Math. Struct. in Comp. Science (1998), *vol.* 8, *pp.* 93–116. Printed in the United Kingdom

Minimal realization in bicategories of automata[†]

ROBERT ROSEBRUGH[‡], N. SABADINI[*]
and R. F. C. WALTERS[§]

[‡]*Department of Mathematics and Computer Science, Mount Allison University, Sackville, N. B. E0A 3C0 Canada*

[*]*Dipartimento di Scienze dell'Informazione, Università di Milano, 30135 Milano, Italy*

[§]*School of Mathematics and Statistics, University of Sydney, Sydney, NSW 2006 Australia*

Received 20 September 1996; revised 5 May 1997

The context of this article is the programme to develop monoidal bicategories with a feedback operation as an algebra of processes, with applications to concurrency theory. The objective here is to study reachability, minimization and minimal realization in these bicategories. In this setting the automata are 1-cells, in contrast with previous studies where they appeared as objects. As a consequence, we are able to study the relation of minimization and minimal realization to serial composition of automata using (co)lax (co)monads. We are led to define suitable behaviour categories and prove minimal realization theorems that extend classical results.

1. Introduction

Katis, Sabadini, Walters and Weld have described bicategories equipped with operations of serial and parallel composition, and feedback modelled as, respectively, composition of 1-cells, a tensor product and an operation called *feedback* (Katis *et al.*, 1997; Sabadini *et al.* 1994). The bicategories are constructed from a base category **C** with a symmetric monoidal tensor $\otimes$. Objects are those of **C** and arrows (or processess) from X to Y are pairs (U, α) where $\alpha : X \otimes U \longrightarrow U \otimes Y$. As mentioned above, composition models serial composition of circuits, there is a tensor product on circuits, and circuits from $X \otimes Q$ to $Q \otimes Y$ have a feedback operation whose result is a circuit from X to Y. In this article we concentrate on serial composition. When the tensor is cartesian product the 1-cells are called circuits and used to study physical devices. When the tensor is sum they are called Elgot automata and used as a model of algorithm. In Katis *et al.* (1997) behaviour functors for these bicategories are also considered.

In this article our objective is to study three bicategories of automata: the bicategory

[†] This work has been supported by NSERC Canada, Italian MURST and the Australian Research Council.

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

of Mealy automata $\mathscr{A}$, which adds an initial state to the circuit model; the bicategory of Elgot automata $\mathscr{E}$; and the bicategory of Σ-automata $\mathscr{F}$, which generalizes Elgot automata by labelling transitions from an alphabet Σ. The corresponding behaviours are, respectively, certain functions between input and output monoids, partial functions with duration and certain matrices of languages. In each case we study reachability and minimization, and prove a minimal realization theorem. Reachability and minimization are described by idempotent (co)monads. Since the automata are arrows rather than objects, we are able to extend classical results to relate serial composition of automata with the reachability and minimization (co)monads found.

In Section 2 we define the bicategory $\mathscr{A}$ whose 1-cells are circuits with an initial state. Except for the lack of a finiteness condition, these are the classical Mealy automata (Hopcroft and Ullman 1979) and we use that name. This provides a setting in which both reachability and minimization can be considered. Reachability is described by a comonad (as has already been noted by Adamek and Trnkova (1990)) on each hom category, and the coalgebras are the reachable automata. Minimization is described by monads on the hom categories, and the algebras are minimal automata. With an appropriate definition of the behaviour of Mealy automata, we are able to prove a minimal realization theorem that extends Nerode's theorem (Nerode 1958). It provides a variant of Goguen's minimal realization theory (Goguen 1972) and we extend this to include serial composition. The local situation, *i.e.*, in a single hom category, is summarized in the following diagram. The reachable automata from X to Y are denoted $\mathscr{A}_R(X,Y)$; the subcategory of minimized automata is $\mathscr{A}_R^M(X,Y)$, and behaviours from X to Y are denoted $\mathbf{B}_{\mathscr{A}}(X,Y)$. In the diagram, F is minimization, E is behaviour and N is minimal realization.

$$\begin{tikzcd}
& \mathscr{A}_R(X,Y) \arrow[dl, shift right=1.5, "F"'] \arrow[dr, shift left=1.5, "E"] & \\
(\mathscr{A}_R)^M(X,Y) \arrow[ur, shift right=1.5, "I"'] \arrow[ur, phantom, "\dashv"] \arrow[rr, shift left=1.5, "E'"] \arrow[rr, phantom, "\simeq"] & & \mathbf{B}_{\mathscr{A}}(X,Y) \arrow[ul, shift left=1.5, "N"] \arrow[ul, phantom, "\dashv"] \arrow[ll, shift left=1.5, "N'"]
\end{tikzcd}$$

In Section 3 we consider the bicategory $\mathscr{E}$ of Elgot automata, which model algorithms and whose natural semantics is a partial function with duration. We again find a local comonad for reachability, and a local monad for minimization. We prove a minimal realization theorem here as well.

In Section 4 we generalize to allow labelled transitions, defining the bicategory $\mathscr{F}$ of Σ-automata. Here the behaviour category has considerable interest – the arrows are matrices of languages with an 'anti-prefix' property.

To extend the results to the full process bicategories, *i.e.*, to take account of serial composition, requires the (co)lax (co)monads introduced by Carboni and Rosebrugh (1991). Reachability, minimization and our minimal realizations are idempotent

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

(co)lax (co)monads. In Section 5 we recall results on lax monads and consider the idempotent case, which concerns us.

Finally, in Section 6 we complete the above picture by showing that the minimization-minimal realization theory is compatible with serial composition. That is, the diagram above is valid in each case without the local restriction.

Throughout this article we are using the category **set** of sets as base category. In each section we use various algebraic properies of **set**. For the definition of the bicategory of Mealy automata we use products only. Elgot automata require sums only, and Σ-automata require the fact that **set** is a distributive category.

2. Mealy automata

We begin by defining a bicategory of circuits with initial state, which we call Mealy automata. The initial state allows us to define the behaviour of a Mealy automaton as a function between free monoids. We define a category of behaviours so that behaviour is a homomorphism of bicategories. For Mealy automata reachability is a useful concept and we find local comonad structures compatible with serial composition. Our main result in this section, Theorem 17, involves a realization of behaviours of reachable Mealy automata using a Nerode-type construction.

Definition 1. The bicategory $\mathscr{A}$ of *Mealy automata* in **set** has

— Objects: the objects $X, Y, \ldots$ of **set**
— Arrows: from X to Y are triples (U, α, u_0) where U is an object of **set**, $\alpha : X \times U \longrightarrow U \times Y$ and $u_0 : 1 \longrightarrow U$ (the input set is X, the output set is Y and the state set is U.)
— Identity arrow: on X is $(1, t, 1_1)$ where $t : X \times 1 \longrightarrow 1 \times X$
— 2-Cells: from (U, α, u_0) to (U', α', u'_0) are arrows $\theta : U \longrightarrow U'$ of **set** such that $\theta u_0 = u'_0$ and $(\theta \times Y) \cdot \alpha = \alpha' \cdot (X \times \theta)$
— Composition of arrows: if $(U, \alpha, u_0) : X \longrightarrow Y$ and $(V, \beta, v_0) : Y \longrightarrow Z$ then $(V, \beta, v_0)(U, \alpha, u_0) = (U \times V, (U \times \beta)(\alpha \times V), (u_0, v_0))$
— Vertical composition of 2-cells: if $\theta : (U, \alpha, u_0) \longrightarrow (U', \alpha', u'_0)$ and $\theta' : (U', \alpha', u'_0) \longrightarrow (U'', \alpha'', u''_0)$ then their vertical composite $\theta' \cdot \theta$ is the arrow $\theta'\theta$ of **set**
— Horizontal composition of 2-cells: if θ and ψ are horizontally composable their composite, denoted $\psi \circ \theta$ is $\theta \times \psi$ in **set**

Remark 2. If all references to initial state are removed from the preceding definition we obtain precisely the bicategory of circuits **Circ** as introduced in Katis *et al.* (1997). There is an evident forgetful homomorphism of bicategories $\mathscr{A} \longrightarrow$ **Circ**.

For further work, we first need to extend the domain of α to words in X^*, the free monoid on X.

Definition 3. Let $(U, \alpha, u_0) : X \longrightarrow Y$ be a Mealy automaton and write $\alpha = \langle \alpha_U, \alpha_Y \rangle$. Define $\alpha^*_U : X^* \times U \longrightarrow U$ inductively by: $\alpha^*_U(\epsilon, u) = u$ and for $w \in X^*$, $x \in X$: $\alpha^*_U(wx, u) = \alpha_U(x, \alpha^*_U(w, u))$. Similarly, $\alpha^*_Y : X^* \times U \longrightarrow Y^*$ is defined inductively by $\alpha^*_Y(\epsilon, u) = \epsilon$ and for $w \in X^*$, $x \in X$: $\alpha^*_Y(wx, u) = \alpha^*_Y(w, u)\alpha_Y(x, \alpha^*_U(w, u))$.

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

Note that $\alpha^*_U(wv,u) = \alpha^*_U(v,\alpha^*_U(w,u))$ and $\alpha^*_Y(wv,u) = \alpha^*_Y(w,u)\alpha^*_Y(v,\alpha^*_U(w,u))$. Both equations are easily proved by induction on the length of v and are needed below.

We say that $f : X^* \longrightarrow Y^*$ *preserves initial subwords* if whenever $w = w_1w_2$, we have $f(w) = f(w_1)v$ for some v in Y^*.

Definition 4. The category, $\mathbf{B}_{\mathscr{A}}$, of *behaviours* has the same objects as **set**. For objects X and Y the set of behaviours from X to Y, $\mathbf{B}_{\mathscr{A}}(X,Y)$, is the set of functions $f : X^* \longrightarrow Y^*$ for which f preserves initial subwords and length. Composition in $\mathbf{B}_{\mathscr{A}}$ is inherited from **set**.

This notion of behaviour is derived from that of *complete sequential machine mapping* (Ginsburg 1966). Under the condition of length preservation, the preservation of initial subwords implies that if $w = w'x$ for x in X then $f(w) = f(w')y$ for some y in Y.

Definition 5. Let $(U,\alpha,u_0) : X \longrightarrow Y$ be in $\mathscr{A}$ and $\alpha = \langle \alpha_U, \alpha_Y \rangle$. The *behaviour of* (U,α,u_0) is the arrow $E(U,\alpha,u_0) : X \longrightarrow Y$ in $\mathbf{B}_{\mathscr{A}}(X,Y)$ defined by $E(U,\alpha,u_0)(w) = \alpha^*_Y(w,u_0)$ for $w \in X^*$.

That the conditions for $E(U,\alpha,u_0)$ to be in $\mathbf{B}_{\mathscr{A}}(X,Y)$ are satisfied is easily proved by induction. So are the following lemmas.

Lemma 6. If there is a 2-cell $\phi : (U,\alpha,u_0) \longrightarrow (V,\beta,v_0)$, then for $w \in X^*$ and $u \in U$ we have $\phi(\alpha^*_U(w,u)) = \beta^*_V(w,\phi(u))$.

Proof. We procede by induction on the length of w. If $w = \epsilon$, we have $\phi(\alpha^*_U(\epsilon,u)) = \phi(u) = \beta^*_V(\epsilon,\phi(u))$. Next suppose $\phi(\alpha^*_U(w,u)) = \beta^*_V(w,\phi(u))$ and $x \in X$. We have

$$\begin{aligned} \phi(\alpha^*_U(wx,u)) &= \phi(\alpha(x,\alpha^*_U(w,u))) \\ &= \beta(x,\phi(\alpha^*(w,u)) \\ &= \beta(x,\beta^*_V(w,\phi(u)) \\ &= \beta^*_V(wx,\phi(u)), \end{aligned}$$

where the second equality is the definition of ϕ. The result follows. □

Lemma 7. If there is a 2-cell $\phi : (U,\alpha,u_0) \longrightarrow (V,\beta,v_0)$, then $E(U,\alpha,u_0) = E(V,\beta,v_0)$.

Proof. Again, we procede by induction on the length of $w \in X^*$. Suppose $w = \epsilon$. Then $E(U,\alpha,u_0)(\epsilon) = \alpha^*_Y(\epsilon,u_0) = \epsilon = \beta^*_Y(\epsilon,v_0) = E(V,\beta,v_0)(\epsilon)$. Next suppose $E(U,\alpha,u_0)(w) = E(V,\beta,v_0)(w)$ and $x \in X$. Then we get

$$\begin{aligned} E(U,\alpha,u_0)(wx) &= \alpha^*_Y(wx,u_0) \\ &= \alpha^*_Y(w,u_0)\alpha_Y(x,\alpha^*_U(w,u_0)) \\ &= \beta^*_Y(w,v_0)\beta_Y(x,\phi(\alpha^*_U(w,u_0))) \\ &= \beta^*_Y(w,v_0)\beta_Y(x,\beta^*_V(w,\phi(u_0))) \\ &= \beta^*_Y(w,v_0)\beta_Y(x,(\beta^*_V(w,v_0))) \\ &= \beta^*_Y(wx,v_0) = E(V,\beta,v_0)(wx), \end{aligned}$$

where the third equality is by the inductive assumption and the fourth uses the previous Lemma. □

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

Recall that the category $\mathbf{B}_{\mathscr{A}}$ may be viewed as a bicategory with discrete hom categories.

Proposition 8. Behaviour, E, extends to homomorphism of bicategories from $\mathscr{A}$ to $\mathbf{B}_{\mathscr{A}}$.

Proof. First, E is locally functorial by Lemma 7. A straightforward calculation using the equations after Definition 5 applied to a composite automaton shows that E preserves composition of 1-cells up to isomorphism. □

Minimization of automata classically proceeds in two steps: first non-reachable states are discarded, and then states with equivalent behaviour are identified. We consider local versions of these steps in the bicategory of Mealy automata beginning with reachability.

Definition 9. For an automaton $(U, \alpha, u_0) : X \longrightarrow Y$, the *reachable states* are $U_R = \{u \in U \mid \exists w \in X^* \ \alpha_U^*(w, u_0) = u\}$. The *reachable kernel* of (U, α, u_0) is $R(U, \alpha, u_0) = (U_R, \alpha_R, u_0)$, where α_R is the restriction of α.

We note immediately that $R : \mathscr{A}(X, Y) \longrightarrow \mathscr{A}(X, Y)$ is functorial: the function $\theta : U \longrightarrow U'$ defining a 2-cell $(U, \alpha, u_0) \longrightarrow (U', \alpha', u_0')$ clearly restricts to $\theta_R : U_R \longrightarrow U'_R$. R is also evidently idempotent and there is an inclusion of (U_R, α_R, u_0) in (U, α, u_0). These inclusions are components of a natural transformation $\rho : R \longrightarrow 1_{\mathscr{A}}(X, Y)$. Thus, we have the following proposition.

Proposition 10. R is an idempotent comonad on $\mathscr{A}(X, Y)$ with counit ρ.

Coalgebras for $R(= R(X, Y))$ are called *reachable* automata, and they define a full subcategory $\mathscr{A}_R(X, Y)$ of $\mathscr{A}(X, Y)$. In Section 6 we will need the following proposition.

Proposition 11. Let $(U, \alpha, u_0) : X \longrightarrow Y$ and $(V, \beta, v_0) : Y \longrightarrow Z$ be Mealy automata. The assignment $r_{UV}(u, v) = (u, v)$ defines a morphism of Mealy automata:

$$r_{UV} : R((V, \beta, v_0)(U, \alpha, u_0)) \longrightarrow R(V, \beta, v_0)R(U, \alpha, u_0) : X \longrightarrow Z.$$

Proof. The underlying function of the comparison is an inclusion that is compatible with the actions. Indeed, let $(u, v) \in (U \times V)_R$. We claim that $(u, v) \in U_R \times V_R$. To see this, recall that $(U \times \beta)(\alpha \times V)(x, u, v) = (\alpha_U(x, u), \beta_V(\alpha_Y(x, u), v), \beta_Z(\alpha_Y(x, u), v))$, and consequently $(U \times \beta)(\alpha \times V)^*_{U \times V}(w, (u, v)) = (\alpha_U^*(w, u), \beta_V^*(\alpha_Y^*(w, u), v))$. Thus, if $(U \times \beta)(\alpha \times V)^*_{U \times V}(w, (u_0, v_0)) = (u, v))$, then $(\alpha_U^*(w, u_0), \beta_V^*(\alpha_Y^*(w, u_0), v_0)) = (u, v)$. □

Next we consider state minimization for Mealy automata. As in classical automata theory, we define an equivalence relation on states and use the quotient set as states in constructing a 'minimal' automaton with the same behaviour.

The equivalence relation on states of (U, α, u_0) is defined by $u \sim_\alpha u'$ iff $\forall w \in X^*$ we have $\alpha_Y^*(w, u) = \alpha_Y^*(w, u')$. Thus states are declared equivalent if they have the same output for all of X^* under α^*. The quotient automaton is $M(U, \alpha, u_0) = (U_M, \alpha_M, [u_0])$, where $U_M = U/\sim_\alpha$ and α_M is defined on classes in U_M by $\alpha_M(x, [u]) = ([\alpha_U(x, u)], \alpha_Y(x, u))$, where $\alpha = \langle \alpha_U, \alpha_Y \rangle$. This construction is well-defined. We give an example.

Example 12. Consider the Mealy automaton from $X = \{a, b\}$ to $Y = \{0, 1\}$ whose states are $U = \{u_0, u_1, u_2\}$, start state is u_0 and action α is indicated in the following picture,

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

where, for example, $\alpha(a, u_0) = (u_1, 1)$

$$\begin{tikzcd}
u_0 \arrow[rr, "a/1"] \arrow[dr, "b/0"'] & & u_1 \arrow[loop right, "b/1"] \arrow[dl, shift left, "a/0"] \\
& u_2 \arrow[ur, shift left, "b/1"] \arrow[loop below, "a/0"] &
\end{tikzcd}$$

To determine $\sim_\alpha$, note that $\alpha_Y(u_0, a) = 1$, while $\alpha_Y(u_1, a) = 0$, so u_0 is not equivalent to u_1. On the other hand, an easy induction shows that $\alpha_Y^*(u_1, w) = \alpha_Y^*(u_2, w)$ for all $w \in X^*$ so $u_1 \sim_\alpha u_2$. Thus the minimized automaton has the following state diagram:

$$\begin{tikzcd}
{[u_0]_\alpha} \arrow[r, shift left, "a/1"] \arrow[r, shift right, "b/0"'] & {[u_1]_\alpha} \arrow[loop above, "a/0"] \arrow[loop below, "b/1"]
\end{tikzcd}$$

The behaviour of both the original and minimized automata is given by $f : X^* \longrightarrow Y^*$, where for w in X^*:

$$f(w) = \begin{cases} \epsilon & \text{if } w = \epsilon \\ 1v & \text{if } w = aw' \\ 0v & \text{if } w = bw' \end{cases}$$

and v is the image of w' under the homomorphism from X^* to Y^* mapping a to 0 and b to 1.

M is functorial on $\mathscr{A}(X, Y)$ and idempotent. The quotient mapping defines a 2-cell $\eta_{(U,\alpha,u_0)} : (U, \alpha, u_0) \longrightarrow M(U, \alpha, u_0)$. The $\eta_{(U,\alpha,u_0)}$ are components of a natural transformation from $1_{\mathscr{A}(X,Y)}$ to M.

Proposition 13. The functor M is an idempotent monad on $\mathscr{A}(X, Y)$ with unit η.

Algebras for $M(= M(X, Y))$ are called *minimal* automata and define a full subcategory $\mathscr{A}^M(X, Y)$ of $\mathscr{A}(X, Y)$. For use in Section 6 we note the following proposition.

Proposition 14. Let $(U, \alpha, u_0) : X \longrightarrow Y$ and $(V, \beta, v_0') : Y \longrightarrow Z$ be Mealy automata. The assignment $m_{UV}([u], [v]) = [(u, v)]$ defines a 1-cell of Mealy automata:

$$m_{UV} : M(V, \beta, v_0')M(U, \alpha, u_0) \longrightarrow M((V, \beta, v_0')(U, \alpha, u_0)) : X \longrightarrow Z.$$

Proof. The underlying function of the comparison is easily described. Indeed, for $([u], [v]) \in U_M \times V_M$ we define $m_{UV}([u], [v]) = [(u, v)] \in (U \times V)_M$. To see that this is well-defined, recall that $(U \times \beta)(\alpha \times V)(x, u, v) = (\alpha_U(x, u), \beta_V(\alpha_Y(x, u), v), \beta_Z(\alpha_Y(x, u), v))$. Suppose that $u \sim u'$ and $v \sim v'$. Denote the action of the composite automaton $(V, \beta, v_0')(U, \alpha, u_0)$ by γ, so for any $x \in X$ we have $\gamma_Z(x, (u, v)) = \beta_Z(\alpha_Y(x, u), v) = \beta_Z(\alpha_Y(x, u'), v) = \beta_Z(\alpha_Y(x, u'), v') = \gamma_Z(x, (u', v'))$. Consequently, for any $w \in X^*$ we have $\gamma_Z^*(w, (u, v)) = \gamma_Z(w, (u', v'))$, so $[(u, v)] = [(u', v')]$. Similar arguments show that m_{UV} is a morphism of $\mathscr{A}$. $\square$

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

We observe that taking the reachable kernel and the minimization for Mealy automata are processes that commute up to isomorphism, *i.e.*, the minimization of the reachable kernel of (U, α, u_0) is isomorphic to the reachable kernel of its minimization. These are simply seen from the definitions above. Consequently, the minimization monad restricts to a monad M' on $\mathscr{A}_R(X, Y)$ and the reachability comonad restricts to a comonad R' on $\mathscr{A}^M(X, Y)$. The category of algebras for the restriction of M is isomorphic to the coalgebras for the restriction of R. The situation we have been describing is summed up in the following diagram. The pairs of functors are adjoint and both the inner and outer squares commute. The I's (respectively, J's) are inclusions adjoint to the reflectors F and F' (respectively, coreflectors G and G'.)

$$\begin{tikzcd}
\mathscr{A}(X,Y) \arrow[rrr, shift left, "G"] \arrow[d, shift right, "F"'] & & & \mathscr{A}_R(X,Y) \arrow[lll, shift left, "J"] \arrow[d, shift left, "F'"] \\
\mathscr{A}^M(X,Y) \arrow[r, shift right, "G'"'] \arrow[u, shift right, "I"'] & (\mathscr{A}^M)_{R'}(X,Y) \arrow[l, shift right, "J'"'] & \cong & (\mathscr{A}_R)^{M'}(X,Y) \arrow[u, shift left, "I'"]
\end{tikzcd}$$

Our next objective is the adjunction between minimal realization and behaviour. Though we have defined behaviour for an arbitrary automaton, the realization of a behaviour constructed below is necessarily reachable, so our adjunction refers to $\mathscr{A}_R(X, Y)$. We begin construction of the minimal realization of a behaviour by defining a crucial equivalence relation. Let $f : X \longrightarrow Y$ be a behaviour from X to Y. For $w, w' \in X^*$, we define $w \sim_f w'$ iff

$$\forall v \in X^* \quad f(wv) = f(w)y \text{ and } f(w'v) = f(w')y' \Rightarrow y = y'.$$

It is easy to check that $\sim_f$ is indeed an equivalence relation and we denote equivalence classes by $[w]_f$.

Definition 15. The *Nerode automaton* of a behaviour $f : X \longrightarrow Y$ is the Mealy automaton $Nf = (X^*/{\sim_f}, \alpha_f, [\epsilon]_f)$ with $\alpha_f(x, [w]_f) = ([wx]_f, y)$ where $w \in X^*, x \in X, y \in Y$ and $f(wx) = f(w)y$.

We first have to ensure that α_f is well-defined, *i.e.*, if $w \sim_f w'$, then $wx \sim_f w'x$ and $y = y'$ where $f(w'x) = f(w')y$. For any $v \in X^*$ we have $f(wxv) = f(wx)z = f(w)yz$ for some $y \in Y$ and $z \in Y^*$. Similarly, $f(w'xv) = f(w'x)z' = f(w')y'z'$. Since $w \sim_f w'$, we conclude that $yz = y'z'$, whence $y = y', z = z'$, and so also $wx \sim_f w'x$, as required.

We also note that Nf is reachable by its definition. As an example, we construct the Nerode automaton for the behaviour found in Example 12 above.

Example 16. Recall that f in $\mathbf{B}_{\mathscr{A}}(X, Y)$ was defined (as a function from X^* to Y^*) by the formula:

$$f(w) = \begin{cases} \epsilon & \text{if } w = \epsilon \\ 1u & \text{if } w = aw' \\ 0u & \text{if } w = bw', \end{cases}$$

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

and u is the image of w' under the homomorphism from X^* to Y^* mapping a to 0 and b to 1. We need to determine the equivalence relation $\sim_f$ and its classes. This is straightforward since it is easy to show that

(i) $a \sim_f b \sim_f aw \sim_f bw$ for any $w \in X^*$, and
(ii) $\epsilon \not\sim_f a$.

To see the relations (i), note that if v is arbitrary in X^*, then for any $w \in \{a, b, aw', bw'\}$ we have $f(wv) = f(w)u$, where u is the image of v under the homomorphism above, independent of $f(w)$. For (ii) it is enough to observe that $f(\epsilon b) = f(\epsilon)0$ while $f(ab) = f(a)1$ and $0 \neq 1$ whence $\epsilon \not\sim_f a$. Now the action on Nf and the isomorphism of Nf with the minimized automaton displayed in Example 12 are obvious.

The following result is a variant of Goguen's adjunction between minimal realization and behaviour. He considered machines that emitted a single output letter after reading the entire input. His behaviours were arbitrary functions from X^* to Y. We have taken account of the entire output sequence and consequently need the more complete definition of behaviours found above. If $Y = \{0, 1\}$ and all objects are finite sets, the result is a version of Nerode's Theorem (Nerode 1958).

Theorem 17. The behaviour of the Nerode automaton of f is f, that is, $ENf = f$. Moreover, we have

$$E \dashv N : \mathbf{B}_{\mathscr{A}}(X, Y) \longrightarrow \mathscr{A}_R(X, Y).$$

Proof. We prove the first statement by induction. First, $E(Nf)(\epsilon) = \epsilon = f(\epsilon)$. Now let $w \in X^*$ and $x \in X$. Assuming $E(Nf)(w) = f(w)$, we have

$$\begin{aligned} E(Nf)(wx) &= E(Nf)(w)(\alpha_f)_Y(x, (\alpha_f)^*_U(w, [\epsilon]_f)) \\ &= f(w)(\alpha_f)_Y(x, (\alpha_f)^*_U(w, [\epsilon]_f)) \\ &= f(w)(\alpha_f)^*_Y(x, [w]_f) \\ &= f(w)y \text{ where } f(w)y = f(wx) \\ &= f(wx). \end{aligned}$$

The desired equality of behaviours follows.

For the stated adjunction, we need to show that 2-cells (in $\mathbf{B}_{\mathscr{A}}(X, Y)$) from $E(U, \alpha, u_0)$ to f are in natural bijection with 2-cells (in $\mathscr{A}_R(X, Y)$) from (U, α, u_0) to Nf. Since $\mathbf{B}_{\mathscr{A}}$ is locally discrete, this amounts to showing that $E(U, \alpha, u_0) = f$ if and only if there is a unique 2-cell from (U, α, u_0) to Nf.

For sufficiency, we observe that if there is $(U, \alpha, u_0) \longrightarrow Nf$, then $E(U, \alpha, u_0) = ENf = f$, by Lemma 7 and the result of the previous paragraph.

For necessity, suppose $E(U, \alpha, u_0) = f$ and we define a unique 2-cell $\phi : (U, \alpha, u_0) \longrightarrow Nf$. We begin by recalling that (U, α, u_0) is reachable and define $\overline{\phi} : U \longrightarrow U_f$ by $\overline{\phi}(u) = [w]_f$ for some $w \in X^*$ such that $\alpha^*_U(w, u_0) = u$. We need to show that $\overline{\phi}$ is well-defined, that it defines a 2-cell in $\mathscr{A}_R$ and that it is the only such 2-cell.

We show first that $\overline{\phi}(u)$ does not depend on the choice of $w \in X^*$ such that $\alpha^*_U(w, u_0) = u$. Indeed, suppose that $\alpha^*_U(w, u_0) = u = \alpha^*_U(w', u_0)$. For $v \in X^*$, let $f(wv) = f(w)y$ and $f(w'v) = f(w')y'$. For brevity, denote $E(U, \alpha, u_0)$ by E and recall that $E = f$, so

$f(w)y = f(wv) = E(wv) = E(w)\alpha^*_Y(v,u) = f(w)\alpha^*_Y(v,u)$, and we conclude $y = \alpha^*_Y(v,u)$. Similarly, $f(w')y' = f(w')\alpha^*_Y(v,u)$. Thus $y = \alpha^*_Y(v,u) = y'$.

Next, $\overline{\phi}(u_0) = [\epsilon]_f$, the initial state of Nf, and to see that $\alpha_f(X \times \overline{\phi}) = (\overline{\phi} \times Y)\alpha$, let $(x,u) \in X \times U$. Suppose $\alpha^*_U(w,u_0) = u$, since (U,α,u_0) is reachable. Now $\alpha_f(X \times \overline{\phi})(x,u) = \alpha_f(x,\overline{\phi}(u)) = \alpha_f(x,\overline{\phi}(\alpha^*_U(w,u_0)) = \alpha_f(x,[w]_f) = ([wx]_f,y)$ and $f(wx) = f(w)y$. On the other hand,

$$\begin{aligned}
(\overline{\phi} \times Y)\alpha(x,u) &= (\overline{\phi}(\alpha_U(x,u)),\alpha_Y(x,u)) \\
&= (\overline{\phi}(\alpha_U(x,\alpha^*_U(w,u_0)),\alpha_Y(x,\alpha^*_U(w,u_0))) \\
&= (\overline{\phi}(\alpha^*_U(wx,u_0),y') \\
&\quad \text{where } \alpha^*_Y(wx,u_0) = E(U,\alpha,u_0)(wx) = f(wx) = f(w)y' \\
&= ([wx]_f,y).
\end{aligned}$$

Finally, we show $\overline{\phi}$ is the unique 2-cell from (U,α,u_0) to Nf. If ψ is another such 2-cell, then $\psi(u_0) = [\epsilon] = \overline{\phi}(u_0)$ is necessary, so $\psi = \overline{\phi}$ on all states reachable from u_0 by words of length 0. Now assume $\psi = \overline{\phi}$ on all states reachable from u_0 by words of length $\mid w \mid$ or less. For any $x \in X$ if $u = \alpha^*_U(wx,u_0)$, letting $u' = \alpha^*_U(w,u_0)$, we have

$$\begin{aligned}
\psi(u) &= \psi(\alpha_U(x,u')) \text{ by definition of } u' \\
&= (\alpha_f)_U(x,\psi(u')) \text{ since } \psi \text{ in } \mathscr{A}_R \\
&= (\alpha_f)_U(x,\overline{\phi}(u')) \text{ by hypothesis} \\
&= \overline{\phi}(\alpha_U(x,u')) \\
&= \overline{\phi}(u).
\end{aligned}$$

So $\psi = \overline{\phi}$. □

We consider the equivalence of reachable minimized automata and behaviours in Section 6. With that exception, the theorem above completes the description of reachability, minimization and minimal realization for Mealy automata summarized in the diagram in the Introduction.

3. Elgot automata

This section studies a bicategory of automata that can be used to model algorithms. The name arises from Elgot's work on sequential algorithms. Elgot automata have been used by Sabadini, Walters and Vigna (Sabadini *et al.* 1996) to define partial recursive functions, and by Vigna (Vigna 1996) to define the Blum–Shub–Smale computable functions (Blum *et al.* 1989).

Definition 18. (Katis *et al.*, 1997) The bicategory $\mathscr{E}$ of *Elgot automata* in **set** has

— Objects: the objects $X, Y, \ldots$ of **set**

— Arrows: from X to Y are pairs (U,α) where U is an object of **set** (called the *internal states* of (U,α)) and $\alpha : X + U \longrightarrow U + Y$ (the *transition morphism*)

— Identity arrow: on X is $(0, 1_X)$

— 2-Cells: from (U,α) to (U',α') are functions $\theta : U \longrightarrow U'$ of **set** such that $(\theta + Y)\cdot\alpha = \alpha' \cdot (X + \theta)$

— Composition of arrows: if $(U,\alpha) : X \longrightarrow Y$ and $(V,\beta) : Y \longrightarrow Z$, then $(V,\beta)(U,\alpha) = (U+V,(U+\beta)(\alpha+V))$
— Vertical composition of 2-cells: if $\theta : (U,\alpha) \longrightarrow (U',\alpha')$ and $\theta' : (U',\alpha') \longrightarrow (U'',\alpha'')$, their vertical composite $\theta' \cdot \theta$ is the function $\theta'\theta$ of **set**
— Horizontal composition of 2-cells: if θ and ψ are horizontally composable, their composite, denoted $\psi \circ \theta$, is $\theta + \psi$ in **set**

The semantics of an Elgot automaton might be viewed simply as the partial function from X to Y given, where defined, by the unique value in Y resulting from iterating α one or more times. To obtain our minimal realization theorems, we will need also to record the 'duration' of the process. We use the notation '$\rightharpoonup$' to denote a partial function.

Definition 19. Let $(U,\alpha) : X \longrightarrow Y$ be an Elgot automaton. The *behaviour* of (U,α) is the partial function $E(U,\alpha) : X \rightharpoonup Y \times \mathbb{N}$ defined by $E(U,\alpha)(x) = (y,n)$ if $\alpha^{n+1}(x) = y \in Y$ (and undefined otherwise).

Motivated by the preceding definition, we define a category of behaviours $\mathbf{B}_{\mathscr{E}}$ to have the same objects as **set**, and as arrows from X to Y, the partial functions from X to $Y \times \mathbb{N}$. In $\mathbf{B}_{\mathscr{E}}$ the composite of $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ is defined by $gf(x) = (z, m+n)$ when both $f(x) = (y,n)$ and $g(y) = (z,m)$ are defined, and undefined otherwise. As we observed for Mealy automata, we have the following lemma.

Lemma 20. If there is a 2-cell $\phi : (U,\alpha) \longrightarrow (U',\alpha')$ in $\mathscr{E}$, then $E(U,\alpha) = E(U',\alpha')$.

Proof. The proof is by induction from the fact that a 2-cell of automata is a function between state objects that commutes with the action. □

Viewing $\mathbf{B}_{\mathscr{E}}$ as a bicategory with discrete hom categories we get the following proposition.

Proposition 21. Behaviour, E, extends to homomorphism of bicategories from $\mathscr{E}$ to $\mathbf{B}_{\mathscr{E}}$.

Proof. First, E is locally functorial by Lemma 20. It is easy to see that the behaviour of a serial composite of Elgot automata is the composition in $\mathbf{B}_{\mathscr{E}}$ of their behaviours. □

Definition 22. Let $(U,\alpha) : X \longrightarrow Y$ be an Elgot automaton. The object of *reachable states* of (U,α) is

$$U_R = \{u \in U \mid \exists x \in X \; \exists n \in \mathbb{N} \; \alpha^n(x) = u\}.$$

The *reachable kernel* of (U,α) is the automaton $R(U,\alpha) = (U_R,\alpha_R) : X \longrightarrow Y$, where $\alpha_R : X + U_R \longrightarrow U_R + Y$ is the restriction of α.

The first thing to observe is that $R : \mathscr{E}(X,Y) \longrightarrow \mathscr{E}(X,Y)$ is functorial and idempotent, and that there is a 2-cell $\rho_{(U,\alpha)} : R(U,\alpha) \longrightarrow (U,\alpha)$ that is the component of a natural transformation from R to $1_{\mathscr{E}(X,Y)}$. Each of these facts follows after a short diagram chase. Moreover, it is easy to see that $R\rho = \rho R$, since each amounts to a transformation with identity components. We summarize this in the following proposition.

Proposition 23. The functor R is an idempotent comonad on $\mathscr{E}(X,Y)$ with counit ρ.

Corollary 24. The behaviour of the reachable kernel, $R(U,\alpha)$, of an Elgot automaton, (U,α), is the same as that of (U,α).

Coalgebras for the local reachability comonads are 'reachable' Elgot automata, *i.e.*, automata all of whose internal states are visited under the iterated action of α on at least one $x \in X$. We note, for later use, a comparison between the reachable kernel of a composite and the composite of reachable kernels.

Proposition 25. If $(U, \alpha) : X \longrightarrow Y$ and $(V, \beta) : Y \longrightarrow Z$, there is a canonical 2-cell $r_{UV} : R((U, \alpha)(V, \beta)) \longrightarrow R(U, \alpha)R(V, \beta)$.

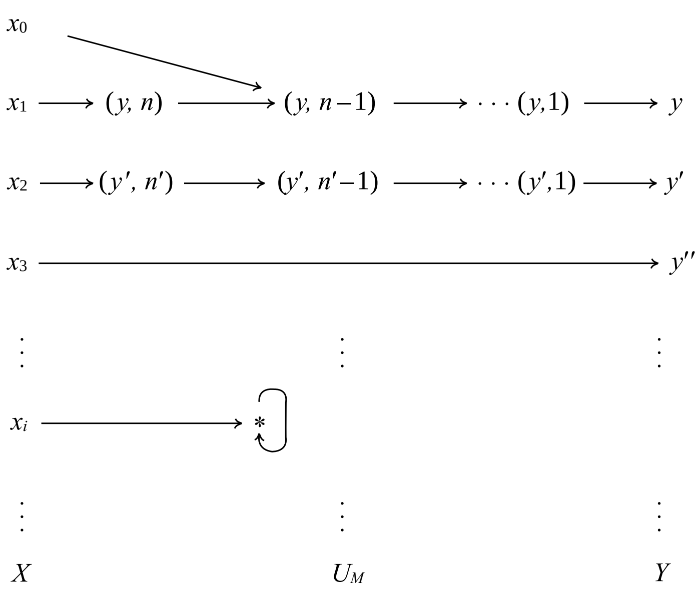
Proof. To see this, we observe that if $w \in (U + V)_R$, then $w \in U_R + V_R$, and that the appropriate restrictions of α and β are defined. □

We have a minimization theory for Elgot automata that will lead to a particularly simple description of minimized automata. We begin with an equivalence relation on states of (U, α): $u \sim_\alpha u'$ iff for all $n > 0$, for all $y \in Y$, $\alpha^n(u) = y$ iff $\alpha^n(u') = y$. Thus states are declared equivalent if they reach the same point in Y after the same duration, or if they both never reach Y. We can construct a 'quotient' automaton $M(U, \alpha) = (U_M, \alpha_M)$. We define $U_M = U/\sim_\alpha$, and α_M is defined on $X + U_M$ by

$$\alpha_M(x) = \begin{cases} [u] & \text{if } \alpha(x) = u \in U \\ y & \text{if } \alpha(x) = y \in Y \end{cases} \qquad \alpha_M([u]) = \begin{cases} [u'] & \text{if } \alpha(u) = u' \in U \\ y & \text{if } \alpha(u) = y \in Y. \end{cases}$$

Proposition 26. The quotient arrow $\eta : U \longrightarrow U_M$ underlies a 2-cell in $\mathscr{E}$ denoted $\eta_{(U,\alpha)} : (U, \alpha) \longrightarrow M(U, \alpha)$. Applying M to it gives an isomorphism, and M is an idempotent monad on $\mathscr{E}(X, Y)$.

Any algebra for M is a reachable automaton isomorphic to one of the following. States are (some of) the pairs consisting of an element of y in Y and a positive integral 'duration to Y' (plus possibly a 'non-terminating' state). The action on input x is direct transition from X to Y or direct transition from X to an internal state. On state (y, n) the action is 'reduction of duration' to $(y, n - 1)$ when $n \geqslant 2$, and $(y, 1)$ to y. The picture below illustrates the idea, and guides the proof of the preceding proposition.



https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

Our next objective is a minimal realization of any behaviour in $\mathbf{B}_{\mathscr{E}}$. The idea is simply to construct an automaton like that pictured above for a specified behaviour. Let $f : X \longrightarrow Y \times \mathbb{N}$ be a behaviour from X to Y. The state set for the minimal realization automaton Nf is

$$U_f = \begin{cases} \{(y,m) \mid \exists n \geqslant m \geqslant 1 \quad \exists x \in X \quad f(x) = (y,n)\} & \text{if } f \text{ is fully defined} \\ \{(y,m) \mid \exists n \geqslant m \geqslant 1 \quad \exists x \in X \quad f(x) = (y,n)\} \bigcup \{*\} & \text{otherwise.} \end{cases}$$

The action for the minimal automaton is defined on X by

$$\alpha_f(x) = \begin{cases} f(x) & \text{if } p_2(f(x)) > 0 \\ p_1(f(x)) & \text{if } p_2(f(x)) = 0 \\ * & \text{if } f(x) \text{ not defined,} \end{cases}$$

where the p_i are projections from $Y \times \mathbb{N}$. On U_f we define

$$\alpha_f(u) = \begin{cases} (y, n-1) & \text{if } u = (y,n) \text{ and } n > 1 \\ y & \text{if } u = (y,1) \\ * & \text{if } u = *. \end{cases}$$

This makes $\alpha_f : X + U_f \longrightarrow U_f + Y$, and we note that the automaton $Nf = (U_f, \alpha_f)$ is reachable.

Proposition 27. The behaviour of $Nf = (U_f, \alpha_f)$ is f. Furthermore, we have E is left adjoint to $N : \mathbf{B}_{\mathscr{E}}(X,Y) \longrightarrow \mathscr{E}_R(X,Y)$.

Proof. The diagram above indicates why the first statement holds: the constructed automaton simply has states that provide transitions of correct duration for elements of X where f is defined and a loop elsewhere.

For the adjunction, we show that 2-cells (in $\mathbf{B}_{\mathscr{E}}(X,Y)$) from $E(U,\alpha)$ to f correspond to 2-cells (in $\mathscr{E}_R(X,Y)$) from (U,α) to Nf. Since $\mathbf{B}_{\mathscr{E}}$ is locally discrete, that is to show that $E(U,\alpha) = f$ if and only if there is a unique 2-cell from (U,α) to Nf.

For sufficiency, observe that if there is $(U,\alpha) \longrightarrow Nf$, then $E(U,\alpha) = ENf = f$, by Lemma 20 and the previous paragraph.

For necessity, we suppose $E(U,\alpha) = f$ and seek to define a unique 2-cell $\phi : (U,\alpha) \longrightarrow Nf$. Recalling that (U,α) is reachable, we define $\phi : U \longrightarrow U_f$ by

$$\phi(u) = \begin{cases} (y,n) & \text{if } \alpha^n(u) = y \text{ for some } n > 0 \\ * & \text{if there is no such } y. \end{cases}$$

We need to show that ϕ is well-defined, that it defines a 2-cell in $\mathscr{E}_R$, and that it is the only such 2-cell. The first two follow immediately from $E(U,\alpha) = f$. For the last, simply observe that (y,n) is the only state of U_f for which $\alpha_f^n(y,n) = y$, while $*$ is the only 'looping' state. Hence, the requirement that ϕ be a morphism leaves no choice in the definition of $\phi(u)$. □

4. Σ-automata and matrices of languages

Let Σ be an alphabet, which we fix for this section. The model in the preceding section is here generalized to allow deterministic state transitions labelled by elements of Σ.

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

The resulting behaviours are certain matrices of languages. Non-deterministic automata whose behaviours are also matrices have been considered by Bloom, Sabadini and Walters (Bloom *et al.* 1994).

Definition 28. The bicategory $\mathscr{F}$ of Σ-*automata* in **set** has

— Objects: the objects $X, Y, \ldots$ of **set**
— Arrows: from X to Y are pairs (U, α) where U is an object of **set** (called the *internal states* of (U, α)) and $\alpha : X+(U\times\Sigma) \longrightarrow U+Y$ the *transition morphism*, with components $\alpha_X : X \longrightarrow U + Y$ and $\alpha_U : U \times \Sigma \longrightarrow U + Y$
— Identity arrow: on X is $(0, 1_X)$
— 2-Cells: from (U, α) to (U', α') are functions $\theta : U \longrightarrow U'$ of **set** such that $(\theta + Y)\cdot\alpha = \alpha' \cdot (X + (\theta \times \Sigma))$
— Composition of arrows: if $(U, \alpha) : X \longrightarrow Y$ and $(V, \beta) : Y \longrightarrow Z$, then $(V, \beta)(U, \alpha) = (U + V, (U + \beta)(\alpha + (V \times \Sigma))(X + \delta))$, where $\delta : (U + V) \times \Sigma \longrightarrow U \times \Sigma + V \times \Sigma$ is the distributive law
— Vertical composition of 2-cells: if $\theta : (U, \alpha) \longrightarrow (U', \alpha')$ and $\theta' : (U', \alpha') \longrightarrow (U'', \alpha'')$, their vertical composite $\theta' \cdot \theta$ is the function $\theta'\theta$ of **set**
— Horizontal composition of 2-cells: if θ and ψ are horizontally composable, their composite, denoted $\psi \circ \theta$, is $\theta + \psi$ in **set**

The idea here is that transitions among states of a Σ-automaton are labelled by elements of Σ. An Elgot automaton is essentially the special case where Σ has one element. We write Σ^+ for the free semi-group on Σ (or the words of length one or more in Σ^*.)

Definition 29. Let $(U, \alpha) : X \longrightarrow Y$ be a Σ-automaton. Define a partial function $\alpha_U^+ : U \times \Sigma^+ \rightharpoonup U + Y$ as follows. For $a \in \Sigma$ and $w \in \Sigma^+$

$$\alpha_U^+(u, a) = \alpha_U(u, a) \quad \alpha_U^+(u, wa) = \begin{cases} \alpha_U(\alpha^+(u, w), a) & \text{if } \alpha_U^+(u, w) \in U \\ \text{undefined otherwise.} \end{cases}$$

Define a partial function $\alpha^* : X \times \Sigma^* \rightharpoonup U + Y$ by $\alpha^*(x, \epsilon) = \alpha_X(x)$ and for $w \in \Sigma^+$,

$$\alpha^*(x, w) = \begin{cases} \alpha^+(\alpha(x), w) & \text{if } \alpha_X(x) \in U \\ \text{undefined otherwise.} \end{cases}$$

This extension of α to Σ^* allows us to define the behaviour of a Σ-automaton. For each $x \in X$ and each $y \in Y$ we have a language over Σ that is the set of labels of paths under the action of α from x to y. Together we obtain an $X \times Y$ matrix of languages. More precisely, we have the following definition.

Definition 30. Let $(U, \alpha) : X \longrightarrow Y$ be a Σ-automaton. The *behaviour* of (U, α) is the $X \times Y$ matrix of Σ-languages $E(U, \alpha)_{x,y}$, where $E(U, \alpha)_{x,y} = \{w \in \Sigma^* | \alpha^*(x, w) = y\}$.

Notice that this definition can be interpreted as generalizing that of behaviour for an Elgot automaton. If we have an Elgot automaton $(U, \alpha) : X \longrightarrow Y$, we can define a Σ-automaton for $\Sigma_a = \{a\}$ as $\overline{\alpha} : X + (U \times \Sigma_a) \longrightarrow U + Y$, where $\overline{\alpha}_X(x) = \alpha_X(x)$ for $x \in X$ and $\overline{\alpha}_U(u, a) = \alpha_U(u)$. Then observe that $\overline{\alpha}^*(x, a^n) = \alpha^{n+1}(x)$ and both sides of the equation are either defined or undefined. Thus $E(U, \alpha)(x) = (y, n)$ iff $E(U, \overline{\alpha})_{x,y} = \{a^n\}$ (and $E(U, \alpha)(x)$ is undefined iff $E(U, \overline{\alpha})_{x,y} = \emptyset$ for all y).

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

We note some important properties of behaviours. First, since our automata are deterministic, for a fixed x the $E(U,\alpha)_{x,y}$ are pairwise disjoint. Moreover, if $w \in E(U,\alpha)_{x,y}$ and $v \in \Sigma^+$, then $wv \notin E(U,\alpha)_{x,y'}$ for any y' (including y). This motivates the following definitions.

Definition 31. A language $L \subseteq \Sigma^*$ is *anti-prefix* if for all $w, v \in \Sigma^*(w \in L$ and $wv \in L$ imply $v = \epsilon)$. Languages L_1 and L_2 are *anti-prefix-disjoint* if for all $w, v \in \Sigma^*(w \in L_1$ implies $wv \notin L_2)$, and *vice versa.*

Note that we may take $v = \epsilon$ in the second definition, so anti-prefix-disjoint languages are disjoint.

Proposition 32. Let $\mathbf{L}$ and $\mathbf{M}$ be $X \times Y$ and $Y \times Z$ matrices of languages such that the entries of $\mathbf{L}$ and $\mathbf{M}$ are anti-prefix languages and the entries in each row of $\mathbf{L}$ and $\mathbf{M}$ are pairwise anti-prefix-disjoint. The $X \times Z$ matrix $\mathbf{K} = \mathbf{LM}$ with entries defined by $K_{x,z} = \bigcup_{y \in Y} L_{x,y}M_{y,z}$ has anti-prefix entries and entries in each row are pairwise anti-prefix-disjoint.

Proof. We first show that the entries of $\mathbf{LM}$ are anti-prefix. Let $w \in LM_{x,z}$ so there are $y \in Y, w_1 \in L_{x,y}, w_2 \in M_{y,z}$ such that $w = w_1w_2$. Now suppose $wv \in LM_{x,z}$ so there are $y' \in Y, v_1 \in L_{x,y'}, v_2 \in M_{y',z}$ such that $w = v_1v_2$. We distinguish 3 cases:
Case 1: $|v_1| < |w_1|$. In this case $w_1 = v_1v_3$ for some v_3 with $|v_3| > 0$. If $y = y'$, this contradicts the prefix property of $L_{x,y}$. Otherwise, since entries in a row of $\mathbf{L}$ are pairwise anti-prefix-disjoint, $v_1 \in L_{x,y'}$ implies $w_1 = v_1v_3 \notin L_{x,y}$, which is a contradiction.
Case 2: $|v_1| = |w_1|$. In this case $v_1 = w_1$ so $y = y'$ since the entries in a row of $\mathbf{L}$ are anti-prefix disjoint. Then $w_2 \in M_{y,z}$ and $w_2v = v_2 \in M_{y,z}$ imply that $v = \epsilon$.
Case 3: $|v_1| > |w_1|$. In this case $v_1 = w_1v_3$ for some v_3 and a contradiction similar to Case 1 ensues.
We conclude that $v = \epsilon$.

Next we show that the entries in a row of $\mathbf{K}$ are pairwise anti-prefix disjoint. Let $w \in K_{x,z}$ so there are $y \in Y, w_1 \in L_{x,y}, w_2 \in M_{y,z}$ such that $w = w_1w_2$. Let $v \in \Sigma^*$. We must show that $wv \notin K_{x,z'}$ for $z' \neq z$. Again we have 3 cases:
Case 1: $|v_1| < |w_1|$. In this case $w_1 = v_1v_3$ for some v_3 with $|v_3| > 0$. As above this contradicts the properties of $\mathbf{L}$.
Case 2: $|v_1| = |w_1|$. In this case $v_1 = w_1$, so $y = y'$ since the row entries of $\mathbf{L}$ are anti-prefix disjoint. Then $w_2 \in M_{y,z}$ implies $w_2v = v_2 \notin M_{y,z'}$ since the row entries of $\mathbf{M}$ are anti-prefix-disjoint.
Case 3: $|v_1| > |w_1|$. Again this is similar to Case 1. We conclude that $wv \notin K_{x,z'}$. □

The preceding Proposition allows the definition of a suitable receiving category for the behaviours of Σ-automata. The category $\mathbf{B}_{\mathscr{F}}$ has objects sets, arrows from X to Y given by $X \times Y$ matrices of anti-prefix languages over Σ with entries in each row pairwise anti-prefix-disjoint. Composition is defined using the matrix multiplication of the preceding proposition. We define the matrix of the composite of $\mathbf{L} : X \longrightarrow Y$ and $\mathbf{M} : Y \longrightarrow Z$ to be the matrix $\mathbf{K} = \mathbf{LM}$, with the product taken in the diagrammatic order. Thus the composite is an arrow of $\mathbf{B}_{\mathscr{F}}$ by Proposition 32.

Lemma 33. If there is a 2-cell $\phi : (U,\alpha) \longrightarrow (U',\alpha')$ in $\mathscr{F}$, then $E(U,\alpha) = E(U',\alpha')$.

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

Viewing $\mathbf{B}_{\mathscr{F}}$ as a bicategory with discrete hom categories we get the following proposition.

Proposition 34. Behaviour, E, extends to homomorphism of bicategories from $\mathscr{F}$ to $\mathbf{B}_{\mathscr{F}}$.

Proof. First, E is locally functorial by the remarks after Definition 30 and Lemma 33. The behaviour of a serial composite of Σ-automata is the composite in $\mathbf{B}_{\mathscr{F}}$ of their behaviours. To see this, note that the concatenation of a Σ^* word from the first behaviour with one from the second simply describes a path through the composite automaton. □

Definition 35. Let $(U, \alpha) : X \longrightarrow Y$ be a Σ-automaton. The object of *reachable states* of (U, α) is

$$U_R = \{u \in U \mid \exists x \in X\, \exists w \in \Sigma^*\, \alpha^*(x, w) = u\}.$$

The *reachable kernel* of (U, α) is the automaton $R(U, \alpha) = (U_R, \alpha_R) : X \longrightarrow Y$, where $\alpha_R : X + U_R \longrightarrow U_R + Y$ is the restriction of α.

We again observe that $R : \mathscr{F}(X, Y) \longrightarrow \mathscr{F}(X, Y)$ is functorial and idempotent, and that there is a 2-cell $\rho_{(U,\alpha)} : R(U, \alpha) \longrightarrow (U, \alpha)$ that is the component of a natural transformation from R to $1_{\mathscr{F}(X,Y)}$. Moreover, $R\rho = \rho R$. We have the following analogues of results for Elgot automata.

Proposition 36.

(1) The functor R is an idempotent comonad on $\mathscr{F}(X, Y)$ with counit ρ.
(2) The behaviour of the reachable kernel, $R(U, \alpha)$, of a Σ-automaton (U, α) is the same as that of (U, α).
(3) If $(U, \alpha) : X \longrightarrow Y$ and $(V, \beta) : Y \longrightarrow Z$, there is a canonical 2-cell $r_{UV} : R((U, \alpha)(V, \beta)) \longrightarrow R(U, \alpha)R(V, \beta)$.

The coalgebras for the local reachability comonads are the reachable Σ-automata.

The minimization theory we obtain in the case of Σ- automata is also similar to that of the preceding section. We begin with an equivalence relation on states of (U, α): $u \sim_\alpha u'$ iff for all $w \in \Sigma^*$ and for all $y \in Y$ we have $\alpha_U^+(u, w) = y$ iff $\alpha_U^+(u', w) = y$.

Again, we can construct a 'quotient' automaton denoted $M(U, \alpha) = (U_M, \alpha_M)$. We define $U_M = U/\sim_\alpha$, and α_M is defined on $X + U_M$ by

$$\alpha_M(x) = \begin{cases} [u] & \text{if } \alpha(x) = u \\ y & \text{if } \alpha(x) = y \in Y \end{cases}$$

and for $a \in \Sigma$

$$\alpha_M([u], a) = \begin{cases} [u'] & \text{if } \alpha(u, a) = u' \in U \\ y & \text{if } \alpha(u, a) = y \in Y. \end{cases}$$

Proposition 37.

(1) The quotient function $\eta : U \longrightarrow U_M$ underlies a 2-cell in $\mathscr{F}$ denoted $\eta_{(U,\alpha)} : (U, \alpha) \longrightarrow M(U, \alpha)$. Applying M to it gives an isomorphism, and (M, η) is an idempotent monad on $\mathscr{F}(X, Y)$.
(2) The behaviour of $M(U, \alpha)$ is the same as that of (U, α).
(3) If $(U, \alpha) : X \longrightarrow Y$ and $(V, \beta) : Y \longrightarrow Z$ in $\mathscr{F}$, there is a canonical function $m_{UV} : M(U, \alpha)M(V, \beta) \longrightarrow M((U, \alpha)(V, \beta))$.

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

The algebras for M have a unique state associated with each path to an element of Y that actually occurs in the behaviour of (U, α).

Our next objective is a minimal realization of any behaviour in $\mathbf{B}_{\mathscr{F}}$. Let $\mathbf{L} = (L_{x,y})$ be a behaviour from X to Y. For $x, x' \in X$ and $w, w' \in \Sigma^*$, we write $(x, w) \sim_L (x', w')$ if and only if for all $v \in \Sigma^*$, for all $y \in Y$ $\;wv \in L_{x,y} \iff w'v \in L_{x',y'}$. The state set for the minimal realization automaton is $U_L = (X \times \Sigma^*)/\sim_L$. The action for the minimal realization is defined on $X + U_L$ by:

$$\alpha_L(x) = \begin{cases} y & \text{if } L_{x,y} = \{\epsilon\} \\ [(x, \epsilon)] & \text{otherwise} \end{cases}$$

and

$$\alpha_L([(x, w)], a) = \begin{cases} y & \text{if } wa \in L_{x,y} \\ [(x, wa)] & \text{otherwise.} \end{cases}$$

With the observation that α_L is well-defined, we have $\alpha_L : X + (U_L \times \Sigma) \longrightarrow U_L + Y$, and we note that the automaton $N\mathbf{L} = (U_L, \alpha_L)$ is reachable.

Proposition 38. The behaviour of $N\mathbf{L}$ is $\mathbf{L}$. Furthermore, we have E is left adjoint to $N : \mathbf{B}_{\mathscr{F}}(X, Y) \longrightarrow \mathscr{F}_R(X, Y)$.

Proof. The situation is similar to that for Elgot automata: the constructed automaton has states that correspond to equivalent deterministic transitions to an output state, plus possibly a loop state. Thus the first statement follows immediately.

For the adjunction, we show that 2-cells (in $\mathbf{B}_{\mathscr{F}}(X, Y)$) from $E(U, \alpha)$ to $\mathbf{L}$ correspond to 2-cells (in $\mathscr{F}_R(X, Y)$) from (U, α) to $N\mathbf{L}$. Since $\mathbf{B}_{\mathscr{F}}$ is locally discrete, this means we need to show that $E(U, \alpha) = \mathbf{L}$ if and only if there is a unique 2-cell from (U, α) to $N\mathbf{L}$.

First observe that if there is $(U, \alpha) \longrightarrow N\mathbf{L}$, then $E(U, \alpha) = EN\mathbf{L} = \mathbf{L}$, by Lemma 33 and the first paragraph.

For necessity, we suppose $E(U, \alpha) = \mathbf{L}$ and seek to define a unique 2-cell $\phi : (U, \alpha) \longrightarrow N\mathbf{L}$. Recalling that (U, α) is reachable, we define $\phi : U \longrightarrow U_L$ by

$$\phi(u) = [(x, w)] \qquad \text{if } \alpha^*(x, w) = u.$$

We need to show that ϕ is well-defined, that it defines a 2-cell in $\mathscr{F}_R$, and that it is the only such 2-cell. The first two follow immediately from $E(U, \alpha) = \mathbf{L}$. For the last, simply observe that $[(x, w)]$ satisfies $\alpha_L^*(x, w) = [(x, w)]$, so $\phi(\alpha^*(x, w)) = \alpha_L^*(x, w) = [(x, w)]$, and the requirement that ϕ be a morphism determines the definition of $\phi(u)$. $\square$

5. Lax monads

We first recall some definitions for bicategory morphisms and lax monads. In particular, we consider morphisms of bicategories that are identity on objects and that have the structure of a monad on each hom category, and then we give conditions sufficient to guarantee that the hom-category monads define a monoid in a suitable category of bicategory morphisms.

To establish notation, we recall that a *morphism of bicategories* from $\mathscr{B}$ to $\mathscr{C}$ is a pair (F, ϕ) in which: F maps objects and 1-cells of $\mathscr{B}$ to objects and 1-cells of $\mathscr{C}$; for every

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

object B of $\mathscr{B}$, there is a 2-cell $\phi_B : 1_{FB} \longrightarrow F(1_B)$; and whenever $f : B \longrightarrow B'$ and $g : B' \longrightarrow B''$ are composable, there is a 2-cell $\phi_{gf} : FgFf \longrightarrow Fgf : B \longrightarrow B''$. The data are subject to equations found in Bénabou (1967). We denote the action of F on a hom category by $F(B,B') : \mathscr{B}(B,B') \longrightarrow \mathscr{C}(FB,FB')$. We will also need to consider oplax transformations between morphisms. An oplax transformation $\sigma : (F,\phi) \longrightarrow (G,\gamma)$ is given by arrows $\sigma_B : FB \longrightarrow GB$ for all objects B in $\mathscr{B}$, and 2-cells $\sigma_f : \sigma_{B'}Ff \longrightarrow Gf\sigma_B$, whenever $f : B \longrightarrow B'$ is in $\mathscr{B}$, subject to equations again in Bénabou (1967). Our interest, as noted above, will be in rather special morphisms and transformations. They arise in examples and ensure that we obtain a monoidal category in which to define lax monads.

Proposition 39. (Carboni and Rosebrugh 1991, Proposition 2.1) For any class X the following data determine a bicategory, which we denote $\mathscr{M}(X)$:

1 Objects are bicategories with class of objects X.
2 One-cells are morphisms of bicategories that are identity on objects.
3 Two-cells are oplax transformations whose object components are all identities.

Definition 40. (Carboni and Rosebrugh 1991) A *lax monad* on $\mathscr{B}$ with objects $\mathbf{B}_o$ is a monoid in $\mathscr{M}(\mathbf{B}_o)(\mathscr{B},\mathscr{B})$.

We can give explicit criteria of a more elementary sort providing a characterization of morphisms that are lax monads.

Proposition 41. (Carboni and Rosebrugh 1991, Proposition 2.3) An endomorphism (T, τ) of $\mathscr{B}$ in $\mathscr{M}(\mathbf{B}_o)$ together with, for every pair B, B' in $\mathscr{B}$, natural transformations

$$\eta_{BB'} : \mathbf{1}_{\mathscr{B}}(B,B') \longrightarrow T(B,B') \longleftarrow T^2(B,B') : \mu_{BB'}$$

extends to a lax monad if

1 each $(T(B,B'), \eta_{BB'}, \mu_{BB'})$ is a monad on $\mathscr{B}(B,B')$
2 for all B,

$$\tau_B = \eta_{\mathbf{1}_B}$$

3 if $f : B \longrightarrow B'$ and $g : B' \longrightarrow B''$ are 1-cells in $\mathscr{B}$, then

$$\tau_{gf}(\eta_g \circ \eta_f) = \eta_{gf} : gf \longrightarrow Tgf$$

and

$$\mu_{gf} T\tau_{gf}\tau_{TgTf} = \tau_{gf}(\mu_g \circ \mu_f) : T^2gT^2f \longrightarrow Tgf.$$

Conversely, a lax monad determines transformations $\eta_{BB'}$ and $\mu_{BB'}$ satisfying 1, 2 and 3.

In fact the lax monads and colax comonads considered below are (locally) *idempotent*, that is, for any arrow $f : B \longrightarrow B'$ in $\mathscr{B}$ we have $T\eta_f = \eta_{Tf}$, and the common value is inverted by μ_f, so that (locally) $T^2 \cong T$. In this case we have a simplification given by the following proposition.

Proposition 42. If $((T,\tau),\eta,\mu)$ is idempotent, the equation involving μ in Part 3 of the preceding proposition follows from the other data.

Proof. To show that $\mu_{gf}T\tau_{gf}\tau_{TgTf} = \tau_{gf}(\mu_g \circ \mu_f)$, we show that $(\mu_{gf})^{-1}\tau_{gf} = T\tau_{gf}\tau_{TgTf}(\mu_g \circ \mu_f)^{-1}$. Now $(\mu_{gf})^{-1} = \eta_{Tgf}$ and $(\mu_g \circ \mu_f)^{-1} = \mu_g^{-1} \circ \mu_f^{-1} = \eta_{Tg} \circ \eta_{Tf}$. Note that the following diagram commutes by naturality of η and the equation above

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

involving η.

$$\begin{tikzcd}
 & TgTf \arrow[r, "\tau_{gf}"] \arrow[dl, "\eta_{Tg}\circ\eta_{Tf}"'] \arrow[d, "\eta_{TgTf}"] & Tgf \arrow[d, "\eta_{Tgf}"] \\
T^2gT^2f \arrow[r, "\tau_{TgTf}"'] & T(TgTf) \arrow[r, "T\tau_{gf}"'] & T^2gf
\end{tikzcd}$$

Thus $(\mu_{gf})^{-1}\tau_{gf} = \eta_{Tgf}\tau_{gf} = T(\tau_{gf})\eta_{TgTf} = \tau_{TgTf}(\eta_{Tg}\circ\eta_{Tf}) = T\tau_{gf}\tau_{TgTf}(\mu_g\circ\mu_f)^{-1}$. $\square$

In view of Condition 1 of Proposition 41, there is a local category of (Eilenberg–Moore) algebras for each pair of objects. The main result of Carboni and Rosebrugh (1991) constructs a bicategory with these algebras as hom categories assuming local exactness conditions on the underlying morphism (T,τ). This construction of algebras simplifies the case when the monad is idempotent. In fact, no exactness is required of the local monads in this case.

Proposition 43. Let $((T,\tau),\eta,\mu)$ be an idempotent lax monad on a bicategory, $\mathscr{B}$. The following data determine a bicategory denoted $\mathscr{B}^T$:

1 objects are those of $\mathscr{B}$
2 for objects B and B' of $\mathscr{B}$, the hom category is $\mathscr{B}^T(B,B')$
3 composition of 1-cells $f : B \longrightarrow B', g : B' \longrightarrow B''$ is defined by $T(gf)$
4 horizontal composition of 2-cells is also defined by application of T.

Proof. In Carboni and Rosebrugh (1991) it is shown that the underlying arrow of the composite of (f,ϕ) in $\mathscr{B}^T(B,B')$ with (g,γ) in $\mathscr{B}^T(B',B'')$ is the joint coequalizer of two parallel pairs of 2-cells, one of which is $T(g\circ\phi) : T(gTf) \longrightarrow T(gf)$ and $\mu_{gf}T\tau_{gf}T(\eta_g\circ Tf) : T(gTf) \longrightarrow T(gf)$ (and the other pair just interchanges the roles of ϕ and γ). We claim that these 2-cells are equal. Now assuming that the monad is idempotent means that $\mu_{gf} = (T\eta_{gf})^{-1}$, and since algebras for $T(B,B')$ are the objects (of $T(B,B')$) for which the unit is invertible, we have $\phi = \eta_f^{-1}$. Thus our claim holds if $T\eta_{gf}T(g\circ\eta_f^{-1}) = T\tau_{gf}T(\eta_g\circ Tf)$, so if $\eta_{gf}g\circ\eta_f^{-1} = \tau_{gf}\eta_g\circ Tf$. But from the first of Equations 3 in Proposition 41, $\eta_{gf} = \tau_{gf}(\eta_g\circ\eta_f)$, and thus $\eta_{gf}g\circ\eta_f^{-1} = \tau_{gf}(\eta_g\circ\eta_f)g\circ\eta_f^{-1} = \tau_{gf}\eta_g\circ Tf$, as required. The other pair of 2-cells is similarly equal, so the required joint coequalizer is just $T(gf)$, which is then the composite of (f,ϕ) and (g,γ) as claimed. The horizontal composite is similar. $\square$

In the next section we will need to consider a dual of the concepts described above, namely colax comonads. A *comorphism of bicategories* is $(G,\gamma) : \mathscr{B} \longrightarrow \mathscr{C}$ where G maps objects and 1-cells of $\mathscr{B}$ to objects and 1-cells of $\mathscr{C}$. For every object B of $\mathscr{B}$ there is a 2-cell $\gamma_B : G(1_B) \longrightarrow 1_{GB}$; and whenever $f : B \longrightarrow B'$ and $g : B' \longrightarrow B''$ are composable, there is a 2-cell $\gamma_{gf} : Ggf \longrightarrow GgGf : B \longrightarrow B''$, subject to appropriate equations. An opcolax transformation $\sigma : (G,\gamma) \longrightarrow (H,\nu)$ between comorphisms is given by arrows

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

$\sigma_B : GB \longrightarrow HB$ for all objects B in $\mathscr{B}$, and 2-cells $\sigma_f : Hf\sigma_B \longrightarrow \sigma_{B'}Gf$,whenever $f : B \longrightarrow B'$ is in $\mathscr{B}$, again subject to equations. As above, we obtain a bicategory $\mathscr{C}(\mathscr{X})$ of identity on objects comorphisms and define a *colax comonad* on $\mathscr{B}$ to be a comonoid in $\mathscr{C}(\mathbf{B}_l)(\mathscr{B},\mathscr{B})$. We will not state the obvious duals of propositions in this section, but we will use them without further comment in the next section.

6. Applications to automata

In Sections 2, 3 and 4 we have identified various local (co-)monads for reachability and minimization. Our purpose in this section is to apply the results given in the preceding section to demonstrate that these local (co-)monads extend to lax (co-)monads defined on the bicategories of automata concerned. That is, they are compatible with serial composition up to a comparison morphism. We show further that the Nerode adjunctions described above also extend to the (bi-)categories in question.

We begin with Mealy automata, considering reachability first and then minimization. The notation is from Section 2. In the case of reachability we deal with identity on objects endocomorphisms.

Proposition 44. The comonads $R(X,Y)$ defined on $\mathscr{A}$ extend to an idempotent colax comonad $R : \mathscr{A} \longrightarrow \mathscr{A}$, and the algebras for $R(X,Y)$ are the one-cells of a bicategory, denoted $\mathscr{A}_R$ of reachable automata.

Proof. We first need to show that the local functors $R(X,Y)$ have the structure of a comorphism (R,r) on $\mathscr{M}$. Recall that the identity Mealy automaton 1_X on X is (essentially the identity arrow) $(1,t,1_1) : X \longrightarrow X$ where $t : X \times 1 \longrightarrow 1 \times X$. Since $RX = X$, and since we easily see $R(1_X) = 1_X$, we simply take $r_X : R(1_X) \longrightarrow 1_{RX}$ to be the identity. Let $(U,\alpha,u_0) : X \longrightarrow Y$ and $(V,\beta,v_0) : Y \longrightarrow Z$. The comparison 2-cell for their composite is r_{UV} from Proposition 11.

Since the $R(X,Y)$ are idempotent comonads, by Proposition 42 we have to check only Equations 2 and the first of Equations 3 in Proposition 41 to see that the $R(X,Y)$ extend. Both r_X and ρ_{1_X} are identities, so Equations 2 are satisfied. For the first of Equations 3 we note that

$$(\rho_V \circ \rho_U)r_{UV} : (U \times V)_R \longrightarrow U_R \times V_R \longrightarrow U \times V$$

is simply the inclusion $\rho_{UV} : (U \times V)_R \longrightarrow U \times V$. □

The colax structure provides a comparison between the reachable kernel of a serial composite (in $\mathscr{A}$) and the serial composite of reachable kernels. Serial composition of reachable automata is just composition of 1-cells in the bicategory of coalgebras. The explicit description of composition in $\mathscr{A}_R$ is simply that the composite of reachable automata in $\mathscr{A}_R$ is the reachable kernel of their composite in $\mathscr{A}$. Comments of the same sort apply to the colax comonads for reachability and lax monads for minimization described below.

The next result follows immediately from Carboni and Rosebrugh (1991, Theorem 3.6), and the preceding Proposition.

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

Corollary 45. The idempotent colax comonad $(R, \rho) : \mathscr{A} \longrightarrow \mathscr{A}$ factors as

$$\mathscr{A} \xrightarrow{G} \mathscr{A}_R \xrightarrow{J} \mathscr{A},$$

where G is a bicategory homomorphism. For all X, Y we have $J(X, Y) \dashv G(X, Y)$, so $\mathscr{A}_R(X, Y)$ is a coreflective subcategory of $\mathscr{A}(X, Y)$ with coreflector G

We now turn to the similar results for minimization of Mealy automata.

Proposition 46. The monads $M(X, Y) : \mathscr{A}(X, Y) \longrightarrow \mathscr{A}(X, Y)$ extend to a lax monad on $\mathscr{A}$. Algebras for $M(X, Y)$ are the one-cells of a bicategory, denoted $\mathscr{A}^M$ of minimized automata.

Proof. Again we first show that the local functors $M(X, Y)$ have the structure of a morphism (M, m) on $\mathscr{M}$. The identity Mealy automaton 1_X on X is (essentially the identity arrow) $(1, t, 1_1) : X \longrightarrow X$ where $t : X \times 1 \longrightarrow 1 \times X$. Thus $M(1_X) = 1_X$, and since $MX = X$, we take $m_X : 1_{MX} \longrightarrow M(1_X)$ to be the identity. The comparison 2-cell for a composite $(V, \beta, v_0)(U, \alpha, u_0) : X \longrightarrow Z$ is m_{UV} from Proposition 14.

Since the $M(X, Y)$ are idempotent monads, by Proposition 42 we have to check only Equations 2 and the first of Equations 3 in Proposition 41 to see that the $M(X, Y)$ extend. Both m_X and μ_{1_X} are identities, so Equations 2 are satisfied. For the first of Equations 3 we note that $m_{UV}(\mu_V \circ \mu_U)$ is simply the quotient $\mu_{UV} : U \times V \longrightarrow (U \times V)_M$. □

Composition in $\mathscr{A}^M$ is easy to describe: the composite of minimized automata in $\mathscr{A}^M$ is the minimization of their composite in $\mathscr{A}$.

Corollary 47. The idempotent lax monad $(M, \eta) : \mathscr{A} \longrightarrow \mathscr{A}$ factors as

$$\mathscr{A} \xrightarrow{F} \mathscr{A}^M \xrightarrow{I} \mathscr{A},$$

where F is a bicategory homomorphism, for all X, Y we have $F(X, Y) \dashv I(X, Y)$, so $\mathscr{A}^M(X, Y)$ is a reflective subcategory of $\mathscr{A}(X, Y)$ with reflector F.

The situation we have been describing is summed up in the following proposition. The pairs of functors are locally adjoint and provide examples of the various notions of local adjunction in the literature.

Proposition 48. In the following diagram both the inner and outer squares commute. The I's (respectively, J's) are locally reflective (respectively, coflective) inclusions.

$$\begin{tikzcd}
\mathscr{A} \arrow[rr, shift left, "G"] \arrow[d, shift right, "F"'] & & \mathscr{A}_R \arrow[ll, shift left, "J"] \arrow[d, shift left, "F'"] \\
\mathscr{A}^M \arrow[u, shift right, "I"'] \arrow[r, shift right, "G'"'] & (\mathscr{A}^M)_{R'} \arrow[l, shift right, "J'"'] \arrow[r, phantom, "\cong"] & (\mathscr{A}_R)^{M'} \arrow[u, shift left, "I'"]
\end{tikzcd}$$

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

Theorem 49. The Nerode automaton construction extends to a morphism of bicategories, $N : \mathbf{B}_{\mathscr{A}} \longrightarrow \mathscr{A}_R$, and E and N determine a local adjunction. Moreover, N factors as $N = I'N'$, E factors as $E = E'F'$, and N' and E' determine an equivalence $\mathbf{B}_{\mathscr{A}} \cong (\mathscr{A}_R)^{M'}$ as indicated in the diagram.

$$\begin{tikzcd}
& \mathscr{A}_R \arrow[dl, shift right, "F'"'] \arrow[dr, shift left, "E"] & \\
(\mathscr{A}_R)^{M'} \arrow[ur, shift right, "I'"'] \arrow[rr, shift left, "E'"] \arrow[rr, phantom, "\simeq"] & & \mathbf{B}_{\mathscr{A}} \arrow[ul, shift left, "N"'] \arrow[ll, shift left, "N'"]
\end{tikzcd}$$

Proof. Suppose that $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are composable behaviours. We need a 2-cell $v_{gf} : NgNf \longrightarrow Ngf$ in $\mathscr{A}_R$. Note that the composite $NgNf$ is in $\mathscr{A}_R$, so is that described in Corollary 44. Recall that the composite in $\mathscr{A}$ of Nf and Ng has internal states $X^*/\sim_f \times Y^*/\sim_g$, that is, pairs $([w]_f, [v]_g)$ where $w \in X^*, v \in Y^*$. An easy calculation shows that the reachable states are pairs of the form $([w]_f, [f(w)]_g)$. After this observation, it is easy to see that defining $v([w]_f, [f(w)]_g) = v([w]_{gf})$ provides the required structure.

We also need $v_X : 1_X \longrightarrow N1_X$, but this 2-cell can be taken to be an identity since $N1_X$ has only one internal state (the equivalence relation $\sim_{1_X}$ is the all relation).

In the diagram above, we can define E' to be EI' and N' to be $F'N$. To establish the theorem, we verify that these provide factorizations of E and N as $E \cong E'F'$ and $N \cong I'N'$, and then show that both composites of E' and N' are isomorphic to the identity.

First, $E'F' = EI'F'$ by definition. Since F' is minimization, applying it does not affect behaviour, so we have $EI'F' = E$, and the first iso is established. Next, $I'N' = I'F'N$ by definition. Now Nf is a minimized automaton, so application of $I'F'$ is essentially the identity and the second iso follows.

For the equivalence, note that $E'N' = EI'F'N$ by definition, and as just observed, $I'F'N \cong N$, so $E'N' \cong EN$, but $ENf = f$ by the previous Theorem. Finally, $N'E' = F'NEI'$ by definition. Now a reachable, minimal Mealy automaton $I'A$ is isomorphic to the Nerode automaton of its behaviour, that is, $NEI'A \cong I'A$. (The unit of the adjunction in the previous Theorem provides the comparison, which is epic by reachability and monic by minimality.) Thus $N'E'A = F'NEI'A \cong F'I'A \cong A$. □

Using notation from Section 4, we consider the situation for Σ-automata. Recall that we can view the Elgot automata of Section 3 as a special case.

Proposition 50. The monads $R(X, Y)$ defined on $\mathscr{F}$ extend to an idempotent colax comonad $R : \mathscr{F} \longrightarrow \mathscr{F}$, and the algebras for $R(X, Y)$ are the one-cells of a bicategory, denoted $\mathscr{F}_R$ of reachable Σ-automata.

Proof. Recall first from Proposition 36 that if $(U, \alpha) : X \longrightarrow Y$ and $(V, \beta) : Y \longrightarrow Z$,

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

there is a comparison 2-cell $r_{UV} : R((U,\alpha)(V,\beta)) \longrightarrow R(U,\alpha)R(V,\beta)$. The equations for a colax comonad are trivial in this situation so the result follows from Proposition 43. □

The composite of reachable automata in $\mathscr{F}_R$ is the reachable coreflection of their composite in $\mathscr{F}$.

Corollary 51. The idempotent colax comonad $(R,\rho) : \mathscr{F} \longrightarrow \mathscr{F}$ factors as

$$\mathscr{F} \xrightarrow{G} \mathscr{F}_R \xrightarrow{J} \mathscr{F},$$

where G is a bicategory homomorphism. For all X, Y we have $J(X,Y) \dashv G(X,Y)$, so $\mathscr{F}_R(X,Y)$ is a coreflective subcategory of $\mathscr{F}(X,Y)$ with coreflector G.

Once again, the minimal realization-behaviour adjunction is essentially the minimization local adjunction. We begin with minimization.

Proposition 52. The monads $M(X,Y) : \mathscr{F}(X,Y) \longrightarrow \mathscr{F}(X,Y)$ extend to a lax monad on $\mathscr{F}$. Algebras for $M(X,Y)$ are the one-cells of a bicategory, denoted $\mathscr{F}^M$ of minimized Σ-automata.

Proof. Recall first from Proposition 37 that if $(U,\alpha) : X \longrightarrow Y$ and $(V,\beta) : Y \longrightarrow Z$, there is a comparison 2-cell $m_{UV} : M(U,\alpha)M(V,\beta) \longrightarrow M((U,\alpha)(V,\beta))$. The equations for a lax monad also follow easily, so the result follows from Proposition 43. □

The composite of minimized Σ-automata in $\mathscr{F}^M$ is the minimization of their composite in $\mathscr{F}$.

Corollary 53. The idempotent lax monad $(M,\eta) : \mathscr{F} \longrightarrow \mathscr{F}$ factors as

$$\mathscr{F} \xrightarrow{F} \mathscr{F}^M \xrightarrow{I} \mathscr{F},$$

where F is a bicategory homomorphism. For all X, Y we have $F(X,Y) \dashv I(X,Y)$, so $\mathscr{F}^M(X,Y)$ is a reflective subcategory of $\mathscr{F}(X,Y)$ with reflector F.

As with Mealy automata, it is easy to see that reachable kernel and minimization commute. Thus we obtain lax monads for reachability (respectively, minimization) on $\mathscr{F}^M$ (respectively, $\mathscr{F}_R$) whose local algebras coincide, that is, $(\mathscr{F}^M)_{R'} \cong (\mathscr{F}_R)^{M'}$, where the primed monads act as R (respectively, M) did on $\mathscr{F}$. We sum up with the following proposition.

Proposition 54. In the following diagram both the inner and outer squares commute. The I's (respectively, J's) are locally reflective (respectively, coflective) inclusions.

$$\begin{tikzcd}
\mathscr{F} \arrow[rr, shift left, "G"] \arrow[d, shift right, "F"'] & & \mathscr{F}_R \arrow[ll, shift left, "J"] \arrow[d, shift left, "F'"] \\
\mathscr{F}^M \arrow[u, shift right, "I"'] \arrow[r, shift right, "G'"'] & (\mathscr{F}^M)_{R'} \arrow[l, shift right, "J'"'] \arrow[r, phantom, "\cong"] & (\mathscr{F}_R)^{M'} \arrow[u, shift left, "I'"]
\end{tikzcd}$$

https://doi.org/10.1017/S0960129597002454 Published online by Cambridge University Press

Finally, the minimized reachable Σ-automata have the same relation to their behaviours as Mealy automata.

Theorem 55. The minimal realization construction on $\mathbf{B}_{\mathscr{F}}$ extends to a morphism of bicategories, $N : \mathbf{B}_{\mathscr{F}} \longrightarrow \mathscr{F}_R$, and E and N determine a local adjunction. Moreover, N factors as $N = I'N'$, E factors as $E = E'F'$, and N' and E' determine an equivalence $\mathbf{B}_{\mathscr{F}} \cong (\mathscr{F}_R)^{M'}$ as indicated in the diagram.

$$\begin{tikzcd}
& \mathscr{F}_R \arrow[dl, shift right, "F'"'] \arrow[dr, shift left, "E"'] & \\
(\mathscr{F}_R)^{M'} \arrow[ur, shift right, "I'"'] \arrow[rr, shift left, "E'"] \arrow[rr, phantom, "\simeq"] & & \mathbf{B}_{\mathscr{F}} \arrow[ul, shift left, "N"'] \arrow[ll, shift left, "N'"]
\end{tikzcd}$$

Proof. The proof is very similar to the proof in the case of Mealy automata once we make the observation that N is actually a lax functor in this situation also. $\square$

Acknowledgement

The authors wish to acknowledge discussions with Stephen Bloom.

References

Adamek, J. and Trnkova, V. (1990) *Automata and Algebras in Categories*, Kluwer.

Bénabou, J. (1967) Introduction to bicategories. *Springer-Verlag Lecture Notes in Mathematics* **47** 1–77.

Betti, R. and Power, J. (1988) On local adjunctions of distributive bicategories. *Bolletino della Unione Matematica Italiana* **7** (2-B) 931–947.

Blum, L., Shub, M. and Smale, S. (1989) On a theory of computation and complexity over the real numbers: NP completeness, recursive functions and universal machines. *Bulletin of the American Mathematical Society* **21** 1–46.

Bloom, S., Sabadini, N. and Walters, R. F. C. (1996) Matrices, machines and behaviors. *Applied Categorical Structures* **4** 343–360.

Carboni, A. and Rosebrugh, R. (1991) Lax monads. *Journal of Pure and Applied Algebra* **76** 13–32.

Ginsburg, S. (1966) *The Mathematical Theory of Context-Free Languages*, McGraw-Hill.

Goguen, J. A. (1972) Minimal realization of machines in closed categories. *Bulletin of the AMS* **78** 777–783.

Hopcroft, J. and Ullman, J. D. (1979) *Introduction to automata theory, languages and computation*, Addison-Wesley.

Jay, B. (1988) Local adjunctions. *Journal of Pure and Applied Algebra* **53** 227–238.

Katis, P., Sabadini, N. and Walters, R. F. C. (1997) Bicategories of processes. *Journal of Pure and Applied Algebra.* **115** (2) 141–178.

Nerode, A. (1958) Linear automaton transformations. *Proc. Amer. Math. Soc.* **9** 541–544.

Sabadini, N., Walters, R. F. C. and Vigna, S. (1996) A note on recursive functions. *Mathematical structures in computer science* **6** 127–139.

Sabadini, N., Weld, H. and Walters, R. F. C. (1994) On categories of asynchronous circuits. *School of mathematics and statistics research reports, University of Sydney* (93-34).

Vigna, S. (1996) On relations between distributed computability and the BSS model. *Theoretical Computer Science* **162** 5–21.