this category do indeed generalize functor categories. A description of these power objects, which we call *functor graphs*, can already be found in Coppey (1976). Surprisingly, strong identitivity plays a central role in this context; in fact, strong identitivity of the functor graph translates directly into functoriality of the objects and naturality of the morphisms. We proceed directly to the definition of functor graphs the realm of precategories, where things remain reasonably simple:

Definition 3.1. Let $\mathbf{A}$ and $\mathbf{B}$ be precategories. A *transformation* from F to G, where $F, G : \mathbf{A} \to \mathbf{B}$ are functors, is a family $\nu = (\nu_f)$, indexed over all morphisms f in $\mathbf{A}$, such that $\nu_f : Fdf \to Gcf$ is a morphism in $\mathbf{B}$ for each f. ν is called *natural* if for each f in $\mathbf{A}$, $Gf\nu_{df}$ and $\nu_{cf}Ff$ are defined in $\mathbf{B}$ and equal to ν_f. The *functor graph* $[\mathbf{A}, \mathbf{B}]$ is defined as the composition graph with all functors from $\mathbf{A}$ to $\mathbf{B}$ as objects and all 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 $\mu_f \nu_{df}$ (equivalently: $\mu_{cf}\nu_f$) is defined in $\mathbf{B}$ for each f in $\mathbf{A}$; in this case, $(\mu\nu)_f = \mu_f\nu_{df} = \mu_{cf}\nu_f$.

Thus defined, $[\mathbf{A}, \mathbf{B}]$ is in fact a precategory. We omit the verification that the evaluation map $ev : [\mathbf{A}, \mathbf{B}] \times \mathbf{A} \to \mathbf{B}$ which 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 graph now allows us to generalize Lawvere's comma construction to precategories. Let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ be precategories, and let $F : \mathbf{A} \to \mathbf{C}$ and $G : \mathbf{B} \to \mathbf{C}$ be functors; then the *comma graphs* $(F, \mathbf{C})$, $(\mathbf{C}, G)$, and (F, G) are defined by the following diagram of pullbacks:

$$\begin{tikzcd}
(F,G) \arrow[r] \arrow[d] & (F,\mathbf{C}) \arrow[r] \arrow[d] & \mathbf{A} \arrow[d, "F"] \\
(\mathbf{C},G) \arrow[r] \arrow[d] & {[\mathbf{2},\mathbf{C}]} \arrow[r, "P_1"'] \arrow[d, "P_2"] & \mathbf{C} \\
\mathbf{B} \arrow[r, "G"'] & \mathbf{C} &
\end{tikzcd},$$

where $\mathbf{2}$ denotes the category $1 \to 2$; the arrows $P_1, P_2 : [\mathbf{2}, \mathbf{C}] \to \mathbf{C}$ in the diagram are the two obvious projections. Whether the limit is taken in the quasicategory of composition graphs or that of precategories is irrelevant, since the latter is closed under limits in the former. $[\mathbf{2}, \mathbf{C}]$ is the composition graph analogue of the arrow category.

We give a more explicit description of (F, G):

The objects of (F, G) are (F, G)*-structured morphisms* (A, f, B), where A and B are objects in $\mathbf{A}$ respectively $\mathbf{B}$ and $f : FA \to GB$ is a morphism in $\mathbf{C}$. A morphism $\nu : (A_1, f_1, B_1) \to (A_2, f_2, B_2)$ can be regarded as a pair (a_1, b_1), where $a_1 : A_1 \to A_2$ and $b_1 : B_1 \to B_2$ are morphisms in $\mathbf{A}$ respectively $\mathbf{B}$, such that $f_2 F a_1$ and $G b_1 f_1$ are defined and equal:

$$\begin{tikzcd}
FA_1 \arrow[r, "Fa_1"] \arrow[d, "f_1"'] & FA_2 \arrow[d, "f_2"] \\
GB_1 \arrow[r, "Gb_1"'] & GB_2
\end{tikzcd}.$$

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

If $\mu = (a_2, b_2) : (A_2, f_2, B_2) \to (A_3, f_3, B_3)$ is a further morphism, then $\mu\nu$ is defined in (F, G) iff $(a_2, b_2)(a_1, b_1)$ is defined in $\mathbf{A} \times \mathbf{B}$ and $(a_2 a_1, b_2 b_1) : (A_1, f_1, B_1) \to (A_3, f_3, B_3)$ is a morphism in (F, G) (for the latter, it is sufficient that one of the terms $f_3 F(a_2 a_1)$ and $G(b_2 b_1) f_1$ be defined); i.e. the underlying functor of (F, G) is initial. In particular, any *concrete* graph morphism between two comma categories over $\mathbf{A} \times \mathbf{B}$, i.e. any graph morphism that commutes with the underlying functors, is already an initial functor.

Composition graphs of the form $(F, \mathbf{C})$ and $(\mathbf{C}, G)$ can be treated as special cases (namely, as $(F, id\ \ )$ respectively $(id\ \ , G)$); the notation for the objects can be simplified here: Objects of $(\mathbf{C}, G)$ are G*-structured morphisms* (f, B), where $f : C \to GB$, and objects of $(F, \mathbf{C})$ are F*-costructured morphisms* (A, f), where $f : FA \to C$.

4. Adjunctions

The generalization of the concept of an adjunction to the setting of composition graphs is not quite as unambiguous as the situation with functor graphs. All of the classical definitions pose severe difficulties in this context:

The definitions via universal arrows respectively adjoint situations (cf. Adámek et al. (1990)) are not easily translated, since it is not clear in what sense the involved equations should be required to hold (weakly, strongly, etc.). Moreover, it would be desirable to have an equivalence between these two definitions; this apparently cannot be obtained with any of the naive generalizations, 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 on both sides; hence it is not suitable for a setting with little or no associativity requirements.

An alternative definition which 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 composition 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 $(\mathbf{B}, G)$ (cf. Herrlich and Strecker (1979), 27O.). 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'.

In this context, we need a weakened concept of adjointness which can handle situations where the universal property of the units holds only with certain restrictions, i.e. where only structured morphisms with additional properties factor through the units. The approach via comma graphs turns out to be readily adaptable to this problem: Instead of an isomorphism between the full comma graphs $(\mathbf{B}, G)$ and $(\mathbf{A}, F)$, one has an isomorphism between suitable subgraphs.

To formulate the adjointness theorem, we need a certain amount of terminology:

Definition 4.1. Let $F : \mathbf{A} \to \mathbf{C}$ and $G : \mathbf{B} \to \mathbf{C}$ be functors between precategories, and let $\mathcal{K}$ be a subclass of $\mathrm{Ob}\,(F, G)$. An object $S = (A, f, B) \in \mathcal{K}$ is called

(i) *($\mathcal{K}$-strictly) G-wellcomposable* if (A, Ggf, B_1) is defined (and in $\mathcal{K}$) for each morphism $g : B \to B_1$.

(ii) *strongly G-wellcomposable* if S is G-wellcomposable, and whenever $g : B \to B_1$ and $h : B_1 \to B_2$ are morphisms in $\mathbf{B}$ such that $(A, Gh(Ggf), B_2)$ is defined, then hg (and hence $G(hg)f$) is defined.

(iii) *strongly $\mathcal{K}$-partially G-wellcomposable* if, whenever $g : B \to B_1$ and $h : B_1 \to B_2$

are morphisms in $\mathbf{B}$ such that (A, Ggf, B_1) and $(A, Gh(Ggf), B_2)$ are defined and in $\mathcal{K}$, then hg (and hence $G(hg)f$) is defined.

$\mathcal{K}$ is called *stable* if, whenever $(a,b) : (A_1, f_1, B_1) \to (A_2, f_2, B_2)$ is a morphism in (F,G) and $(A_i, f_i, B_i) \in \mathcal{K}$, $i = 1,2$, then $(A_1, Gbf_1, B_2) = (A_1, f_2Fa, B_2) \in \mathcal{K}$:

$$\begin{tikzcd} FA_1 \arrow[r, "Fa"] \arrow[d, "f_1"'] \arrow[dr] & FA_2 \arrow[d, "f_2"] \\ GB_1 \arrow[r, "Gb"'] & GB_2 \end{tikzcd} .$$

All of these terms will also be used for subgraphs $\mathbf{K}$ of (F,G); e.g. $\mathbf{K}$ will be called stable if $\mathrm{Ob}\,\mathbf{K}$ is stable.

Note that the word '(F,G)-strictly', but not the word '(F,G)-partially', can be left out without altering the meaning.

Remark 4.2. $\mathcal{K}$-strictly wellcomposable morphisms belong to $\mathcal{K}$.

Definition 4.3. Let $F : \mathbf{A} \to \mathbf{C}$ and $G : \mathbf{B} \to \mathbf{C}$ be functors between precategories. A functor Φ with codomain (F,G) is called *semifull* if each morphism of the form (id, b) or (a, id) in $\hom(\Phi K, \Phi L)$ has a preimage under Φ in $\hom(K, L)$. A subgraph is called semifull if its inclusion functor is semifull.

Lemma 4.4. Let $F : \mathbf{A} \to \mathbf{C}$ and $G : \mathbf{B} \to \mathbf{C}$ be functors between precategories, and let $\beta = (id, b) : S \to T$ and $\alpha = (a, id) : T \to U$ be morphisms in (F,G). Then $\alpha\beta = (a,b) : S \to U$ is defined in (F,G).

Remark 4.5. It is in fact easy to see that, in the notation of the above definition, morphisms of the form (id, b) on the one hand and of the form (a, id) on the other hand 'almost' form a 'factorization structure' (a notion that has yet to be defined) on (F,G); i.e. the required factorizations and unique diagonals exist. Closedness under isomorphisms is, of course, missing; this problem is somewhat difficult to settle in general, because a canonical definition of an isomorphism in a composition graph is not yet available. In practice, however, it is often clear what should be regarded as an isomorphism, and in these cases, the closure of the said construction under composition with such isomorphisms can surely be regarded as a factorization structure.

The above-defined wellcomposability conditions can be restated in terms of the existence of certain objects and morphisms in (F,G): The condition that, say, Ggf as in the above definition is defined and in $\mathcal{K}$ is equivalent to the existence of a morphism $(g, id) : (A, f, B) \to T$ in (F,G) with $T \in \mathcal{K}$, whereby T is, of course, uniquely determined; the remaining conditions are easily translated along the same lines. In particular, stability of $\mathcal{K}$ just means that any morphism (a,b) between $\mathcal{K}$-objects factors in (F,G) as $(a, id)(id, b)$ via a $\mathcal{K}$-object. These observations make the following lemma more or less trivial:

Lemma 4.6. Let $F_i : \mathbf{A} \to \mathbf{C}_i$ and $G_i : \mathbf{B} \to \mathbf{C}_i$ be functors between precategories, $i = 1, 2$, let $\mathbf{K}$ be a semifull concrete subgraph of (F_1, G_1); let $\Phi : \mathbf{K} \to (F_2, G_2)$ be a concrete functor, let $S = (A, f, B) \in \mathrm{Ob}\,(F_1, G_1)$, and let $\Phi S = (A, \tilde{f}, B)$.

(i) If $(A, G_1 gf, B_1)$ is defined and in $\mathbf{K}$ for some $g : B \to B_1$ in $\mathbf{B}$, then

$$\Phi(A, G_1 gf, B_1) = (A, G_2 g\tilde{f}, B_1).$$

(ii) If Φ is semifull, $\mathbf{K}$ is full, and $\Phi[\mathbf{K}]$ is stable, then Φ is full and $\mathbf{K}$ is stable.
(iii) If Φ is full and $\mathbf{K}$ is stable, then $\Phi[\mathbf{K}]$ is stable.
(iv) If S is $\mathbf{K}$-strictly G_1-wellcomposable, then ΦS is $\Phi[\mathbf{K}]$-strictly G_2-wellcomposable.
(v) If Φ is semifull and S is strongly $\mathbf{K}$-partially G_1-wellcomposable, then ΦS is strongly $\Phi[\mathbf{K}]$-partially G_2-wellcomposable.
(vi) If $\mathbf{K} = (F_1, G_1)$, Φ is an isomorphism, and S is strongly G_1-wellcomposable, then ΦS is strongly G_2-wellcomposable.

Proof. (i), (iv), (v): As an example, let Φ be semifull, let S be strongly $\mathbf{K}$-partially G_1-wellcomposable, and let $g : B \to B_1$ and $h : B_1 \to B_2$ be morphisms in $\mathbf{B}$ such that $T = (A, G_2 g\tilde{f}, B_1)$ and $U = (A, G_2 h(G_2 g\tilde{f}), B_2)$ are defined and in $\Phi[\mathbf{K}]$, say $T = \Phi\bar{T}$ and $U = \Phi\bar{U}$. Then we have morphisms

$$\begin{tikzcd} \Phi S \arrow[r, "(id,g)"] & \Phi\bar{T} \arrow[d, "(id,h)"'] \\ & \Phi\bar{U} \end{tikzcd}$$

in (F_2, G_2). We have to show that $(id, hg) : \Phi S \to \Phi\bar{U}$ is a morphism in (F_2, G_2). Now the diagram has a preimage in $\mathbf{K}$ by semifullness of Φ; thus $(id, hg) : S \to \bar{U}$ is a morphism in (F_1, G_1) by biwellcomposability of S, and hence in $\mathbf{K}$ by semifullness of $\mathbf{K}$. Application of Φ yields the desired morphism in (F_2, G_2). The remaining statements are shown analogously.

(ii): Let $(a, b) : \Phi S \to \Phi T$ be a morphism in (F_2, G_2). By stability of $\Phi[\mathbf{K}]$, we have a factorization

$$\begin{tikzcd} \Phi S \arrow[r, "(id,b)"] \arrow[dr, "(a,b)"'] & \Phi U \arrow[d, "(a,id)"] \\ & \Phi T \end{tikzcd}$$

in (F_2, G_2). Since Φ is semifull, $(id, b) : S \to U$ and $(a, id) : U \to T$ are morphisms in $\mathbf{K}$. By lemma 4.4, this implies that $(a, b) : S \to T$ is a morphism in (F_1, G_1) and hence in $\mathbf{K}$; thus Φ is full. To see that $\mathbf{K}$ is stable, just note that, whenever $(a, b) : S \to T$ is a morphism in (F_1, G_1), then $(a, b) : \Phi S \to \Phi T$ is a morphism in (F_2, G_2), which gives rise to $U \in \mathrm{Ob}\,\mathbf{K}$ as above.

(iii): Let $(a, b) : \Phi S \to \Phi T$ be a morphism in (F_2, G_2); then $(a, b) : S \to T$ is a

morphism in $\mathbf{K}$. Since $\mathbf{K}$ is stable and semifull, we have a factorization

$$\begin{tikzcd} S \arrow[r, "{(id,b)}"] \arrow[dr, "{(a,b)}"'] & U \arrow[d, "{(a,id)}"] \\ & T \end{tikzcd}$$

in $\mathbf{K}$; applying Φ yields the required factorization of (a,b) in (F_2, G_2).

(vi): Follows from (iv) and (v). □

Definition 4.7. Let $F : \mathbf{A} \to \mathbf{B}$ be a functor between precategories. A subclass of $\mathrm{Ob}\,(\mathbf{B}, F)$ (or a subgraph of $(\mathbf{B}, F)$) is called *F-covering* if it contains all structured morphisms of the form (Ff, A_2), where $f : A_1 \to A_2$; a functor Φ with codomain $(\mathbf{B}, F)$ is called *F-covering* if its image is F-covering.

Remark 4.8. If a subclass $\mathcal{K}$ of $\mathrm{Ob}\,(\mathbf{B}, F)$ as above is F-covering, then $\mathcal{K}$ is stable iff (Fgf, A_2) belongs to $\mathcal{K}$ whenever Fgf is defined for some (f, A_1) in $\mathcal{K}$ and some morphism $g : A_1 \to A_2$ in $\mathbf{A}$.

We are now prepared to state the adjointness theorem:

Theorem 4.9. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories, and let $\mathbf{K}$ be a full and stable concrete subgraph of $(\mathbf{B}, G)$. Equivalent are

(i) There exists a functor $F : \mathbf{B} \to \mathbf{A}$ and an F-cocovering semifull concrete embedding $\mathbf{K} \hookrightarrow (F, \mathbf{A})$.

(ii) For each object B in $\mathbf{B}$, there exists a $\mathbf{K}$-*partial G-universal arrow*, i.e. a $\mathbf{K}$-strictly $\mathbf{B}$-cowellcomposable and strongly $\mathbf{K}$-partially G-wellcomposable G-structured morphism $\eta_B : B \to GA_B$ in $\mathbf{K}$ such that, for each structured morphism $f : B \to GA$ in $\mathbf{K}$, there exists a unique $f^{\#} : A_B \to A$ such that $Gf^{\#}\eta_B = f$ is defined.

Proof. *(i)⇒ (ii):* By assumption, we have an F-cocovering semifull concrete subgraph $\mathbf{L}$ of $(F, \mathbf{A})$ and a concrete isomorphism $\Phi : \mathbf{K} \to \mathbf{L}$. For $B \in \mathrm{Ob}\,\mathbf{B}$, let

$$(\eta_B, A_B) = \Phi^{-1}(B, id_{FB})$$

(i.e. $A_B = FB$). (B, id_{FB}) is $\mathbf{L}$-strictly F-cowellcomposable because $\mathbf{L}$ is F-cocovering, and strongly $\mathbf{L}$-partially $\mathbf{A}$-wellcomposable; hence (η_B, FB) has the required wellcomposability properties by Lemma 4.6.

Now let $f : B \to GA$ be a G-structured morphism in $\mathbf{K}$. Any factorization $f = Gf^{\#}\eta_B$, where $f^{\#} : A_B \to A$, fulfills

$$\Phi(f, A) = \Phi(Gf^{\#}\eta_B, A) = (FB, f^{\#} id_{FB}) = (FB, f^{\#})$$

by Lemma 4.6, which establishes uniqueness of the factorization. On the other hand, defining $f^{\#}$ by the above equation, we have

$$(f, A) = \Phi^{-1}(FB, f^{\#} id_{FB}) = (Gf^{\#}\eta_B, A)$$

as required.

(ii)⇒(i): Define F on objects by $FB = A_B$ for each B, and given $f : B_1 \to B_2$ in $\mathbf{B}$, let $Ff : FB_1 \to FB_2$ be the unique morphism that makes the diagram

$$\begin{tikzcd} B_1 \arrow[r, "\eta_{B_1}"] \arrow[d, "f"'] & GFB_1 \arrow[d, dotted, "GFf"] \\ B_2 \arrow[r, "\eta_{B_2}"'] & GFB_2 \end{tikzcd}$$

commute; this morphism exists, because $(\eta_{B_2} f, FB_2)$ belongs to $\mathbf{K}$ by $\mathbf{K}$-strict $\mathbf{B}$-cowellcomposability of (η_{B_2}, FB_2).

It is obvious that F is a graph morphism, i.e. preserves identities. To check that F is a functor, let $g : B_2 \to B_3$ be a further morphism in $\mathbf{B}$ such that gf is defined. By well-associativity, the terms

$$\begin{aligned} \eta_{B_3}(gf) &= (\eta_{B_3} g) f \\ &= (GFg\eta_{B_2}) f \\ &= GFg(\eta_{B_2} f) \\ &= GFg(GFf\eta_{B_1}) \end{aligned}$$

are defined. Hence $G(FgFf)\eta_{B_1} = \eta_{B_3}(gf)$ is defined by strong $\mathbf{K}$-partial G-wellcomposability, which implies $FgFf = F(gf)$ as required. (Note in particular that $(GFf\eta_{B_1}, FB_1) = (\eta_{B_2} f, FB_2)$ and $(GFg(GFf\eta_{B_1}, FB_3)) = (\eta_{B_3}(gf), FB_3)$ belong to $\mathbf{K}$.)

Now define a concrete functor $\Phi : \mathbf{K} \to (F, \mathbf{A})$ on objects by $\Phi(f, A) = (B, f^{\#})$ for $f : B \to GA$, where $f^{\#} : FB \to A$ is the unique morphism such that $Gf^{\#}\eta_B = f$. It suffices to show that Φ is a semifull concrete graph morphism.

To see that Φ is a graph morphism, let $(a, b) : (f_1, A_1) \to (f_2, A_2)$ be a morphism in $\mathbf{K}$, where $f_i : B_i \to GA_i$, $i = 1, 2$. Then the terms

$$\begin{aligned} Gb(Gf_1^{\#}\eta_{B_1}) &= (Gf_2^{\#}\eta_{B_2})a \\ &= Gf_2^{\#}(\eta_{B_2} a) \\ &= Gf_2^{\#}(GFa\eta_{B_1}) \end{aligned}$$

are defined, which implies definedness of $bf_1^{\#} = f_2^{\#} Fa$; i.e. $(a, b) : \Phi(f_1, A_1) \to \Phi(f_2, A_2)$ is a morphism in $(F, \mathbf{A})$ as required. (Again note that all relevant terms are in $\mathbf{K}$ by stability of $\mathbf{K}$.)

It remains to show that Φ is semifull. To begin, let $(id_B, a) : \Phi(f_1, A_1) \to \Phi(f_2, A_2)$ be a morphism in $(F, \mathbf{A})$, i.e. we have $af_1^{\#} = f_2^{\#}$. Then the terms

$$\begin{aligned} f_2 &= Gf_2^{\#}\eta_B \\ &= G(af_1^{\#})\eta_B \\ &= Ga(Gf_1^{\#}\eta_B) \\ &= Gaf_1 \end{aligned}$$

are defined, i.e. $(id, a) : (f_1, A_1) \to (f_2, A_2)$ is a morphism as required.

Now let $(b, id_A) : \Phi(f_1, A) \to \Phi(f_2, A)$ be a morphism in $(F, \mathbf{A})$, where $f_i : B_i \to A$, $i = 1, 2$; i.e. we have $f_2^{\#} Fb = f_1^{\#}$. Then the terms

$$\begin{aligned} f_1 &= Gf_1^{\#}\eta_{B_1} \\ &= G(f_2^{\#}Fb)\eta_{B_1} \\ &= Gf_2^{\#}(GFb\eta_{B_1}) \\ &= Gf_2^{\#}(\eta_{B_2}b) \\ &= f_2 b \end{aligned}$$

are defined, i.e. $(b, id_A) : (f_1, A) \to (f_2, A)$ is a morphism. $\square$

Definition 4.10. A functor that satisfies the equivalent conditions in the above theorem is called $\mathbf{K}$*-partially adjoint (to F)*.

Remark 4.11. By lemma 4.6, the embedding $\mathbf{K} \hookrightarrow (F, \mathbf{A})$ as in the above theorem is full iff its image is stable.

Remark 4.12. If $\mathbf{K}$ is G-covering and G is $\mathbf{K}$-partially adjoint, then $\mathbf{K}$ consists (in the notation of condition (ii) in the above theorem) of all G-structured morphisms of the form $(Gf\eta_B, A)$, where $B \in \mathrm{Ob}\,\mathbf{B}$, $f : A_B \to A$, and $Gf\eta_B$ is defined (cf. Remark 4.8). In this case, $\mathbf{K}$-partial adjointness can also be characterized by means of adjoint situations; i.e. if $\mathbf{K}$ is G-covering, then the conditions above are equivalent to

(iii) There exists a functor $F : \mathbf{B} \to \mathbf{A}$ and natural transformations $\eta : \mathbf{B} \to GF$ and $\varepsilon : FG \to \mathbf{A}$ such that for each B, (η_B, FB) is $\mathbf{K}$-strictly $\mathbf{B}$-cowellcomposable and strongly $\mathbf{K}$-partially G-wellcomposable, dually for ε (where all conditions are w.r.t. $\mathcal{L} = \{(B, f) \in \mathrm{Ob}\,(F, \mathbf{A}) \mid Gf\eta_B$ is defined$\}$), and such that $G\varepsilon\eta_G = id_G$ and $\varepsilon_F F\eta = id_F$ are defined.

In particular, in the realm of categories, such a partial adjoint is already adjoint.

Proof. Theorem 4.9 (i),(ii)⇒ (iii): It is easily verified that the given partial universal arrows extend to a natural transformation $\eta : \mathbf{B} \to GF$, where $(\eta_f, FB_2) = \Phi^{-1}(B_1, Ff)$ for $f : B_1 \to B_2$ in $\mathbf{B}$, and that a dual construction yields a natural transformation $\varepsilon : FG \to \mathbf{A}$ (where $(GA_1, \varepsilon_g) = \Phi(Gg, A_2)$ for $g : A_1 \to A_2$ in $\mathbf{A}$).

Now let $f : A_1 \to A_2$ be a morphism in $\mathbf{A}$; we have to show $G\varepsilon_f\eta_{GA_2} = Gf$. We have a commutative square

$$\begin{tikzcd} FGA_1 \arrow[r, "Fid_{GA_1}"] \arrow[d, "id_{FGA_1}"'] & FGA_1 \arrow[d, "\varepsilon_f"] \\ FGA_1 \arrow[r, "\varepsilon_f"'] & A_2 \end{tikzcd} \quad ;$$

the image of the corresponding $\mathbf{L}$-morphism under Φ^{-1} (in the notation of the proof of

Theorem 4.9) yields the desired equation:

$$\begin{tikzcd} GA_1 \arrow[r, "id_{GA_1}"] \arrow[d, "\eta_{GA_1}"'] & GA_1 \arrow[d, "Gf"] \\ GFGA_1 \arrow[r, "G\varepsilon_f"'] & GA_2 \end{tikzcd} .$$

The other equation is shown dually.

(iii) ⇒ (ii) in Theorem 4.9: To show the universal property of (η_B, FB) for $B \in \operatorname{Ob}\mathbf{B}$, let $f : B \to GA$ be a G-structured morphism in $\mathbf{K}$. Then the terms $f = (G\varepsilon_A \eta_{GA})f = G\varepsilon_A(\eta_{GA} f) = G\varepsilon_A(GFf\eta_B)$ are defined; by strong $\mathbf{K}$-partial G-wellcomposability of (η_B, FB), this implies that $\varepsilon_A Ff$ is defined and extends f as required. Dually, $\varepsilon_A F(G\bar{f}\eta_B) = \bar{f}$ is defined for each F-costructured morphism $\bar{f} : FB \to A$ in $\mathcal{L}$; hence definedness of $G\bar{f}\eta_B = f$ implies $\bar{f} = \varepsilon_A Ff$, i.e. the factorization is unique. □

Combining this observation with Theorem 4.9 and Lemma 4.6 (vi) (and recalling the constructions in the respective proofs), we obtain the characterization theorem for 'proper' adjointness as in Schröder (1999):

Theorem 4.13. Let $G : \mathbf{A} \to \mathbf{B}$ be a functor between precategories. Equivalent are

(i) There exists a functor $F : \mathbf{B} \to \mathbf{A}$ and a concrete isomorphism $(\mathbf{B}, G) \cong (F, \mathbf{A})$.

(ii) For each object B in $\mathbf{B}$, there exists a *G-universal arrow*, i.e. a $\mathbf{B}$-cowellcomposable and strongly G-wellcomposable G-structured morphism $\eta_B : B \to GA_B$ such that, for each structured morphism $f : B \to GA$, there exists a unique $f^{\#} : A_B \to A$ such that $Gf^{\#}\eta_B = f$ is defined.

(iii) There exists a functor $F : \mathbf{B} \to \mathbf{A}$ and natural transformations $\eta : \mathbf{B} \to GF$ and $\varepsilon : FG \to \mathbf{A}$ such that for each B, (η_B, FB) is $\mathbf{B}$-cowellcomposable and strongly G-wellcomposable, dually for ε, and such that $G\varepsilon\eta_G = id_G$ and $\varepsilon_F F\eta = id_F$ are defined.

Remark 4.14. The composability conditions for ε in (iii) above can be left out; cf. Schröder (1999).

Remark 4.15. Using the construction of comma graphs via pullbacks, it is easily seen that, for any functors K and H with suitable domains and codomains, the comma categories (K, GH) and (FK, H) 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_2G_1 is adjoint to F_1F_2, since

$$(\mathbf{C}, G_2G_1) \cong (F_2, F_1) \cong (F_1F_2, \mathbf{A}).$$

Natural Isomorphisms

In order to use universal properties, i.e. adjunctions, as definitions for construction, 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; sticking to the general philosophy outlined above, we choose a definition via comma graphs, which, again, can be equivalently described by a suitable modification of the naive definition:

Theorem 4.16. Let $F, G : \mathbf{A} \to \mathbf{B}$ be functors between precategories, let $\mathbf{K}$ be an F-covering semifull concrete subgraph of $(\mathbf{B}, F)$, and let $\bar{\mathbf{K}}$ denote the full concrete subgraph of $(\mathbf{B}, F)$ spanned by $\mathbf{K}$. Equivalent are

(i) There exists a G-covering semifull concrete embedding $\mathbf{K} \hookrightarrow (\mathbf{B}, G)$.
(ii) There exists a G-covering full embedding $\bar{\mathbf{K}} \hookrightarrow (\mathbf{B}, G)$.
(iii) There exist natural transformations $\eta : F \to G$ and $\nu : G \to F$ such that $\eta\nu = id_G$ and $\nu\eta = id_F$; furthermore, $\eta_a f$ is defined whenever Faf is defined for $(f, A_1) \in \operatorname{Ob}\mathbf{K}$ and $a : A_1 \to A_2$, and an analogous condition w.r.t. G and $\mathcal{L}$ holds for ν, where $\mathcal{L} = \{(\eta_A f, A) \mid (f, A) \in \operatorname{Ob}\mathbf{K}\}$.

Proof. *(i)⇒(iii):* By assumption, we have a concrete isomorphism $\Phi : \mathbf{K} \to \mathbf{L}$, where $\mathbf{L}$ is a G-covering semifull concrete subgraph of $(\mathbf{B}, G)$ with $\operatorname{Ob}\mathbf{L} = \mathcal{L}$. Define η by

$$(\eta_f, A_2) = \Phi(Ff, A_2) \quad \text{for } f : A_1 \to A_2;$$

ν is constructed analogously using Φ^{-1}. Then $\eta : F \to G$ is natural: If $f : A_1 \to A_2$ is a morphism in $\mathbf{A}$, then we have a morphism $(id_{FA_1}, f) : (Fid_{A_1}, A_1) \to (Ff, A_2)$ in $\mathbf{K}$; application of Φ yields $Gf\eta_{A_1} = \eta_f$ as required:

$$\begin{tikzcd} FA_1 \arrow[r, "id_{FA_1}"] \arrow[d, "Fid_{A_1}"'] & FA_1 \arrow[d, "Ff"] & & FA_1 \arrow[r, "id_{FA_1}"] \arrow[d, "\eta_{A_1}"'] & FA_1 \arrow[d, "\eta_f"] \\ & & \overset{\Phi}{\mapsto} & & \\ FA_1 \arrow[r, "Ff"'] & FA_2 & & GA_1 \arrow[r, "Gf"'] & GA_2 \end{tikzcd} .$$

The second naturality equation and naturality of ν are shown analogously.

Thus we have, for f as above, a morphism $(\nu_{A_1}, id_{A_2}) : (\nu_f, A_2) \to (Ff, A_2)$ in $\mathbf{K}$; again, we obtain $\eta_f \nu_{A_1} = Gf$ by application of Φ, i.e. we have $\eta\nu = id_G$:

$$\begin{tikzcd} GA_1 \arrow[r, "\nu_{A_1}"] \arrow[d, "\nu_f"'] & FA_1 \arrow[d, "Ff"] & & GA_1 \arrow[r, "\nu_{A_1}"] \arrow[d, "Gf"'] & FA_1 \arrow[d, "\eta_f"] \\ & & \overset{\Phi}{\mapsto} & & \\ FA_2 \arrow[r, "Fid_{A_2}"'] & FA_2 & & GA_2 \arrow[r, "Gid_{A_2}"'] & GA_2 \end{tikzcd} .$$

The other equation is shown analogously.

Finally, if Faf is defined for $(f, A_1) \in \operatorname{Ob}\mathbf{K}$ and $a : A_1 \to A_2$, then we have a morphism $(f,a):(f,A_1)\to(Fa,A_2)$ in $\mathbf{K}$, and we obtain definedness of $\eta_a f$ via Φ:

$$\begin{tikzcd} \bullet \arrow[r, "f"] \arrow[d, "f"'] & FA_1 \arrow[d, "Fa"] \\ FA_1 \arrow[r, "Fa"'] & FA_2 \end{tikzcd} \quad \overset{\Phi}{\mapsto} \quad \begin{tikzcd} \bullet \arrow[r, "f"] \arrow[d] & FA_1 \arrow[d, "\eta_a"] \\ GA_1 \arrow[r, "Ga"'] & GA_2 \end{tikzcd} .$$

The definedness condition for ν is shown analogously.

(iii)⇒(ii): Denoting by $\mathbf{L}$ the full concrete subgraph of $(\mathbf{B},G)$ spanned by $\mathcal{L}$, we define a concrete functor $\Phi:\bar{\mathbf{K}}\to\mathbf{L}$ on objects by

$$\Phi(f,A)=(\eta_A f,A);$$

note that $\eta_A f$ is defined by assumption. Analogously, using ν instead of η, we define $\Psi:\mathbf{L}\to\bar{\mathbf{K}}$. It is clear that Φ and Ψ are mutually inverse on objects.

To see that Φ is actually a concrete functor, is suffices to show that Φ is a concrete graph morphism. To this end, let $(b,a):(f_1,A_1)\to(f_2,A_2)$ be a morphism in $\bar{\mathbf{K}}$, i.e. $Faf_1=f_2b$. Then $\eta_a f_1=(\eta_{A_2}Fa)f_1$ is defined by assumption, so by wellassociativity of $\mathbf{B}$ and by naturality of η, the terms

$$\begin{aligned}(\eta_{A_2}f_2)b &= \eta_{A_2}(f_2b)\\ &= \eta_{A_2}(Faf_1)\\ &= (\eta_{A_2}Fa)f_1\\ &= (Ga\eta_{A_1})f_1\\ &= Ga(\eta_{A_1}f_1)\end{aligned}$$

are also defined; i.e. $(b,a):\Phi(f_1,A_1)\to\Phi(f_2,A_2)$ is a morphism in $\mathbf{L}$ as required. Analogously, Ψ is a concrete functor; hence, Φ and Ψ are mutually inverse concrete isomorphisms — equivalently: $\Phi:\bar{\mathbf{K}}\to(\mathbf{B},G)$ is a full concrete embedding.

(ii)⇒(i): Trivial. □

Remark 4.17. Condition (ii) above is the only point up to now where well-associativity is apparently used in a crucial way.

Definition 4.18. Functors F and G that fulfill the equivalent conditions in the above theorem are called $\mathbf{K}$*-partially naturally isomorphic*; η as in condition (iii) is called a $\mathbf{K}$*-partial natural isomorphism.*

This now easily produces the proper uniqueness statement for partially adjoint functors:

Theorem 4.19. If G is $\mathbf{K}$-partially adjoint to F, i.e. we have $\Phi:\mathbf{K}\hookrightarrow(F,\mathbf{A})$ as in Theorem 4.9, then G is $\mathbf{K}$-partially adjoint to F' iff F and F' are $\Phi[\mathbf{K}]$-partially naturally coisomorphic.

Remark 4.20. The above theorem immediately produces the uniqueness statement for left adjoints of properly adjoint functors; in this case, $\Phi[K] = (F, \mathbf{A})$, i.e. the natural coisomorphism between two left adjoints and its inverse are objectwise wellcomposable.

Remark 4.21. While the definition of a natural isomorphism is somewhat difficult in the general context, as the above theorem demonstrates, in many precategories this will not be a problem at all, namely in the quite common case that any morphism with a left and right inverse is composable with arbitrary morphisms on both sides. In this situation, the reference to $\mathbf{K}$ in the above definition is, of course, redundant; i.e. a natural isomorphism can just be defined as a natural transformation with a twosided inverse.

This completes our theoretical considerations; we now turn to the applications in the area of probabilistic systems.

5. Probability Presentations

Recall that a *probability space* in all its generality is a triple $(\Omega, \quad, P)$, where Ω is a nonempty set (the *outcome space*), is a σ-algebra over Ω, and P is a probability measure on (see for instance Halmos (1950)). However, for the purpose of studying finite or even countable automata, it suffices to work with the following notion of a (countable) probability presentation.

Definition 5.1. A *probability presentation* is a pair (Ω, p) where Ω is a countable nonempty set and $p : \Omega \to {}^{+}_{0}$ 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 indistinctly for all probability presentations.

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

As discussed in Mateus et al. (1999), in order to obtain the desired universal characterizations of the aggregation, restriction, and interconnection of probabilistic automata, one should define morphisms between probability presentations as follows:

Definition 5.2. 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

$$P_1(B | f(\Omega_1)) = P(f^{-1}(B))$$

for each $B \subset \Omega_2$. The precategory of probability presentations and their morphisms is denoted by $\mathbf{Prob}$, the obvious forgetful functor to $\mathbf{Set}$ by Out (for 'outcome space').

Unlike the classical *measure preserving transformations* as defined e.g. in Halmos (1950), which are always probabilistically surjective, this definition allows the formation of actual subspaces of probability spaces; in the case of probabilistically surjective maps, the two definitions coincide. However, while measure preserving transformations behave well with respect to composition, probability presentation morphisms need not compose:

$$\begin{tikzcd}
 & x : 1/2 \arrow[dr] & \\
a : 1/2 \arrow[r] & y : 1/4 \arrow[r] & u : 3/4 \\
b : 1/2 \arrow[r, swap, "f"] & z : 1/4 \arrow[r, swap, "g"] & v : 1/4
\end{tikzcd}$$

Although both f and g are probability presentation morphisms, gf is not; i.e. **Prob** is not a category.

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

Lemma 5.3. 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 $p_2(\omega)/P_2(f(\Omega_1)) = P_1(f^{-1}(\omega))$ for each $\omega \in f(\Omega_1)$.

The composability properties discussed in Section 4 can be conveniently characterized in **Prob**:

Definition 5.4. Let (Ω_1, p_1) and (Ω_2, p_2) be probability presentations, and let $f : \Omega_1 \to \Omega_2$ be a map. f is called

(i) *probabilistically constant* if there exists $\omega \in f(\Omega_1)$ such that $p_2(\omega) = P_2(f(\Omega_1))$;
(ii) *probabilistically surjective* if $P_2(f(\Omega)) = 1$;
(iii) *probabilistically injective* if, whenever $f(\omega_1) = f(\omega_2)$, then $\omega_1 = \omega_2$ or $p_1(\omega_1) = p_1(\omega_2) = 0$.

Proposition 5.5. Let f be a morphism in **Prob**.

(i) f is wellcomposable iff f is probabilistically surjective or constant;
(ii) f is cowellcomposable iff f is probabilistically injective or probabilistically constant;
(iii) f is strongly wellcomposable iff f is surjective;
(iv) f is strongly cowellcomposable iff f is probabilistically injective.

Independent products and conditional probabilities

We now present some universal constructions in **Prob** that are going to be useful for combining probabilistic automata.

We recall the classical definition of the product of two probability spaces:

Definition 5.6. 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 $p_{12}(\omega_1, \omega_2) = p_1(\omega_1)p_2(\omega_2)$.

By definition, a precategory **A** has (binary) products iff the diagonal functor $\Delta : \mathbf{A} \to \mathbf{A}^2$ is coadjoint. Now it is easy to see that **Prob** does not have products in this sense; however, $\Delta : \mathbf{Prob} \to \mathbf{Prob}^2$ is partially coadjoint w.r.t. a suitable subgraph of

$(\Delta, \mathbf{Prob}^2)$, and the associated couniversal arrow for a pair of probability presentations is indeed their independent product.

Definition 5.7. A 2-source (i.e. a Δ-costructured morphism) (f_1, f_2) in **Prob** is called *independent* if

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

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

Proposition 5.8. Let $\mathcal{S} = (f_1, f_2)$ 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 $\mathcal{S}$, applied to the $g_i^{-1}(B_i)$, we obtain the required condition:

$$P(f_1^{-1} g_1^{-1}(B_1) \cap f_2^{-1} g_2^{-1}(B_2)) = P(f_1^{-1} g_1^{-1}(B_1)) \times P(f_2^{-1} g_2^{-1}(B_2)).$$

□

Corollary 5.9. **Ind** is a stable subgraph of $(\Delta, \mathbf{Prob}^2)$.

Lemma 5.10. Let (f_1, f_2) be an independent 2-source with domain (Ω, p); then

$$P_{12}(f_1(\Omega) \times f_2(\Omega) \setminus \langle f_1, f_2 \rangle(\Omega)) = 0.$$

Proof. Let $(\omega_1, \omega_2) \in f_1(\Omega) \times f_2(\Omega) \setminus \langle f_1, f_2 \rangle(\Omega)$; then $f_1^{-1}(\omega_1) \cap f_2^{-1}(\omega_2) = \emptyset$. Therefore $P(f_1^{-1}(\omega_1) \cap f_2^{-1}(\omega_2)) = 0$, and by independence $P(f_1^{-1}(\omega_1))P(f_2^{-1}(\omega_2)) = 0$. Since both f_1 and f_2 are morphisms, $p(\omega_1)p(\omega_2) = 0$ and thus $p(\omega_1, \omega_2) = 0$. □

Theorem 5.11. Independent products are **Ind**-partial Δ-couniversal arrows; i.e. Δ is **Ind**-partially coadjoint.

Proof. By Corollary 5.9, **Ind** is stable. Let (Ω_1, p_1) and (Ω_2, p_2) be probability presentations; denote the usual product projections of their independent product by π_1, π_2. Then the 2-source $\mathcal{P} = (\pi_1, \pi_2)$ is independent by construction.

Factorization property: Let (f_1, f_2) be an independent 2-source with domain (Ω, p) and codomains (Ω_1, p_1) and (Ω_2, p_2). We have to show that the map $f = \langle f_1, f_2 \rangle : (\Omega, p) \to (\Omega_1 \times \Omega_2, p_{12})$ is a morphism. Now for $(\omega_1, \omega_2) \in f(\Omega)$, we have

$$\begin{aligned} P(f^{-1}(\omega_1, \omega_2)) &= P(f_1^{-1}(\omega_1) \cap f_2^{-1}(\omega_2)) \\ &= P(f_1^{-1}(\omega_1))P(f_2^{-1}(\omega_2)) \\ &= \frac{p_1(\omega_1)p_2(\omega_2)}{P_1(f_1(\Omega))P_2(f_1(\Omega))} = \frac{p_{12}(\omega_1, \omega_2)}{P_{12}(f(\Omega))}, \end{aligned}$$

where the last equality was obtained by Lemma 5.10. Hence, by Lemma 5.3, f is a morphism in **Prob**.

*$\mathcal{P}$ is **Ind**-strictly $\mathbf{Prob}^2$-wellcomposable:* Let g_i be a morphism with domain (Ω_i, p_i), $i = 1, 2$. Then $g_i \pi_i$ is defined by Proposition 5.5, and by Lemma 5.8, $(g_1 \pi_1, g_2 \pi_2)$ is an independent 2-source.

$\mathcal{P}$ *is strongly* **Ind**-*partially* Δ-*cowellcomposable:* Let g and h be morphisms in **Prob** such that $(\pi_1 g, \pi_2 g)$ and $\mathcal{S} = ((\pi_1 g)h, (\pi_2 g)h)$ are independent 2-sources. Now gh is the factorization of $\mathcal{S}$ on the level of maps, and hence is a morphism by (i). $\square$

Restrictions of probabilistic automata will require the formation of Cartesian lifts in **Prob**; for the time being, we stick to a definition carried over verbatim from category theory:

Definition 5.12. Let $F : \mathbf{A} \to \mathbf{B}$ be an initial functor between precategories. A morphism $f : A \to C$ in $\mathbf{A}$ is called *Cartesian* (or *initial*) w.r.t. F if, whenever $Ffh = Fg$ is defined for morphisms $g : D \to C$ and $h : FD \to FA$, then h 'is an $\mathbf{A}$-morphism', i.e. there exists a morphism $\bar{h} : D \to A$ such that $F\bar{h} = h$. A *Cartesian lift* of an F-structured morphism (f, A) 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; in particular, Cartesian lifts w.r.t. *Out* are unique. We show that Cartesian lifts w.r.t. *Out* exist for most inclusions.

Proposition 5.13. Let (Ω, p) be a probability presentation, and let $m : \Omega_0 \hookrightarrow \Omega$ be an inclusion in *Set* such that $P(\Omega_0) > 0$. Then (Ω_0, p_0), where $p_0(\omega) = P(\{\omega\}|\Omega_0)$, is a Cartesian lift of $(m, (\Omega, p))$ w.r.t. *Out*.

Proof. $m : (\Omega_0, m_0) \to (\Omega, p)$ 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 $m \circ h : (\Omega_1, p_1) \to (\Omega, p)$ is a morphism; we have to verify that $h : (\Omega_1, p_1) \to (\Omega_0, p_0)$ is a morphism. Now for $\omega \in h(\Omega_1)$, we have

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

$\square$

6. The precategory of generative discrete probabilistic automata

We now provide the above mentioned universal characterization of the aggregation, interconnection, and restriction of generative probabilistic automata. As discussed in van Glabeek et al. (1995), the latter constitute a more general 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 Glabeek et al. (1995). Generative probabilistic automata are used as a semantic model of probabilistic concurrent programs (van Glabeek et al. 1995) and, in the more restrictive form of (a simplified type of) decision trees, in machine learning and knowledge representation (Pinto et al. 1999).

Definition 6.1. 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) $(\perp, s) \in T^s$, and if $(\perp, t) \in T^s$ then $t = s$;
(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. Each T^s is the set of *possible transitions* in s. Finally, each δ^s is the *transition probability distribution* in s.

Given the above results on probability spaces, the classical definition of a morphism between automata now has a natural generalization:

Definition 6.2. 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 $s \in S_1$.

The precategory of probabilistic automata and their morphisms is denoted by **PA**.

We adapt several classical notions to the setting of probabilistic automata:

Definition 6.3. Let M be a probabilistic automaton. We write $s \to^a t$ if $at \in T^s$. 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 6.4. 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 $w \in \{\perp\}^*$;
(iii) *without merges* if, for $s_1, s_2, t \in S$ and $a, b \in A$, $s_1 \to^a t$ and $s_2 \to^b t$ imply $s_1 = s_2$ and $a = b$.

The full subprecategory of **PA** spanned by all accessible automata is denoted by **AcPA**.

Combining probabilistic automata

We now turn our attention to three basic mechanisms for combining probabilistic automata: aggregation, action restriction and interconnection. 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 characterizations, is already present in van Glabeek et al. (1995).

Aggregation

It is now straightforward to define the independent product of two automata:

Definition 6.5. 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 subgraph of $(\Delta, \mathbf{PA}^2)$ spanned by all independent 2-sources is denoted by **IndA**.

Proposition 6.6. 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 *independent product*, given by $(S_1 \times S_2, A_1 \times A_2, \delta)$, 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. Straightforward, using Theorem 5.11. □

This product can be regarded as the free aggregation of M_1 and M_2, which we denote by $M_1 \| M_2$. Interaction between two automata will modeled as a restriction of their free aggregation.

Restriction and Interconnection

We now construct action restrictions of generative probabilistic automata as Cartesian lifts.

Let $Alph : \mathbf{PA} \to \mathbf{Set}_*$ denote the functor that maps each probabilistic automaton $M = (A, S, \delta)$ to its alphabet A.

Proposition 6.7. Let $M = (A, S, \delta)$ be a probabilistic automaton, and let $m : A_0 \hookrightarrow A$ be the inclusion of a subset containing $\perp$. Then (A, S_0, δ_0), where $T_0^s = T^s \cap (A_0 \times S_0)$ and δ_0^s is the Cartesian lift of $(m \times id^s|_{T^s}, \delta^s)$ w.r.t. Out, is a Cartesian lift of (m, M) w.r.t. $Alph$.

Proof. Straightforward. □

Definition 6.8. Let $M = (A, S, \delta)$ be a probabilistic automaton, and let $m : A_0 \to A$ be the inclusion of a subset containing $\perp$. Then the *restriction of* M *to* A_0, denoted by $M|_{A_0}$, is the Cartesian lift of (m, M).

So if we want to impose some interaction, for instance a *calling interaction* where an action a_1 of an automaton M_1 is only allowed in connection with an action a_2 of an automaton M_2, we have to compute the Cartesian lift of a suitable inclusion of alphabets. In the restricted alphabet, a_1 should always go together with a_2; the other pairs involving a_1 are dropped. Therefore, we reach the following notion of interconnection by action calling:

Definition 6.9. 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 respectively M_2. 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

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

7. Other probabilistic models

Accessible automata

In practice it is typical to reduce an automaton to its accessible part, since inaccessible states are redundant. In the theoretical framework built up until now, this construction can be formally captured in the following statement:

Proposition 7.1. **AcPA** is a coreflexive 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. (Proposition 7.1) Let $E : \mathbf{AcPA} \hookrightarrow \mathbf{PA}$ denote the inclusion functor; we have to show that E is coadjoint, i.e. that there exists an E-couniversal arrow for each probabilistic automaton $M = (A, S, \delta)$.

Now let S_a be the set of accessible states of M, and denote the inclusion $S_A \hookrightarrow S$ by i. Then $\varepsilon = (id_A, i) : M_a = (A, S_a, \delta) \to M$ is the required couniversal arrow: M_a is accessible by construction; since the restriction $\varepsilon^s = id_A \times i : \delta^s \to \delta^{i(s)}$ is an identity for each s, ε is wellcomposable and strongly E-cowellcomposable. Finally, if $M_1 = (A_1, S_1, \delta_1)$ 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. □

Decision trees

In many cases, it is useful to require that each state remembers the history of the automaton. In the case of generative probabilistic automata, this leads to the concept of decision trees.

Definition 7.3. A *decision tree* is an accessible and acyclic probabilistic automaton without merges. The full subprecategory of **AcPA** spanned by the decision trees is denoted by **DT**.

Let M be a probabilistic automaton. We define the set of *proper transitions* in M as $tr_M = \{(s, a, t) \in S \times (A \setminus \{\perp\}) \times S \mid s \to^a t\}$. A *path* in M is a sequence $w = (s_1, a_1, t_1) \dots (s_n, a_n, t_n)$ of proper transitions in M such that if $n > 0$ then $s_1 = i$ and $t_k = s_{k+1}$, $k = 1 \dots n-1$. We denote the set of all paths in M by pth_M. A morphism $f : M_1 \to M_2$ between two probabilistic automata $M_i = (S_i, A_i, \delta_i)$ 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 $lst_M : pth_M \to S$ in the following way:

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

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 coreflexive 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,

$$T^w_{pth} = \{(\perp, w)\} \cup \{(a, w(lst(w), a, t)) \mid a \neq \perp \text{ and } (a,t) \in T^{lst(w)}\},$$

and

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

is a decision tree, and

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

is an E-couniversal arrow: The required composability properties of ε follow from the fact that the restrictions

$$\varepsilon^w = (id_A, lst_M) : \delta^w_{pth} \to \delta^{lst(w)}$$

are, by construction, isomorphisms. Now let $D = (A_1, S_1, \delta_1)$ 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 are isomorphisms, is a morphism; since obviously $\varepsilon\tilde{f} = f$, we have the factorization required by the couniversal property. ☐

Remark 7.6. By Remark 4.15, **DT** is also a coreflexive subprecategory of **PA**.

8. Conclusions

The definition of adjointness in precategories, which was chosen for purely theoretical reasons, does describe several well-known constructions in the theory of probabilistic automata in an adequate manner; in this way, the theoretical basis for the said constructions is broadened, and the unified approach thus provided facilitates the extension of these and possibly other concepts to different settings. Further applications might conceivably be found in the areas of probability theory and stochastic processes.

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.

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.

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

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

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

Herrlich, H. and Strecker, G. E. (1979) *Category theory*, 2nd ed. Heldermann Verlag, Berlin.

Mateus, P., Sernadas, A. and Sernadas, C. (1999) Precategories for combining probabilistic automata. Proceedings of CTCS 99, *Electronic Notes in Theoretical Computer Science*, in print.

Pinto, J., Sernadas, A., Sernadas, C. and Mateus, P. (1999) Non-determinism and uncertainty in the situation calculus. In A. Kumar and I. Russell (editors), *Proceedings of the FLAIRS'99 - the 12th International Florida AI Research Symposium*, 454–460. AAAI Press, Menlo Park, California.

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

Schröder, L. (1999) *Composition graphs* (in German). PhD thesis, University of Bremen; also: Logos-Verlag, Berlin.

Schröder, L. and Herrlich, H. (1999) Free adjunction of morphisms. To appear in *Appl. Cat Struct.*

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