Technical Report

UCAM-CL-TR-523
ISSN 1476-2986

Number 523

UNIVERSITY OF CAMBRIDGE

Computer Laboratory

Bigraphical reactive systems: basic theory

Robin Milner

September 2001

JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500

http://www.cl.cam.ac.uk/

Technical reports published by the University of Cambridge
Computer Laboratory are freely available via the Internet:

http://www.cl.cam.ac.uk/TechReports/

Series editor: Markus Kuhn

ISSN 1476-2986

Bigraphical reactive systems: basic theory

Robin Milner

University of Cambridge Computer Laboratory
William Gates Building, J J Thomson Avenue, Cambridge CB3 0FD, UK

September 2001

Abstract A notion of *bigraph* is proposed as the basis for a model of mobile interaction. A bigraph consists of two independent structures: a *topograph* representing locality and a *monograph* representing connectivity. Bigraphs are equipped with reaction rules to form *bigraphical reactive systems (BRSs)*, which include versions of the π-calculus and the ambient calculus. Bigraphs are shown to be a special case of a more abstract notion, *wide reactive systems (WRSs)*, not assuming any particular graphical or other structure but equipped with a notion of *width*, which expresses that agents, contexts and reactions may all be widely distributed entities.

A behavioural theory is established for WRSs using the categorical notion of *relative pushout*; it allows labelled transition systems to be derived uniformly, in such a way that familiar behavioural preorders and equivalences, in particular bisimilarity, are congruential under certain conditions. Then the theory of bigraphs is developed, and they are shown to meet these conditions. It is shown that, using certain functors, other WRSs which meet the conditions may also be derived; these may, for example, be forms of BRS with additional structure.

Simple examples of bigraphical systems are discussed; the theory is developed in a number of ways in preparation for deeper application studies.

CONTENTS

APPENDICES

1 Introduction

It is nearly forty years since Petri devised the first substantial model of concurrent computation, and it was a graphical model [26]. Since that time a great many models have been studied. They are not always graphical, but the spatial metaphor is never far away; we often use terms like linkage, location, mobility, and so on. As it was for Petri, it remains a challenge for us to deploy spatial intuition but to retain rigour. This challenge grows as mobility grows in importance.

Ten years ago I began the study of *action calculi* [22], an algebraic theory with a graphical interpretation, in which the dynamics of a system consists of the reconfiguration of its graph. Action calculi –of which we assume no knowledge here– provide a common frame for a variety of concurrency models, including Petri nets and the π-calculus; each model can be specified in the frame by a signature (a set of node types) and a set of reaction rules. Such a general framework yields understanding of the family of models, and of their differences. What it must also do is to provide a central theory which can be specialised to each individual model. I now believe that such a theory is best achieved by treating graphs as the primary mathematical objects, not only as a means to visualise an algebraic or other theory. That belief lies behind the present work, which advances from action calculi in two ways: in graphical structure, and in behavioural theory.

Graphical structure The first advance is in the form of graph on which the model is based. The model introduced here treats the *locality* and *connectivity* of agents –their nesting and wiring– as orthogonal, co-ordinated only by the nodes of the graph. Hence the term *bigraph*; each graph is the superposition of two graphical structures.

The model inherits many features from action calculi. In particular, a node of a graph represents a variety of things, for example a π-calculus input term $x(y).P$, a 'solution' of the Chemical Abstract Machine [1], an administrative region, or a host machine running several processes. A node possesses ports, linked by arcs to other ports; a node may also enclose, nested within it, other nodes similarly linked. However, in action calculi there are constraints both on the arcs themselves (arcs are directed, and an output port carries exactly one outgoing arc) and on the relation between arcs and nesting (a node may be linked to siblings, parents, aunts and uncles but not to cousins in the nesting hierarchy). These constraints in action calculi made sense from the algebraic viewpoint adopted there. In bigraphs the constraints are removed; thereby we not only model a broader class of systems but also achieve a tractable behavioural theory. Several influences have led to this treatment:

(1) The mobile ambients of Cardelli and Gordon [2] have emphasized the value of dealing with mobility in terms of (physical) location, rather than coding it in terms of wiring (as can be done to some extent in the π-calculus). In terms of node-nesting, action calculi appear to be able to model ambients as originally defined; but they fail to model certain natural developments of the ambient notion.

(2) The fusion systems of Gardner [10], which evolved from action calculi, represent a process-calculus framework whose graphical form is implicit in its process-algebraic formulation. Two of its innovations are adopted in the present work; the unconstrained connectivity of the kind mentioned above, and the explicit fusions –here called aliases and coaliases– of Gardner and Wischik [11], developed from the fusion calculus of Parrow and Victor [25]. These authors are further developing a calculus of fusion graphs.

(3) It may be argued that to allow arcs to link nodes which are distant cousins, i.e. enclosed within distinct parent nodes arbitrarily far apart in the nesting structure, is contrary to reality. But we wish to model not only the *reality*, e.g. how communication is implemented, but also the *fiction* –e.g. 'action-at-a-distance'– which the world wide web permits us to adopt. By embracing both views in the same model, one can hope to validate complex communication protocols in a mobile environment. This is well-argued by Wojciechowski and Sewell in their Nomadic Pict [29]; Sewell and Unyapoth [28] develop verification techniques for 'infrastructure algorithms', which implement high-level location-independent communication primitives in terms of low-level location-dependent ones.

(4) A strong motivation has been to model fully-fledged mobile interaction, as increasingly found on the internet. A certain complexity is essential in a model which will support the analysis of such real systems. By treating locality and connectivity as independently as possible, it appears that the model can nevertheless remain mathematically tractable.

Behavioural theory In process calculi, activity is often first expressed as *reactions* of the form $a \longrightarrow a'$, where a and a' are agents, and then refined or extended somehow to *labelled transitions* of the form $a \xrightarrow{\ell} a'$, where the label ℓ is drawn from some vocabulary expressing the interactions which an agent may perform with its environment. Transitions (we henceforth omit 'labelled') have the advantage that they support the definition of behavioural preorders and equivalences, such as traces, failures and bisimilarity. This kind of behavioural theory has been successful; but for each calculus one may ask where the labels come from. Typically they are simple, but not always atomic; in the π-calculus we have transitions like $a \xrightarrow{\nu x.\overline{y}\langle x\rangle} a'$, meaning that a can send a new private name x along the channel y. For action calculi, but more generally for any *reactive system* (under a precise definition of this concept), we may ask whether somehow these labels can be *derived* uniformly from any given set of reaction rules $r \longrightarrow r'$. A natural approach is to take the labels to be (a special class of) contexts; then $a \xrightarrow{F} a'$ implies the reaction $F \circ a \longrightarrow a'$, i.e. in the context F, a can react to become a'. But we don't want *all* contexts as labels. How to find a suitably minimal set remained open for some six years. Sewell [27] was able uniformly to derive satisfactory context-labelled transitions for parametric term rewrite systems with parallel composition and blocking, and showed bisimilarity to be a congruence. It remained a problem to do it for reactive systems dealing with connectivity. In recent work we have achieved a uniform derivation for all reactive systems satisfying certain conditions. We arrive at quite tractable transition systems for which, moreover, bisimilarity and other familiar behavioural preorders and equivalences are *guaranteed* to be congruences. The first results of this kind appear in Leifer and Milner [19]; these were extended and refined in Leifer's PhD Dissertation [18], which is the most comprehensive reference.

The notion of *reactive system* is defined categorically, and one of the required conditions is that, in the appropriate precategory, sufficient *relative pushouts* (RPOs) exist. In a pilot study, Cattani *et al* [4], we showed that this condition holds for certain action calculi involving no nesting of nodes. We believe that the result extends to calculi with nesting, but verifying the condition becomes uncomfortably hard. The pilot study led us to expect our RPO condition to be met in interesting cases, but also led us to believe that it would be hard to verify for nested systems in which locality and connectivity are interdependent.

We come now to the second advance represented by bigraphs: their behavioural theory –being simpler– has been further developed. In fact, as reported here, the RPO condition holds for a wide class of bigraphical reactive systems (BRSs). The paper includes an example of a derived transition system for a simple BRS; it also develops techniques –based upon certain functors– by which the RPO condition can be transferred to other models than BRSs.

Discussion The nature of this report is dictated by the need, mentioned in the second paragraph of the introduction, for any general model for interactive systems to pay its way by providing a central theory that can be applied usefully in specific instances – i.e. in calculi with special orientation. So the aim here is not to polish each stage of development, but rather to reach a point at which applications of the theory can be carried out and assessed. This point is represented, in Section 6, by a preliminary taxonomy of derived transition rules followed by a conjectured definition of *parametric reaction rule*. This will allow us, as a first experiment, to calibrate the theory against that developed over the last decade for the π-calculus; this is a natural choice for a test case, since many calculi share features with the π-calculus.

We call the proposed definition conjectural, because bigraph theory cannot simply assert the way to define reaction; its proposal needs to be assessed against what is found useful and natural in practice. There are several ways in which the proposal can be judged: Does it lead to transition systems which relate nicely to those that we know? Do the resulting behavioural relations match those that we know, or are they new? And does the theory explain phenomena or resolve questions in specific calculi – e.g. why the original strong bisimilarity for the π-calculus was not a congruence? (This last example is not chosen randomly; I believe that it yields a very satisfactory explanation.)

Because the primary concern has been to reach a point at which the theory can be tested, many questions about the theoretical path chosen are not fully answered here. For example: is it best to use precategories, as we do, instead of categories? Has the best graphical formulation been chosen? Cannot the demonstration of RPOs for bigraphs be simplified? These questions will be all the more worth addressing if the theory, as it stands, is seen to contribute to specific applications.

There are also many links to be established with existing work, especially with the long tradition of research recorded in the Handbook of Graph Grammars and Computing by Graph Transformation [7]. Some of these links are discussed in the concluding section.

How the paper in organised In Section 2 we introduce bigraphs informally, by means of examples of familiar reactions and a term language for describing bigraphs. In Section 3 we review the necessary category theory, and in particular the RPO theory previously developed; we then introduce the abstract notion of *wide reactive system (WRS)* (of which bigraphs will provide a concrete instance), define transitions for them, and prove that under two conditions the resulting bisimilarity is a congruence. In Section 4 we define bigraphs in terms topographs and bigraphs, establishing several important properties; some of these follow from the existence of RPOs for bigraphs, which are proved in Appendix A for topographs and Appendix C for monographs. We also characterise the *idem-pushouts* (IPOs) constructed in Appendices A and B, which will underlie our derived transitions for bigraphs. Thus the reader can gain an understanding of the entire paper without reading any appendices.

In Section 5 we give the central definition of *bigraphical reactive system (BRS)*, a special kind of WRS; we show that bigraphs satisfy the two properties required for the congruence theorem, and hence deduce that bisimilarity in BRSs is a congruence. We also show how certain other WRSs enjoy the same result, provided they are related to BRSs by well-behaved functors; a particular case of this is a derived BRS with added structure that deals with scoping and binding of names. In Section 6 we fill out the theory of BRSs in many ways in preparation for application studies; in particular we justify the term language introduced earlier, we examine certain classes of transition which are (or may turn out to be) redundant in certain analyses, and we propose a definition of *parametric reaction rules* –illustrated in the π-calculus– which harmonises with the treatment of scoping and binding introduced in the preceding section. Finally, in Section 7 we discuss several lines for future work, including the need to establish links between the present theory and other research on related topics.

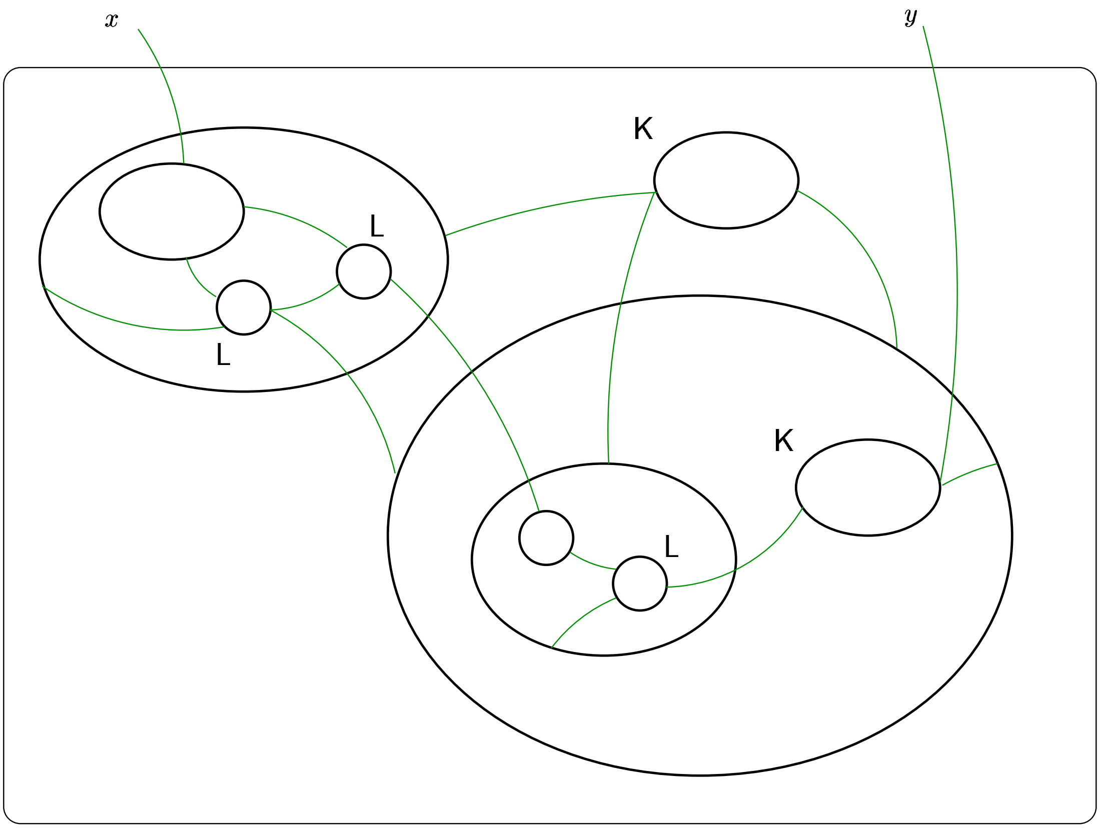


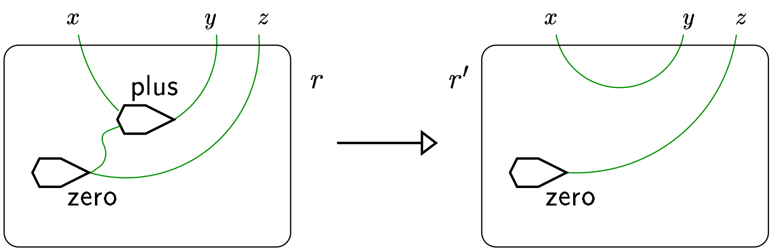
Figure 1: An example of a bigraph

2 Bigraphs in action

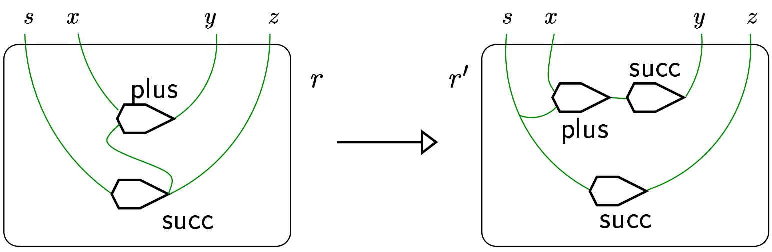
In this section we illustrate bigraphs, with examples showing the kinds of system they can represent, and the kind of mobility that they model. Along the way we introduce a simple term language for describing bigraphs. The whole section is informal and motivational; the subsequent mathematical development depends neither upon the diagrams nor upon the term language of this section. Subsection 2.1 contains examples drawn from familiar calculi; the diagrams allow us to explain how locality and connectivity co-operate. Subsection 2.2 briefly discusses the term language as a possible medium for expressing applications. In Subsection 2.3 some consequences of treating locality and connectivity independently are discussed.

2.1 Examples

Figure 1 shows an uninterpreted example of a bigraph. It has *nodes* (the ovals and circles) which support two kinds of structure; hence the term 'bigraph'. First, nodes may occur inside other nodes, so a bigraph has depth; since a node represents locality, in either a concrete or an abstract sense, we call this nesting structure of a bigraph its *topograph*. Second, nodes may be linked by *edges*, represented here by thin lines which may fork; we call this linked structure of a bigraph, which is independent of locality, its *monograph*. To each node is assigned a *control*, such as K or L, which tells us what kind of node it is. Each control has an *arity*, a finite ordinal; for example, L has arity three, so each L-node has an ordered set of three *ports*, at



$$[\,\mathsf{zero}(z) \mid \mathsf{plus}(x,z,y)\,] \longrightarrow x\!=\!y.\, [\,\mathsf{zero}(z)\,]$$



$$[\,\mathsf{succ}(s,z) \mid \mathsf{plus}(x,z,y)\,] \longrightarrow \nu w.\, [\,\mathsf{succ}(s,z) \mid \mathsf{plus}(x,s,w) \mid \mathsf{succ}(w,y)\,]$$

Figure 2: Two reaction rules for arithmetic

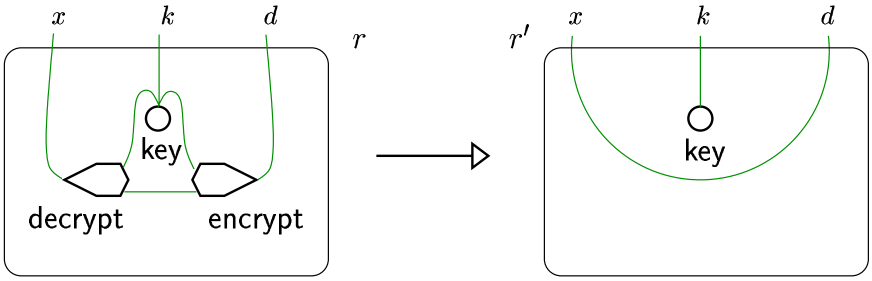
each of which an edge may impinge. It does not matter whether they impinge from inside or from outside the node. The diagram also shows the use of *names* x and y; such names allow a bigraph to be linked into larger bigraph.

The topograph and the monograph of any bigraph are coordinated by a node set, but are otherwise independent structures. But the *dynamics* of bigraphs, i.e. the reactions which may occur, depend upon both structural components: the precondition for any reaction is the presence of a certain pattern of nesting and linkage. A node in a bigraph may represent a great variety of computational objects: a physical location, an administrative region, a data constructor, a π-calculus input guard, an ambient, a cryptographic key, a message, a replicator, and so on. Its behaviour in any such role is determined by one or more *reaction rules*. A control K can be specified as *atomic*, meaning that nothing may be nested within a K node; if non-atomic it may also be specified as *active*, meaning that reactions may occur within K node.

We now give some typical reaction rules. Each rule consists of a precondition and a postcondition, which we call a *redex* and *reactum*. A *reaction* consists of the replacement of a redex occurring in a bigraph by the corresponding reactum; we shall make precise what we mean by ‘occurrence’ and ‘replacement’.

Example 1 (reactions without nesting) Our first examples, shown in Figures 2 and 3, involve no nesting. The graphs r, r' (redex and reactum) have a name set X as outer interface with their environment; here X takes the respective values $\{x, y, z\}$, $\{s, x, y, z\}$ and $\{x, k, d\}$. The shape of nodes is immaterial, except that if a control has arity more than one then its shape should not have rotational symmetry; thus the order of ports is unambiguous.

The first two reaction rules (Figure 2) are for arithmetic, with atomic controls zero, succ and plus having arities one, two and three respectively. They say, roughly, $0 + x \longrightarrow x$ and $s' + x \longrightarrow (s + x)'$ (where a prime here means successor). This strongly resembles the interaction nets of Lafont [17], but note that there is ‘sharing’ in the sense of Hasegawa’s sharing graphs [14]. In both cases one argument of plus is shared via



$$\nu y.\,[\,\mathsf{decrypt}(x,k,y)\mid \mathsf{key}(k)\mid \mathsf{encrypt}(d,k,y)\,] \longrightarrow x\!=\!d.\,[\,\mathsf{key}(k)\,]$$

Figure 3: A reaction rule for decryption

the name z; we envisage a context around the redex which uses the argument for some other purpose. The third reaction rule (Figure 3) is to do with security; a decrypting control succeeds in decrypting the datum named d encrypted on a key named k; as in the arithmetic rules, but more essentially, this key may be shared by other agents. The controls are again atomic.

In the first and third rules the reactum has a link which makes y or d an *alias* for x. Such an alias is essentially an 'explicit fusion' in the sense of Gardner and Wischik [11]; their calculus of explicit fusions was developed from the fusion calculus of Parrow and Victor [25] and from action calculi [22], and has guided the present development. The reactum in the second rule, and the redex in the third rule, both illustrate the use of a closed (i.e. unnamed) edge – respectively between the plus and succ nodes and between the decrypt and encrypt nodes. However, in the latter case it would be equally (or more) realistic for the edge to be named; this would represent the possibility that the encrypted datum may be accessible to other agents.

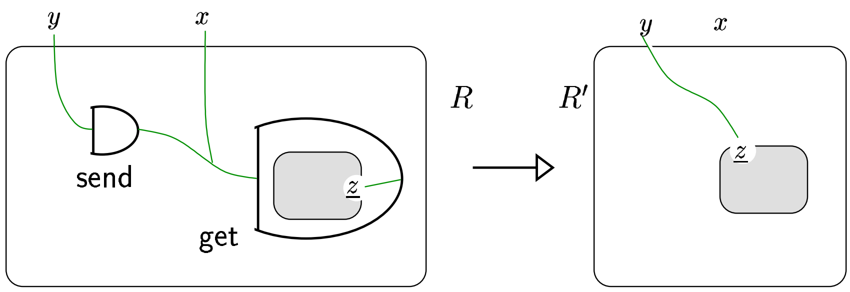
In all three rules, terms that define the redex and reactum are shown. Such terms can be used to define any bigraph. They indicate how to build a bigraph; its body consists of a parallel product of *molecules* (one for each node) and *holes* (see Example 2), with square brackets for grouping. (We use the term 'parallel product' instead of 'parallel composition' in this paper, because it is closer to a tensor product than to categorical composition.) The subterm for a molecule indicates the edges impinging on it by suitable identifiers. The term language is described at the end of this section. The notion of molecule stems from action calculi (Milner [22]); but the use here is more akin to that of fusion systems (Gardner [10]), in that edges are undirected hyperedges. As in that work, the only name-binding mechanism is the restriction operator ν appearing in the term language. Here it is used in the second arithmetic rule to describe a closed edge. ■

Our next three examples are more advanced; they show how bigraphs can define distributed reaction in mobile systems.

Example 2 (reaction in the π-calculus) Our second example (Figure 4) represents the familiar reaction rule of the asynchronous π-calculus (without summation)

$$\overline{x}\langle y\rangle \mid x(z).P \longrightarrow \{y/z\}P\ .$$

To present this reaction rule in terms of bigraphs we need two controls send and get, both with arity two. In the asynchronous π-calculus there are no output guards $\overline{x}\langle y\rangle.(-)$ and reaction is forbidden inside the input guard $x(z).(-)$; to match this we declare send atomic, and get non-atomic but inactive.



$$[\,\mathsf{send}(y,x) \mid \mathsf{get}(x,\underline{z})[-]\,] \longrightarrow y/\underline{z}.\, x.\, [\,-\,]$$

Figure 4: Reaction rule for the π-calculus

Since the rule is parametric, the redex bigraph is now a *context* R with both an *inner* and an *outer* interface as shown. Both interfaces have a width 1 (in this case – but not in the following examples) and a name set. The inner width 1 indicates that the context has a *hole*, as shown. Thus we begin to see that a bigraph (of which R and R' are examples) can be modelled as an arrow in a category with interfaces of the form $\langle m, X\rangle$ where the width m is a finite ordinal and the name set X is finite. The redex and reactum of a rule will always have the same outer interface; normally they will also have the same inner interface, when –as here– the rule is *linear* (see Subsection 2.3). In this example we have $R, R' : \langle 1, \{z\}\rangle \to \langle 1, \{x, y\}\rangle$.

Thus the outer interface of the agent parameter must contain z, and R will link it to the get-node. Our theory will allow polymorphism w.r.t. the parameter's name set; it may include any set W of other names (distinct from z) which are understood to be exported by R and R' to their outer interface – i.e. the reaction rule is transparent to W. Names of the inner interface, referred to as *conames*, are underlined ($\underline{z}$) to distinguish them from those of the outer interface, since the two name sets may have members in common.[1]

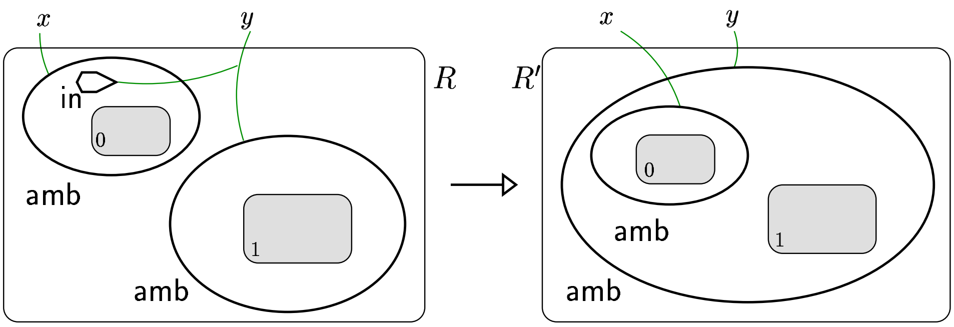
In applying the rule, one must be able to choose arbitrary instances for the names x and y. In fact, an arbitrary substitution for these names can be effected by the surrounding context with the help of *co-aliases*, dual to the notion of alias we have mentioned. We say more about them later. ■

Example 3 (reaction in the ambient calculus) In the ambient calculus of Cardelli and Gordon [2], one of the primitive forms of reaction is the movement of one *ambient* inside another.

Figure 5 shows how bigraphs may represent such a rule. We use two controls, each with arity one: amb for an ambient, and in for a 'command' to move its parent ambient somewhere else. We declare in to be atomic and inactive (in fact, it would make little sense for an atomic control to be active); in contrast with get in the π-calculus we declare amb to be non-atomic and active, since ambients are intended only to localize activity, not to inhibit it. In the redex R, the intent of the in command is to move the ambient named x inside the one named y. The rule has two parameters – the other contents of the x-ambient and the contents of the y-ambient – so the redex has two holes, or sites as we shall call them from now on. Thus in this example the interfaces are given by $R, R' : \langle 2, \emptyset\rangle \to \langle 1, \{x, y\}\rangle$. The parameters, though occupying distinct sites, may be linked by edges; this allows interaction between them – quite independently of their passive participation in the ambient reaction.

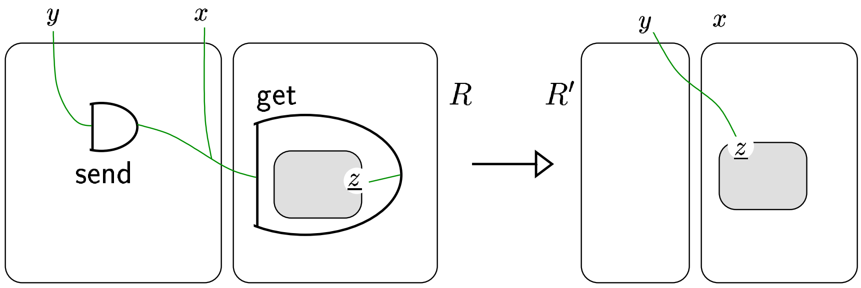
As in Example 2, this 'double agent' parameter may have an arbitrary name set Z, which will be exported at the outer interface of R and R'. Moreover, since locality and linkage are independent, these names pertain

[1]In a term, a singleton prefix such as 'x.' represents a name or coname of the interface that is not mentioned elsewhere.



$$[\,\mathsf{amb}(x)[\mathsf{in}(y) \mid -_0] \mid \mathsf{amb}(y)[-_1]\,] \longrightarrow [\,\mathsf{amb}(y)[\,\mathsf{amb}(x)[-_0] \mid -_1\,]\,]$$

Figure 5: Reaction rule for the ambient calculus



$$[\,\mathsf{send}(y,x)\,] \mid [\,\mathsf{get}(x,\underline{z})[-]\,] \longrightarrow {}^{y}/_{\underline{z}}.\,x.\,[\,] \mid [\,-\,]$$

Figure 6: Remote reaction rule for the π-calculus

to the whole double agent, not to one or other component. In the terms describing the redex and reactum, these component sites are denoted by $-_0$, $-_1$. (Our use of square brackets is not meant to match how the ambient calculus uses them.) ∎

Example 4 (remote reaction in the π-calculus) In the π-calculus reaction rule of Example 2 the redex and reactum have width 1; this means that the rule applies only when the send and get molecules are co-located. To put it another way, any context in which we place the redex will have these two nodes topographically adjacent. To allow a context to site them apart, we just change the outer width of the redex and reactum to 2, as shown in Figure 6; thus in this case we have $R, R' : \langle 1, \{z\}\rangle \to \langle 2, \{x, y\}\rangle$.

Such 'wide' reaction rules make it very easy to model mobility. They are especially interesting in the presence of one or more active controls, because they can be used to separate the components of a distributed redex but still allow it to react. We have already introduced amb as an example of an active control. By choosing Z in Example 3 (Figure 5) to be $\{x', y'\} \uplus W$, we may compose the ambient redex with an instance of the remote π-redex of Figure 6; we then obtain two interwoven but independent redexes, such that neither reaction precludes the other. This is not an unlikely occurrence in the internet, modelled at a suitable level of abstraction. ∎

2.2 Terms

In our examples we have informally adopted a term language for describing bigraphs. These terms may also serve as the core of a bigraph programming language. The elementary ingredients of the terms are *molecules*, such as $\mathsf{amb}(x)[\mathsf{in}(y) \mid -_0]$. If we want our description to include unique node-names $v \in V$ for the nodes of the bigraph, then we shall write this molecule as $\mathsf{amb}_v(x)[\mathsf{in}_u(y) \mid -_0]$ where $u, v \in V$; in that case no node-name v should occur more than once in the term. Here is how terms are built:

- A *term* describing a bigraph $G : \langle m, X\rangle \to \langle n, Y\rangle$ consists of an equivalence relation on the disjoint union $X+Y$, a set of local names written (if any) after ν, and a sequence of n regions. The equivalence relation may be written as equations between names and/or conames, with a singleton (e.g. x, $\underline{z}$) for any name or coname not otherwise mentioned. An equation $y = \underline{x}$ between a name and a coname may also be written $y/\underline{x}$, since it represents a substitution.
- A *molecule* such as $\mathsf{amb}(x)[\mathsf{in}(y) \mid -_0]$ describes a node and its contents; it consists of a control, followed in parentheses by an ordered list of identifiers denoting the edge impinging on each port, followed by a region describing its contents, if any.
- A *region* is a parallel product (in any order) of molecules and sites, in square brackets $[\cdot]$. A *site* is $-_i$ $(0 \le i < m)$. One may write $-$ for $-_0$.
- An *identifier* is either a name x, or a coname $\underline{y}$, or a local name.
- The *local names* are distinct from X and Y, and can be re-named by alpha-conversion.

Note that if two names or conames are equivalent they denote the same edge, so either may be used in a molecule to denote that edge.

As we said at the start of this section, the ensuing mathematical development does not depend upon this term language. The reason for introducing it here is to show how bigraphs relate to longer-established entities. Parallel product (composition) and restriction are of course familiar from many process algebras, and molecules come from action calculi [22]; as we shall see in Section 6, these all arise as derived operations from more primitive bigraph operations. Multi-hole contexts are also a familiar device; the way they are used here is consistent with the way Sewell uses them [27] to represent parametric reaction rules in his derivation of labelled transition systems. We have already mentioned that our equations such as $x = y$, $y = \underline{z}$ are akin to fusions [25, 10]; they are naturally extended here to refer to both names and conames. In fact, as we shall see in Construction 77, all the elements of the term language arise from the algebra of bigraphs.

When this algebra is further developed (the present paper is only a beginning) it will have much in common with the algebra of fusion systems (Gardner [10]). Similar algebra, without fusions, was also adopted by Honda [16] and Yoshida [30] for their combined study of process structures and process combinators. The novel element in our term language is the notion of *region* – representing the idea that the occupant of a multi-hole in a context is not a tuple of separate entities but a loosely connected bigraph. Indeed, although we have not formulated it, the reader may expect that to compose two bigraphical contexts, using the term language, we just insert the regions of one in the sites of the other. This intuition is correct. However, our term language represents only *normal forms* in the algebra of bigraphs, just as molecular forms are normal forms in action calculus. In composing two bigraphs we need an algebra like that of fusion systems to normalise the result, by manipulating restrictions and name equations so that the composite term has all such items at the top level.

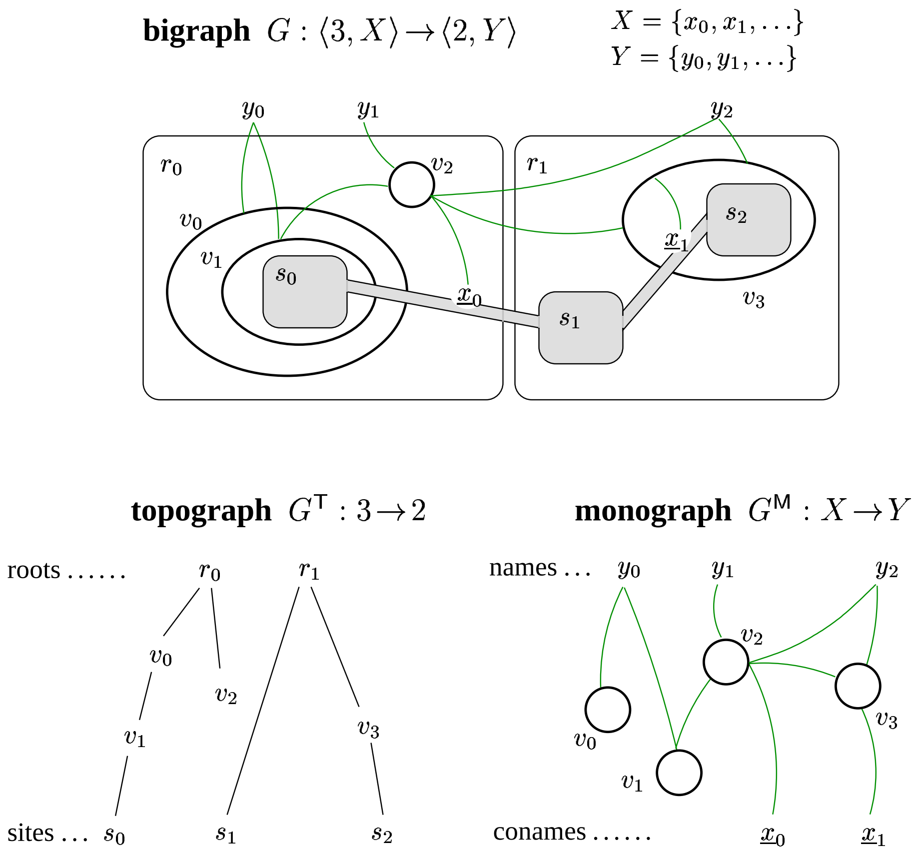


Figure 7: Resolving a bigraph into a topograph and a monograph

2.3 Discussion

The preceding examples illustrate how a bigraph context can be composed, or decomposed, in such a way that a single component may consist of linked regions located separately in the host bigraph. As just mentioned, the term language gives a way of representing such compositions algebraically. The term language also allows us to present the dynamics of bigraphical systems, as we have seen, as reaction rules each described by a pair of terms – a redex and a reactum. But it has been found that, using algebra alone, one cannot develop a behavioural theory based upon suitable *labelled transitions*.

It is therefore necessary to formalise bigraphs directly. To manage their complexity we represent a bigraph as a combination of two independent mathematical structures – a topograph and a monograph. Note that this *combination* is quite distinct from the two kinds of *composition* already discussed: contextual composition and parallel product. (Later we shall treat the former of these as a categorical composition, and the latter as derived from a tensor product.) But it is simply related to them; to compose (in either sense) two bigraphs, we first resolve them into their respective topographs and monographs, then compose these (in the given sense), and finally combine the results into a new bigraph.

Before treating bigraphs mathematically, it is helpful to see an example in Figure 7 of how a bigraph can be resolved into a topograph representing locality, and a monograph representing connectivity. (Controls are not shown in the diagram.) The nodes $v_0, \ldots, v_3$ are common to the two structures, which are otherwise independent. Note that the bigraph's interface $\langle 3, X \rangle \to \langle 2, Y \rangle$ combines the topograph interface $3 \to 2$ with the monograph interface $X \to Y$; there is nothing which determines that the names y_0, y_1, y_2 'belong' to any particular region of the bigraph (= root of the topograph), nor that the conames $\underline{x}_0, \underline{x}_1$ 'belong' to any particular site. In this diagram, though we shall not always do so, we have drawn tunnels linking the three sites to suggest the unity of the multiple bigraph that will occupy them.

The bigraphs of our theory will have an extra restriction: each site $-_i$ ($i < m$) must occur *exactly* once; that is, the contexts in our theory will be *linear*. However, for parametric reaction rules we can be more permissive if (as we shall) we treat each parametric rule (R, R') just as a way of describing a family of *ground* rules (r, r') – one for each way of filling the holes in (R, R'). In this case we shall still require the parametric redex R to be linear, but we may allow R' to be non-linear; this is because we place no restriction on the forking or aborting of links which is involved in copying or discarding an agent. Such rules are described as *left-linear* (see for example Sewell [27]). Non-linear rules are essential in practice; for example, we must be able to model the destruction of all occupants of an arbitrary site.

The reader may find it striking that in a bigraph $G : \langle m, X\rangle \to \langle n, Y\rangle$ we admit no association between the names Y and the roots (regions) n, nor between the conames X and the sites m. It is this dissociation that enables us to treat locality and connectivity independently, yielding a tractable theory. However, at the end of Section 5 we shall revisit this topic. It turns out that, having built the behavioural theory on the assumption of independence, we can introduce name-region dependency –in fact, define the scope of a name to be a certain region– by means of a certain functor, while preserving the important properties of the theory. We shall also see that, for full expressive power in the model, it is necessary to retain *some* names dissociated from regions.

3 Mathematical basis

In this section we develop the mathematical framework in which we shall analyse bigraphs. These definitions and theorems are an adaptation of work which was started in Leifer and Milner [19], and was then refined and extended by Leifer in his PhD Dissertation [18], which is the fullest exposition.

In Subsection 3.1 we introduce *(monoidal) well-supported precategories*. In Subsection 3.2 we recall the theory of *relative pushouts* (RPOs) and *idem pushouts* (IPOs). In Subsection 3.3 we develop *wide reactive systems* (*WRSs*), and derive *labelled transitions* for them. In Subsection 3.4 we lay a foundation for the behavioural theory of bigraphs.

Notation In categories or precategories we use '$\circ$', 'id' and '$\otimes$' for composition, identity and tensor product; we denote the domain I and codomain J of an arrow $f : I \to J$ by $\mathsf{dom}(f)$ and $\mathsf{cod}(f)$. In sets, Id_A will denote the identity function on a set A. We use '$\uplus$' for union between sets known to be disjoint; it should not be confused with the disjoint sum '$+$', which disjoins *any* pair of sets before taking their union. We assume a fixed representation of disjoint sums; for example, $X + P + Y$ means $(\{0\} \times X) \uplus (\{1\} \times P) \uplus (\{2\} \times Y)$, and $\sum_{v \in V} P_v$ means $\bigcup_{v \in V}(\{v\} \times P_v)$. We deal frequently with ordered pairs x_0, x_1; we use $\vec{x}$ to denote the pair. We often use i to range over the set $\{0, 1\}$ indexing such pairs; then we use $\bar{\imath}$ for the complement $1-i$ of i. A natural number m is often interpreted as a finite ordinal, i.e. $m = \{0, 1, \ldots, m-1\}$.

3.1 Well-supported precategories

Definition 1 (precategory, functor) A *precategory* $\mathbf{C}$ is defined exactly as a category, except that the composition of arrows is not always defined. Composition of arrows will be denoted by $\circ$. Composition with the identities id is always defined, and $\mathsf{id} \circ f = f = f \circ \mathsf{id}$. For associativity, either both $h \circ (g \circ f)$ and $(h \circ g) \circ f$ are undefined or both are defined and equal.[2]

A *subprecategory* $\mathbf{C}'$ of $\mathbf{C}$ is defined almost as a subcategory; we take $g \circ f$ to be defined in $\mathbf{C}'$ exactly when it is defined in $\mathbf{C}$. A *functor* $\mathcal{F} : \mathbf{C} \to \mathbf{C}'$ between precategories is defined almost as between categories. $\mathcal{F}$ is a total function on objects and on arrows, and preserves identities. It also preserves composition in this strong sense: if $g \circ f$ is defined in $\mathbf{C}$, then $\mathcal{F}(g) \circ \mathcal{F}(f)$ is defined in $\mathbf{C}'$ and equal to $\mathcal{F}(g \circ f)$. ∎

When considering precategories abstractly we shall use $I, J, K, \ldots$ to stand for their objects and $f, g, h, \ldots$ for their arrows. We shall extend category-theoretic properties to precategories without comment when they are unambiguous. For example, we shall work with epimorphisms; recall that an *epimorphism* (*epi*) in any (pre)category is an arrow f such that whenever $g \circ f = h \circ f$ then $g = h$.

We shall model bigraphs in a special kind of precategory. Every bigraph has an underlying set of nodes; we can compose two bigraphs whose node sets are disjoint, and the node set of the result is their union. This is a particular instance of *support*, defined as follows:

Definition 2 (well-supported precategory) A precategory $\mathbf{C}$ is *well-supported* if it has:

- a *support function* $|\cdot|$ mapping each arrow to a finite set called its *support*, such that $|\mathsf{id}_I| = \emptyset$, and also such that $g \circ f$ is defined — with support $|g \circ f| = |g| \uplus |f|$ — iff $|g| \cap |f| = \emptyset$ and $\mathsf{dom}(g) = \mathsf{cod}(f)$.
- for any arrow $f : I \to J$ and any injective map ρ whose domain includes $|f|$, an arrow $\rho \bullet f : I \to J$ called *the support translation by ρ of f* such that

$$
\begin{array}{ll}
(1) & \rho \bullet \mathsf{id}_I = \mathsf{id}_I \\
(2) & \rho \bullet (g \circ f) = \rho \bullet g \circ \rho \bullet f \\
(3) & \mathsf{Id}_{|f|} \bullet f = f \\
(4) & (\rho_1 \circ \rho_0) \bullet f = \rho_1 \bullet (\rho_0 \bullet f) \\
(5) & \text{if } \rho_0 \restriction |f| = \rho_1 \restriction |f| \text{ then } \rho_0 \bullet f = \rho_1 \bullet f \\
(6) & |\rho \bullet f| = \rho(|f|) \ .
\end{array}
$$

∎

We shall deal with precategories which have a partial tensor product as well as a partial composition. It rests upon an assignment of a set to each *object*, similar to the support of arrows. In bigraphs this will be a set of names, so we adopt that terminology here:

Definition 3 (tensor product, monoidal well-supported precategory) A well-supported precategory is *monoidal* if it has:

- a *naming* function $\| \cdot \|$ mapping each object to a finite *name set*, and a *unit* object ϵ with $\|\epsilon\| = \emptyset$;
- a partial *tensor product* $\otimes$ on objects, with unit ϵ, such that $I \otimes J$ is defined — with name set $\|I \otimes J\| = \|I\| \uplus \|J\|$ — iff $\|I\| \cap \|J\| = \emptyset$;
- a partial *tensor product* on arrows such that if $f_i : I_i \to J_i$ $(i = 0, 1)$ then $f_0 \otimes f_1$ exists — with support $|f_0 \otimes f_1| = |f_0| \uplus |f_1|$ — iff $\|I_0\| \cap \|I_1\| = \emptyset$, $\|J_0\| \cap \|J_1\| = \emptyset$ and $|f| \cap |g| = \emptyset$;

[2] This work was first announced in Milner [23], where a transcription error led to the inclusion of a weaker definition of precategory, in which the associativity condition is: if $h \circ g$ and $g \circ f$ are defined then either both $h \circ (g \circ f)$ and $(h \circ g) \circ f$ are undefined or both are defined and equal. We are only concerned here with precategories which satisfy the present stronger (and simpler) condition.

- a *symmetry* isomorphism $\gamma_{I,J} : I \otimes J \to J \otimes I$ for each pair I, J with disjoint name sets.

If ρ is a support translation then $\rho \cdot (f \otimes g) = \rho \cdot f \otimes \rho \cdot g$. The tensor and symmetries satisfy the equations of a strict symmetric monoidal category whenever both sides exist, i.e.

$$\begin{array}{ll} (1) & f \otimes (g \otimes h) = (f \otimes g) \otimes h \\ (2) & (f_1 \otimes g_1) \circ (f_0 \otimes g_0) = (f_1 \circ f_0) \otimes (g_1 \circ g_0) \\ (3) & \gamma_{I,\epsilon} = \mathsf{id}_I \\ (4) & \gamma_{J,I} \circ \gamma_{I,J} = \mathsf{id}_{I \otimes J} \\ (5) & \gamma_{I,K} \circ (f \otimes g) = (g \otimes f) \circ \gamma_{H,J} \qquad (\text{for } f : H \to I,\ g : J \to K)\ . \end{array}$$ ■

A well-supported precategory is what we need when we wish to preserve the identity of elements in the support of arrows[3]. Sometimes we do not; therefore we define a quotient with respect to arrow support (not object support) as follows:

Definition 4 (support quotient) Let **C** be a well-supported precategory. Two arrows $f, g : I \to J$ in **C** are *support-equivalent*, written $f \simeq g$, if there is a bijection $\rho : |f| \rightleftarrows |g|$ for which $\rho \cdot f = g$. By Definition 2(5) and (6) this is an equivalence relation. Then the *support quotient* of **C** is a category **[C]** whose objects are the objects of **C** and whose arrows are $\simeq$-equivalence classes of arrows in **C**:

$$\mathbf{[C]}(I, J) \triangleq \{ [f]_{\simeq} \mid f \in \mathbf{C}(I, J) \}\ .$$

Identities, composition and tensor product (when it exists) in **[C]** are given by

$$\begin{array}{rcl} \mathsf{id}_m & \triangleq & [\mathsf{id}_m]_{\simeq} \\ [g]_{\simeq} \circ [f]_{\simeq} & \triangleq & [g \circ f]_{\simeq} \\ [g]_{\simeq} \otimes [f]_{\simeq} & \triangleq & [g \otimes f]_{\simeq}\ . \end{array}$$

The funjctor $[\cdot] : \mathbf{C} \to \mathbf{[C]}$ is called the *support quotient functor* for **C**. ■

It is straightforward to show, using the axioms of Definition 2, that **[C]** is category, not merely a precategory. Note that to compose two arrows of **[C]**, which are equivalence classes, we choose representatives of the classes with disjoint supports, compose them in **C**, and then take the equivalence class of the result. The quotient does not affect objects; thus the tensor product in **[C]** is still partial.

3.2 Relative pushouts

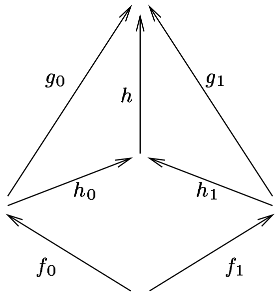
The central technical device in our theory is the notion of relative pushout (RPO), which we now define. This is to prepare for our definition of labelled transitions.

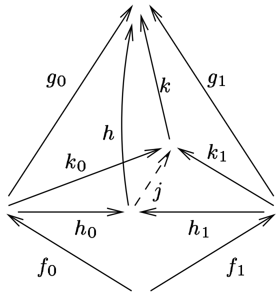
[3] Leifer's development [18] (see Chapter 7) made use of a special category $\mathsf{Track}(\mathbf{C})$ to *keep track of* the support elements of **C**. This allowed the RPO theory to be developed for categories rather than for well-supported precategories; that approach is perhaps more standard in category theory. The present more direct approach is justified by a forgetful functor $\mathcal{U} : \mathsf{Track}(\mathbf{C}) \to \mathbf{C}$ which preserves RPOs in both directions.

Notation In what follows we shall frequently use $\vec{f}$ to denote a pair f_0, f_1 of arrows in a precategory. If, for example, the two arrows share a domain H and have codomains I_0, I_1 we write $\vec{f} : H \to \vec{I}$.

Definition 5 (bound, consistent) If two arrows $\vec{f} : H \to \vec{I}$ share domain H, the pair $\vec{g} : \vec{I} \to K$ share codomain K and $g_0 \circ f_0 = g_1 \circ f_1$, then we say that $\vec{g}$ is a *bound* for $\vec{f}$. If such a pair $\vec{f}$ has any bound, then it is said to be *consistent*. ■

Definition 6 (relative pushout) In a precategory, let $\vec{g} : \vec{I} \to K$ be a bound for $\vec{f} : H \to \vec{I}$. An *RPO-candidate* – or just *candidate* – for $\vec{f}$ w.r.t. $\vec{g}$ is a triple $(\vec{h}, h)$ of arrows such that $\vec{h}$ is a bound for $\vec{f}$ and $h \circ h_i = g_i$ $(i = 0, 1)$. A *relative pushout* (*RPO*) for $\vec{f}$ w.r.t. $\vec{g}$ is a candidate $(\vec{h}, h)$ such that for any candidate $(\vec{k}, k)$ there is a unique arrow j for which $j \circ h_i = k_i$ $(i = 0, 1)$ and $k \circ j = h$. ■





The more familiar notion, a pushout, is a bound $\vec{h}$ for $\vec{f}$ such that *for any* bound $\vec{g}$ there exists an h which makes the left-hand diagram commute. Thus a pushout is a *least* bound, while an RPO provides a *minimal* bound w.r.t. a given bound $\vec{g}$. (In bigraphs we shall find that RPOs exist in cases where there is no pushout; see the discussions following Constructions 31 and 45.)

Now, supposing that we can create an RPO $(\vec{h}, h)$ for $\vec{f}$ w.r.t. $\vec{g}$, what happens if we try to iterate the construction? More precisely, is there an RPO for $\vec{f}$ w.r.t. $\vec{h}$? The answer lies in the following important concept:

Definition 7 (idem pushout) In a precategory, if $\vec{f} : H \to \vec{I}$ is a pair of arrows with common domain, then a pair $\vec{h} : \vec{I} \to J$ is an *idem pushout* (*IPO*) for $\vec{f}$ if $(\vec{h}, \mathsf{id}_J)$ is an RPO for $\vec{f}$ w.r.t. $\vec{h}$. ■

Then it turns out that the attempt to iterate the RPO construction will yield the *same* bound (up to isomorphism); intuitively, the minimal bound for $\vec{f}$ w.r.t. any bound $\vec{g}$ is reached in just one step. This is a consequence of the first two parts of the following proposition, which summarises the essential properties of RPOs and IPOs on which our work relies. They are proved for categories in Leifer's Dissertation [18] (see also Leifer and Milner [19]), and the proofs can be routinely adapted for precategories.[4]

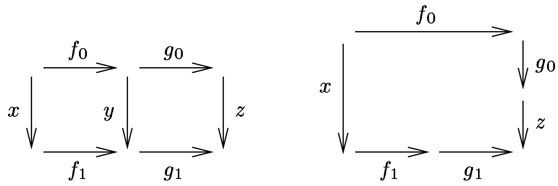
Proposition 8 (properties of RPOs)

(1) If an RPO for $\vec{f}$ w.r.t. $\vec{g}$ exists, then it is unique up to isomorphism.

(2) If $(\vec{h}, h)$ is an RPO for $\vec{f}$ w.r.t. $\vec{g}$, then $\vec{h}$ is an IPO for $\vec{f}$.

[4] This adaptation works for the definition of precategory in Definition 1, which is satisfied by our well-supported precategories. It does not appear to work for the weaker alternative mentioned in a footnote to that definition.

(3) If $\vec{h}$ is an IPO for $\vec{f}$, and an RPO exists for $\vec{f}$ w.r.t. $h \circ h_0, h \circ h_1$, then the triple $(\vec{h}, h)$ is such an RPO.



(4) (IPO pasting) Suppose that both diagrams shown here commute, and that the right-hand diagram has an RPO. Then in the left-hand diagram

- *if the two squares are IPOs then so is the big rectangle;*
- *if the big rectangle and the left square are IPOs then so is the right square.*

Another important consequence of Leifer's development (Proposition 4.8 in [18]) when a precategory is well-supported, is as follows:

Corollary 9 (IPO sliding) *In any well-supported precategory, IPOs are preserved by support translation; that is, if $\vec{g}$ is an IPO for $\vec{f}$ and ρ is any injective map whose domain includes the supports of $\vec{f}$ and $\vec{g}$, then $\rho \cdot \vec{g}$ is an IPO for $\rho \cdot \vec{f}$.*

3.3 Wide reactive systems

We now introduce a kind of dynamical system, of which bigraphs will be an instance.

In previous work [19, 18] a notion of reactive system was defined. In our present terms, this consists essentially of a well-supported precategory whose arrows are contexts –including agents– and two other ingredients: a set of agent-pairs (r, r') called reaction rules, and a subprecategory of so-called *active* contexts. The reaction relation $\longrightarrow\!\!\triangleright$ between agents is taken to be the smallest such that $D \circ r \longrightarrow\!\!\triangleright D \circ r'$ for every active context D and reaction rule (r, r'). For such systems we were able uniformly to derive labelled transitions, in such a way that several behavioural preorders and equivalences based upon these transitions are congruences, subject to two conditions: first, that sufficient RPOs exist in the precategory; second, that 'decomposition preserves activity' – i.e. $D \circ C$ active implies both C and D active.

In subsequent work, sufficient RPOs have been found in interesting cases (Leifer [18], Cattani et al [4]). In each of these cases the condition that decomposition preserves activity is also met, if we limit attention to contexts with a single hole. However, certain derived transition systems are unsatisfactory under this limitation, as Sewell [27] has pointed out with examples. Also, as we saw in Section 2, we wish to consider multi-hole bigraphical contexts – not only to represent parametric reaction rules but also to accommodate multiple or 'wide' agents, as in the remote π-calculus reaction rule in Example 4. There are other reasons for treating wide agents; for example, we would like to understand reactions occurring concurrently at different places in a system.

This gives rise to the possibility of contexts in which some sites may be active, i.e. permit reaction to occur, but not others. The following definitions, leading up to wide reactive systems, offer a way to handle these. We shall refine the above notion of reactive system as little as necessary to achieve that purpose. We shall also see that, if we specialise this new treatment to 'narrow' contexts (those with width one), we recover exactly the original notion of reactive system.

In what follows we shall use **Nat**, the strict symmetric monoidal category whose objects are finite ordinals, and whose arrows are functions between them.

Definition 10 (wide precategory) A *wide precategory* **A** is a well-supported precategory equipped with a *width* functor $\mathsf{width} : \mathbf{A} \to \mathbf{Nat}$ invariant under support translation, and a distinguished object ϵ called the *origin* such that $\mathsf{width}(\epsilon) = 0$. Moreover, for each permutation π of $\mathsf{width}(I)$ there is an isomorphism $\pi_I : I \to I$ in **A** with $\mathsf{width}(\pi_I) = \pi$, such that $(\cdot)_I$ respects identities and compositions.

If **A** –as a precategory– is monoidal with unit ϵ, and width preserves tensor product, then **A** equipped with width and ϵ is a *monoidal wide precategory*.

The objects $I, J, \ldots$ of **A** are called *interfaces*, and its arrows $A, B, \ldots$ are called *contexts*. Contexts with domain ϵ are called *agents*, and denoted by $a, b, \ldots$. ∎

We shall define bigraphs as a wide precategory in the next section. Meanwhile, from our discussion in Section 2 it is easy to see what 'width' means for bigraphs; the width of an interface $I = \langle m, X\rangle$ is just its multiplicity m, and the width of a bigraph $G : \langle m, X\rangle \to \langle n, Y\rangle$ is the function mapping each site $s \in m$ to the unique root $r \in n$ such that $r > s$. Thus width for bigraphs is purely topographical. Of course it says nothing about the nodes and controls in the topograph; but it will allow us to define –at the more abstract level of wide precategories– exactly which sites of a bigraph permit reaction. The notion of *place* will help us to formulate this.

Definition 11 (place) A *place* of an interface I with width m is a subset $\lambda \subseteq m$. We denote by $\mathsf{plac}(I)$ the set of places of I, i.e. the powerset of $\mathsf{width}(I)$.

The width function $f = \mathsf{width}(C)$ of a context $C : I \to J$ is extended to $\mathsf{plac}(I)$ by $f(\lambda) \triangleq \{f(i) \mid i \in \lambda\}$. The *offset* of λ by n is given by $n \cdot \lambda \triangleq \{n + i \mid i \in \lambda\}$. ∎

Places will also have a dual role, which we can illustrate in terms of bigraphs, where a place is just a set of sites. Consider a reaction rule (r, r') with width m. In the informal discussion of Section 2 we tended to assume that, in order for such a rule to 'fire', every root of the redex r should lie at an active site. This is natural for a narrow rule ($m = 1$); but at least for $m > 1$ it is not essential. Take for example the remote π-calculus rule, with width two, in Example 4; we may wish to require only that a part of its redex –say the get node– lie at an active site. One can imagine other rules where this makes good sense; for example, a rule to 'fetch' a value from a remote site may allow that site to be inactive, but require the fetch command itself – represented perhaps by a single node – to be at an active site. To allow this freedom, in our abstract formulation we shall associate with every rule whose interface is I a place $\lambda \in \mathsf{plac}(I)$; the redex will only fire in a context which is active at (every site in) λ.

We are now ready to add dynamics to wide precategories. The following definition introduces an *activity map* to determine the sites at which each context is active; the reader will find the discussion following the definition helpful in understanding the conditions we impose upon the activity map.

Definition 12 (wide reactive system (WRS)) A *(monoidal) wide reactive system (WRS)* over a (monoidal) wide precategory **A**, often written $\mathrm{WR}_{\mathbf{A}}$, has two further components besides **A**:

- a set $\mathsf{reacts} \subseteq \bigcup_I (\mathbf{A}(\epsilon, I)^2 \times \mathsf{plac}(I))$ of *reaction rules*, and

- for each I, J an *activity map* $\mathsf{act} : \mathbf{A}(I,J) \to \mathsf{plac}(I)$ such that

$$\begin{array}{ll} (1) & \mathsf{act}(D \circ C) = \mathsf{act}(C) \cap \mathsf{width}(C)^{-1}(\mathsf{act}(D)),\ \mathsf{act}(\mathsf{id}_I) = \mathsf{width}(I)\ ; \\ (2) & \mathsf{act}(C \otimes D) = \mathsf{act}(C) \uplus \mathsf{width}(\mathsf{dom}(C)) \cdot \mathsf{act}(D)\ \text{(for WR}_\mathbf{A}\text{ monoidal).} \end{array}$$

  We say C is *active at* i if $i \in \mathsf{act}(C)$; similarly C is *active at* λ if $\lambda \subseteq \mathsf{act}(C)$, and C is *active* if $\mathsf{act}(C) = \mathsf{width}(\mathsf{dom}(C))$.

These components both respect support equivalence; that is, (1) if $r \doteq s$, $r' \doteq s'$ and $(r, r', \lambda) \in \mathsf{reacts}$ then $(s, s', \lambda) \in \mathsf{reacts}$, and (2) if $D \doteq E$ then $\mathsf{act}(D) = \mathsf{act}(E)$.

The *reaction relation* $\longrightarrow\!\!\!\!\triangleright$ over agents is defined as follows: $a \longrightarrow\!\!\!\!\triangleright\, a'$ iff there exist a reaction rule (r, r', λ) and a context D with $\lambda \subseteq \mathsf{act}(D)$ such that $a = D \circ r$ and $a' \doteq D \circ r'$. ■

The first part of condition (1) asserts: $D \circ C$ is active at i iff C is active at i and D is active at $\mathsf{width}(C)(i)$. If C has lower width m then condition (2) asserts: $C \otimes D$ is active at i iff either $i < m$ and C is active at i or $i \geq m$ and D is active at $i - m$. We leave it to the reader to check that these conditions make sense – i.e. that they are consistent with the equations governing composition and tensor product.

Now let us revisit the special case in which all contexts are narrow, so that $\mathsf{width}(C) = \{0 \mapsto 0\}$ for all C; so we can say C is *active* if it is active at 0. Condition (1) is then equivalent to requiring that the active contexts form a subprecategory closed under decomposition. Thus, as promised, we have a proper generalisation of the conditions under which the original congruence theorems [18, 19] were proved.

Definition 12 ensures that all its ingredients are closed under support equivalence, allowing us in Definition 14 to divide $\mathrm{WR}_\mathbf{A}$ by $\doteq$, forming a quotient WRS. The following is immediate:

Proposition 13 (support translation of reactions) *Reaction in a WRS is closed under support equivalence, i.e. if $a \doteq b$, $a' \doteq b'$ and $a \longrightarrow\!\!\!\!\triangleright\, a'$ then $b \longrightarrow\!\!\!\!\triangleright\, b'$.*

We may therefore define reaction in the support quotient of any WRS, as follows:

Definition 14 (quotient WRS) Let $\mathrm{WR}_\mathbf{A}$ be a wide reactive system over $\mathbf{A}$. Then a wide reactive system $\mathrm{WR}_{[\mathbf{A}]}$ over $[\mathbf{A}]$ is defined as follows:

- the reaction rules are $([r], [r'], \lambda)$, for each reaction rule (r, r', λ) in $\mathrm{WR}_\mathbf{A}$;
- the active sites are given by $\mathsf{act}([D]) \triangleq \mathsf{act}(D)$.

The preceding results ensure that $[a] \longrightarrow\!\!\!\!\triangleright\, [a']$ in $\mathrm{WR}_{[\mathbf{A}]}$ iff $a \longrightarrow\!\!\!\!\triangleright\, a'$ in $\mathrm{WR}_\mathbf{A}$. ■

Thus reaction in $\mathrm{WR}_{[\mathbf{A}]}$ perfectly mirrors reaction in $\mathrm{WR}_\mathbf{A}$. The importance of the quotient WRS is that, being based upon a category and not merely a precategory, it is amenable to a simpler algebraic theory (which we do not discuss fully in this paper). However, important structural properties are lost in taking the support quotient. In fact, as we shall show in Subsection 4.3, a WRS for bigraphs has sufficient RPOs to allow the derivation of satisfactory labelled transition systems, while its support quotient does not. This has been a crucial factor in determining our theory. In the following subsection we shall show how to derive labelled transition systems for suitable WRSs; we shall also show how these systems may be transferred to their support quotients.

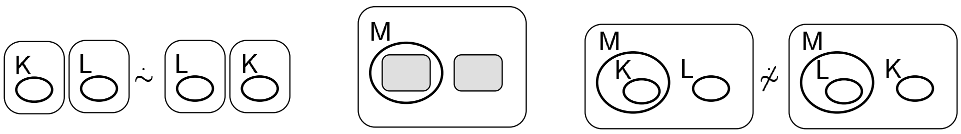
3.4 Transitions and bisimilarity

We now turn to the problem of deriving labelled transitions for an arbitrary WRS, and thence defining behavioural equivalence. The definitions and results of this section will later be applied to bigraphs; they will also be applied to a broader class of WRSs functorially related to bigraphs.

The definition in Leifer and Milner [19] was in terms of IPOs, as follows: $a \xrightarrow{F} a'$ iff there is a reaction rule (r, r') and an active context D for which F, D is an IPO for a, r and $a' = D \circ r'$. We shall do something close to this, with two differences. First, this is an unplaced transition of a; the transition tells us the extra context F needed by a to create a redex, but does not specify *where* those parts of the redex already within a lie. If we think of the transition as an observation of a, then it is an unplaced observation. In fact, even before developing the theory of bigraphs, we can give an example in which the version of bisimilarity arising from such transitions is not a congruence:

Example 5 (unplaced bisimilarity is not a congruence) This example shows that bisimilarity based upon unplaced transitions, which we denote by $\dot{\sim}$, is not in general a congruence for bigraphical systems. Take the signature $\mathcal{K} = \{\mathsf{K}, \mathsf{L}, \mathsf{M}\}$, each with arity zero; let K, L be atomic and M non-atomic but inactive. Ports, names and wiring are irrelevant in this example, so we take interfaces to be just finite ordinals (widths). We write $\mathsf{K}, \mathsf{L} : 0 \to 1$ for agents consisting of a single atomic node, and $\mathsf{M}_1 : 1 \to 1$ for the elementary inactive context based upon M. Let there be a single (unplaced) reaction rule K, L.

Now consider the two agents $a, b : 0 \to 2$ with width two, where $a = \mathsf{K} \mid \mathsf{L}$ and $b = \mathsf{L} \mid \mathsf{K}$, illustrated below. If transitions are unplaced then one can prove $a \mathrel{\dot{\sim}} b$, since no such transition can observe the difference in the place of K (we omit the details).



$a \mathrel{\dot{\sim}} b$ $\qquad$ $C \triangleq [\mathsf{M}_1 \otimes \mathsf{id}_1]$ $\qquad$ $C \circ a \mathrel{\dot{\not\sim}} C \circ b$

But if we put a and b in turn in the context $C \triangleq [\mathsf{M}_1 \mid \mathsf{id}_1] : 2 \to 1$, as shown, it turns out that $C \circ a \mathrel{\dot{\not\sim}} C \circ b$. For $C \circ b$ has a transition $\xrightarrow{\mathsf{id}_1}$ since its K node is a redex with active place; but $C \circ a$ has no such transition since its K node is placed inactively. ■

The second difference from the cited definition is that –as for reaction– we want labelled transitions to be preserved by support translation, so that we also get a good definition of transition for the support quotient $\mathrm{WR}_{[\mathbf{A}]}$. These two differences lead to the following:

Definition 15 (placed transition) Let $\mathrm{WR}_{\mathbf{A}}$ be a wide reactive system. Then the quadruple a, F, λ, a' is a *placed transition* of $\mathrm{WR}_{\mathbf{A}}$, written $a \xrightarrow{F}_{\lambda} a'$, if there exist a reaction rule (r, r', μ) and a context D active at μ for which F, D is an IPO for a, r, $\mathsf{width}(D)(\mu) = \lambda$ and $a' \simeq D \circ r'$. In this case we say that the reaction rule and the IPO *underlie* the transition.[5] ■

[5] By allowing $a' \simeq D \circ r'$, rather than the stricter requirement $a' = D \circ r'$, we allow the derivation of transitions for the support quotient WRS (Definition 16). However, for other purposes –e.g. to keep track of causal dependency among transitions– we may wish to use the stricter condition. This point is taken up in the concluding section.

Returning briefly to Example 5 we now see that with these transitions we shall not have a and b bisimilar; for while a has the transition $\xrightarrow{\mathsf{id}_2}_{\{0\}}$, b has instead the transition $\xrightarrow{\mathsf{id}_2}_{\{1\}}$.

In what follows we shall often omit 'placed', and talk of *transitions* and *transition systems*, since we shall not deal further with unplaced transitions – or if we do, we shall explicitly call them 'pure'.

Crucially, we deduce from IPO sliding (Corollary 9) that

Proposition 16 (support translation of transitions) *In any wide reactive system* $\mathrm{WR}_{\mathbf{A}}$, *transition is preserved by support translation. That is, let* $a \simeq b$, $F \simeq G$ *and* $a' \simeq b'$ *in* $\mathbf{A}$, *with* $F \circ a$ *and* $G \circ b$ *defined; then* $a \xrightarrow{F}_{\lambda} a'$ *iff* $b \xrightarrow{G}_{\lambda} b'$.

This validates the following definition of transition for $\mathrm{WR}_{[\mathbf{A}]}$*:* $[a] \xrightarrow{[F]}_{\lambda} [a]'$ *iff* $F \circ a$ *is defined and* $a \xrightarrow{F}_{\lambda} a'$ *in* $\mathrm{WR}_{\mathbf{A}}$.

Having defined transitions for both WRSs and their support quotients, we may define behavioural equivalences and preorders in familiar ways. Here we shall limit attention to strong bisimilarity. (Throughout this paper we shall omit 'strong' since we do not define or use weak bisimilarity.)

Definition 17 (wide bisimilarity) Let $\mathrm{WR}_{\mathbf{A}}$ be a wide reactive system over $\mathbf{A}$. Then *wide bisimilarity* is the largest symmetric relation $\sim$ between agents of $\mathbf{A}$ with equal codomain such that if $a \sim b$ and $a \xrightarrow{F}_{\lambda} a'$, then whenever $F \circ b$ is defined there exists b' such that $b \xrightarrow{F}_{\lambda} b'$ and $a' \sim b'$.

Wide bisimilarity in $\mathrm{WR}_{[\mathbf{A}]}$, the support quotient of $\mathrm{WR}_{\mathbf{A}}$, is the largest symmetric relation $\sim$ between agents of $[\mathbf{A}]$ with equal codomain such that if $[a] \sim [b]$ and $[a] \xrightarrow{[F]}_{\lambda} [a']$, then there exists $[b']$ such that $[b] \xrightarrow{[F]}_{\lambda} [b']$ and $[a'] \sim [b']$. ■

Note the slight departure from the standard definition; here we must require $F \circ b$ to be defined. This is merely a technical detail; for note that whenever $F \circ a$ is defined there will always exist $F' \simeq F$ for which *both* $F' \circ a$ and $F' \circ b$ are defined.

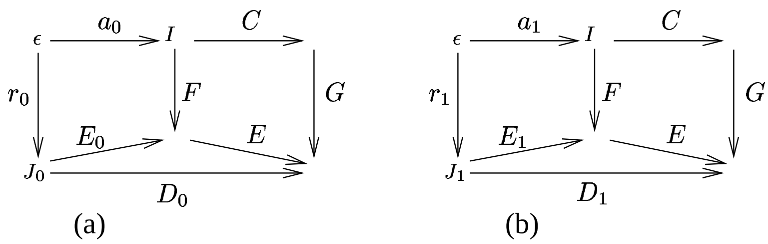
From the definition, together with that of transitions in Definition 15 and Proposition 16, it is straightforward to deduce that wide bisimilarity equivalence in $\mathrm{WR}_{[\mathbf{A}]}$ mirrors that in $\mathrm{WR}_{\mathbf{A}}$:

Proposition 18 (support quotient preserves wide bisimilarity) *Let* $\mathrm{WR}_{[\mathbf{A}]}$ *be the support quotient of a wide reactive system* $\mathrm{WR}_{\mathbf{A}}$. *Then* $a \sim b$ *in* $\mathrm{WR}_{\mathbf{A}}$ *iff* $[a] \sim [b]$ *in* $\mathrm{WR}_{[\mathbf{A}]}$.

We may now add force to the claim that our derived transitions are tractable; for we can prove our central theorem for wide reactive systems —that if there are sufficient RPOs in a WRS then wide bisimilarity is congruential. First, we explain what we mean by 'sufficient RPOs':

Definition 19 (redex-RPOs) A wide reactive system $\mathrm{WR}_{\mathbf{A}}$ *has all redex-RPOs* if, for every reaction rule (r, r', λ), any pair a, r in $\mathbf{A}$ has an RPO w.r.t. any bound F, D in $\mathbf{A}$, where D is active at λ. ■

Theorem 20 (congruence of wide bisimilarity) *In a wide reactive system with all redex-RPOs, wide bisimilarity of agents is a congruence; that is, if* $a_0 \sim a_1$ *then* $C \circ a_0 \sim C \circ a_1$.



Proof The proof is along the lines of Theorem 3.9 in Leifer [18]. We establish the bisimulation

$$\mathcal{S} \triangleq \{(C \circ a_0, C \circ a_1) \mid a_0 \sim a_1,\ C \text{ any context}\} .$$

Suppose then that $a_0 \sim a_1$, and that $C \circ a_0 \xrightarrow{G}_{\mu} b'_0$, for some G such that $G \circ C \circ a_1$ is defined. It is enough to find b'_1 such that $C \circ a_1 \xrightarrow{G}_{\mu} b'_1$ and $(b'_0, b'_1) \in \mathcal{S}$.

There exist a reaction rule $(r_0, r'_0, \kappa_0) : \epsilon \to J_0$ and a context D_0 active at κ_0 such that the large rectangle shown at (a) is an IPO, with $\mathsf{width}(D_0)(\kappa_0) = \mu$ and $b'_0 = D_0 \circ r'_0$. Then because all redex-RPOs exist, there exists a triple (F, E_0, E) forming an RPO as shown, and from Proposition 8 we deduce that the right square at (a) is an IPO.

Now from Definition 12 we know that E_0 is active at κ_0, so $a_0 \xrightarrow{F}_{\lambda} a'_0$ where $\lambda = \mathsf{width}(E_0)(\kappa_0)$ and $a'_0 \triangleq E_0 \circ r'_0$; we also know that E active at λ. From the former, since $a_0 \sim a_1$ there is a transition $a_1 \xrightarrow{F}_{\lambda} a'_1$ with $a'_0 \sim a'_1$. (Note that $F \circ a_1$ is defined, since $E \circ F \circ a_1 = G \circ C \circ a_1$ is defined.) So the left square shown at (b) is an IPO for some reaction rule $(r_1, r'_1, \kappa_1) : \epsilon \to J_1$, where E_1 is active at κ_1 and $\mathsf{width}(E_1)(\kappa_1) = \lambda$, and also $a'_1 = E_1 \circ r'_1$.

Pasting the right IPO square of (a) to this square we obtain – as shown at (b) – a large rectangle which is also an IPO by Proposition 8, and by Definition 12 again $D_1 = E \circ E_1$ is active at κ_1. Also $\mathsf{width}(D_1)(m_1) = \mu$. Hence $C \circ a_1 \xrightarrow{G}_{\mu} b'_1$ where $b'_1 \triangleq D_1 \circ r'_1$. Moreover $(b'_0, b'_1) \in \mathcal{S}$ as required, because $b'_0 = E \circ a'_0$ and $b'_1 = E \circ a'_1$ with $a'_0 \sim a'_1$. ∎

As with 'placed' transition, we shall henceforth often omit the adjective 'wide' when discussing bisimilarity. We should remark that we are taking (strong) bisimilarity as a representative of many preorders and equivalences; Leifer [18] has proved congruence theorems for several others, and we expect that those results can be transferred to the present setting.

In the next section we study bigraphs. In Section 5 we shall establish bigraphical reactive systems (BRSs) as WRSs, and deduce a congruence theorem for them. To broaden the application of that theorem, in Subsection 5.2 we study functors between WRSs in general; then in Subsection 5.3 we show how congruence results can be transferred from BRSs to a variety of other WRSs via suitable functors.

4 Bigraph theory

We are now ready to embark on the theory of bigraphs, in the light of the framework of Section 3.

Definition 21 (signature) A *signature* $\mathcal{K}$ is a set whose elements are called *controls*. For each $K \in \mathcal{K}$ the signature provides a finite ordinal $ar(K)$, its *arity*; it also determines which controls are *atomic*, and which of the non-atomic controls are *active*. ∎

The arity of a control indexes the ports of a node with that control. A node with an atomic control may not contain sites or other nodes. If a control is active (hence also non-atomic) then reactions are permitted inside it. (Thus the term 'active' pertains to what happens inside the node, not how the node itself may react.) In refinements of the theory the signature may carry further information, such as a *sign* and/or a *type* for each port. The sign may be used, for example, to enforce the restriction that each negative port is connected to exactly one positive port, as in action calculi [4, 22]. We shall revisit such refinements in Section 5.

We shall normally work with a fixed but unspecified signature. We give the definition of a bigraph top-down; that is, we define first how a bigraph is built from its two structural components, and then define those components themselves.

Definition 22 (bigraph) A *bigraph* over the signature $\mathcal{K}$ takes the form $G = (V, \mathit{ctrl}, G^{\mathsf{T}}, G^{\mathsf{M}}) : I \to J$ where: $I = \langle m, X\rangle$ and $J = \langle n, Y\rangle$ are its *inner* and *outer interfaces*, each combining a *width* (a finite ordinal) with a finite *name set*; V is a set of *nodes*; $\mathit{ctrl} : V \to \mathcal{K}$ is the *control function* assigning a control to each node; $G^{\mathsf{T}} : m \to n$ and $G^{\mathsf{M}} : X \to Y$ are respectively a topograph and a monograph (see Definitions 23 and 34), each having V as its node set and ctrl as its control function.

We refer to G as the *combination* of G^{T} and G^{M}; we write it as $G = \langle G^{\mathsf{T}}, G^{\mathsf{M}}\rangle$ when the shared parts V and ctrl are understood. ■

Thus each interface combines two components, one for topographs and the other for monographs. The reader may like to recall from Figure 7 how a bigraph combines two components. As we shall see below, bigraphs thus combine two well-studied mathematical structures: trees for the topographs, and equivalence relations (on ports) for the monographs. We proceed to define and study these components in Subsections 4.1 and 4.2 respectively. In Subsection 4.3 we revisit bigraphs, developing their structure by combining attributes from topographs and monographs.

4.1 Topographs

In this subsection we develop the theory of topographs far enough to define their contribution to dynamics.

Definition 23 (topograph) A *topograph* $A = (V, \mathit{ctrl}, \mathit{prt}) : m \to n$ has an *inner width* m and an *outer width* n, both finite ordinals; a finite set V of nodes with a control function $\mathit{ctrl} : V \to \mathcal{K}$; and a *parent function* $\mathit{prt} : m \uplus V \to V \uplus n$ which is *acyclic*, i.e. such that $\mathit{prt}^k(v) \neq v$ for all $k > 0$ and $v \in V$. An *atomic* node – i.e. one whose control is atomic – may not be a parent. We write $w >_A w'$, or just $w > w'$, to mean $w = \mathit{prt}^k(w')$ for some $k > 0$. ■

The acyclicity condition makes the parent function prt represent a forest of n unordered trees; thus n indexes the *roots* of the trees. The other points of each tree are either nodes $v \in V$, or *sites* indexed by m; a site may only occur as a leaf of a tree.

The sites and roots provide the means of composing the forests of two topographs; each root of the first is planted in a distinct site of the second. Formally:

Definition 24 (precategory of topographs) The precategory **Top** has finite ordinals as objects and topographs as arrows. The composition $A_1 \circ A_0 : m_0 \to m_2$ of two topographs $A_i = (V_i, \mathit{ctrl}_i, \mathit{prt}_i) : m_i \to m_{i+1}$ $(i = 0, 1)$ is defined when the two node sets are disjoint; then $A_1 \circ A_0 \triangleq (V, \mathit{ctrl}, \mathit{prt}) : m_0 \to m_2$

where $V = V_0 \uplus V_1$, $ctrl = ctrl_0 \uplus ctrl_1$, and $prt = (\mathsf{Id}_{V_0} \uplus prt_1) \circ (prt_0 \uplus \mathsf{Id}_{V_1})$. The identity topograph at m is $\mathsf{id}_m \triangleq (\emptyset, \emptyset_{\mathcal{K}}, \mathsf{Id}_m) : m \to m$. ■

It is trivial to check that $A \circ \mathsf{id} = A = \mathsf{id} \circ A$. To see that composition is associative, suppose that $A_2 = (V_2, prt_2) : m_2 \to m_3$ is a third topograph with V_2 disjoint from V_0, V_1; then the parent function of $A_2 \circ (A_1 \circ A_0)$ is

$$\begin{aligned}(\mathsf{Id}_{V_0 \uplus V_1} \uplus prt_2) \circ (prt \uplus \mathsf{Id}_{V_2}) &= (\mathsf{Id}_{V_0 \uplus V_1} \uplus prt_2) \circ (((\mathsf{Id}_{V_0} \uplus prt_1) \circ (prt_0 \uplus \mathsf{Id}_{V_1})) \uplus \mathsf{Id}_{V_2}) \\ &= (\mathsf{Id}_{V_0} \uplus \mathsf{Id}_{V_1} \uplus prt_2) \circ (\mathsf{Id}_{V_0} \uplus prt_1 \uplus \mathsf{Id}_{V_2}) \circ (prt_0 \uplus \mathsf{Id}_{V_1} \uplus \mathsf{Id}_{V_2}) ,\end{aligned}$$

and the parent function of $(A_2 \circ A_1) \circ A_0$ expands to the same.

When dealing with many topographs A, B, $\ldots$, instead of indexing their parent functions as prt_A, prt_B etc. we shall find it more convenient to abuse notation and denote the parent function of a topograph A again by A. The context will prevent ambiguity; for example in $B \circ A$ we are talking of topographs, while in $B(A(v))$ we are talking of their parent functions. Thus $(B \circ A)(v)$ means the parent function of the composite topograph $B \circ A$ applied to the node v.

Proposition 25 (isomorphisms in topographs) *An arrow $\iota : m \to n$ in* **Top** *is an isomorphism iff it has no nodes, $m = n$, and its parent function is a bijection.*

Here are some basic properties:

Definition 26 (barren, shallow, deep, active) A node or root in a topograph is *barren* if it has no children. A site or node is *shallow* if its parent is a root, otherwise *deep*. A topograph is *shallow* if all its sites are shallow, otherwise *deep*; it is *active* at a site s if whenever $v > s$ then $ctrl(v)$ is active. ■

Epimorphisms will play an important role, both for topographs and for monographs. Recall that in the category of sets with functions the epis are the surjective functions. Here we find something analogous:

Proposition 27 (epimorphisms in topographs) *In* **Top***, a topograph is an epimorphism iff it has no barren roots.*

Proof ($\Rightarrow$) Let $A : m \to n$ have a barren root, say $0 \in n$. Take any pair $B, B' : n \to p$, with support disjoint from that of A, which are identical except that $B(0) \neq B'(0)$. Then $B \neq B'$ but $B \circ A = B' \circ A$, so A is not epi.

($\Leftarrow$) Assume A has no barren roots, and let $B \circ A = B' \circ A$. Then B and B' must have the same node set V, and for each $v \in V$ we have $B(v) = (B \circ A)(v) = (B' \circ A)(v) = B'(v)$. Also for any $r \in n$, since r is not barren we have $A(w) = r$ for some node or site w of A; hence $B(r) = (B \circ A)(w) = (B' \circ A)(w) = B'(r)$. Hence $B = B'$; but B and B' were arbitrarily chosen, so A is epi. ■

What is a suitable tensor product for **Top**? For example, we do not want $A \otimes B$ to have the effect of merging nodes from A and B. So we choose a partial tensor product, with $A \otimes B$ defined exactly when the supports are disjoint, in which case its support is $|A| \uplus |B|$.

Definition 28 (tensor product) The *tensor product* $\otimes$ in **Top** is defined as follows: On objects, we take $m \otimes n \triangleq m + n$. For two topographs $A_i : m_i \to n_i$ $(i = 0, 1)$ we take $A_0 \otimes A_1 : m_0 + m_1 \to n_0 + n_1$ to be defined when A_0 and A_1 have disjoint node sets; for the parent function, we first adjust the sites and roots of A_1 by adding m_0 and n_0 respectively, then take the union of the two parent functions. ∎

For the rest of this subsection we shall consider a pair $\vec{A} : h \to \vec{m}$ of topographs with common domain h, and we shall adopt a convention for naming their nodes. These may be partitioned into three sets: V_i private to A_i $(i = 0, 1)$ and V_2 common to both. From now on we shall consistently use $v_i, v_i', \ldots$ to range over V_i $(i = 0, 1, 2)$. We shall also take a simplifying liberty; we shall just treat the case where $h = 0$, since the sites in h are treated by the construction exactly as members of V_2. It comes to the same thing if we think of $v_2, v_2' \ldots$ ranging over $h \uplus V_2$.

As the next step towards IPOs, we define certain conditions on $\vec{A}$ which we shall prove to be necessary and sufficient for their consistency (Definition 5).

Definition 29 (consistency conditions) We define four conditions on a pair $\vec{A} : 0 \to \vec{m}$ of topographs as follows, where $i = 0, 1$ (recalling also that $\bar{\imath}$ means $1 - i$):

(C0)	$ctrl_0(v_2) = ctrl_1(v_2)$		
(C1)	$A_i(v_i) \in V_i \uplus m_i$		
(C2)	If $A_i(v_2) \in V_i$ then	(C2.1)	$A_{\bar{\imath}}(v_2) \in m_{\bar{\imath}}$
		(C2.2)	$\not\exists v_{\bar{\imath}}.\ A_{\bar{\imath}}(v_{\bar{\imath}}) = A_{\bar{\imath}}(v_2)$
		(C2.3)	$\forall v_2'.\ A_{\bar{\imath}}(v_2') = A_{\bar{\imath}}(v_2) \Rightarrow A_i(v_2') = A_i(v_2)$
(C3)	If $A_i(v_2) \in V_2$ then $A_{\bar{\imath}}(v_2) = A_i(v_2)$.		

∎

If we assume a bound $\vec{B}$ for $\vec{A}$, then these conditions follow easily from the definition of composition. So:

Proposition 30 (consistency in topographs) *If the pair $\vec{A}$ is consistent then conditions (C0) - (C3) hold.*

Let us express the consistency conditions more informally, in words rather than in symbols. Talk of A_0, A_1 as 'left' and 'right'; so for example the left parent of a shared node v_2 refers to $A_0(v_2)$ –which may be a node or a site– and so on. The conditions declare the following (also with 'left' and 'right' exchanged) for every shared node v_2:

- (C0) v_2 has the same control, left and right.
- (C1) All the children of v_2 are shared.
- (C2) If the left parent of v_2 is an unshared node then
  - (C2.1) its right parent is a root;
  - (C2.2) all its right siblings are shared …
  - (C2.3) … and are also among its left siblings;
- (C3) If the left parent of v_2 is shared then this is also its right parent.

Example 6 (consistent topographs) Consider the pair $\vec{A}$ of topographs in Figure 8, each with two roots and no sites; as above, nodes with subscript 2 are shared. (Controls are not shown). It is worth checking that conditions (C1)–(C3) hold. What happens if an extra node u is added to A_1 as a sibling of v_2? If u is unshared then (C2.2) is violated, so consistency is lost. If u is shared, then to preserve consistency –especially (C2.3)– u must also become a sibling of v_2 in A_0. ∎

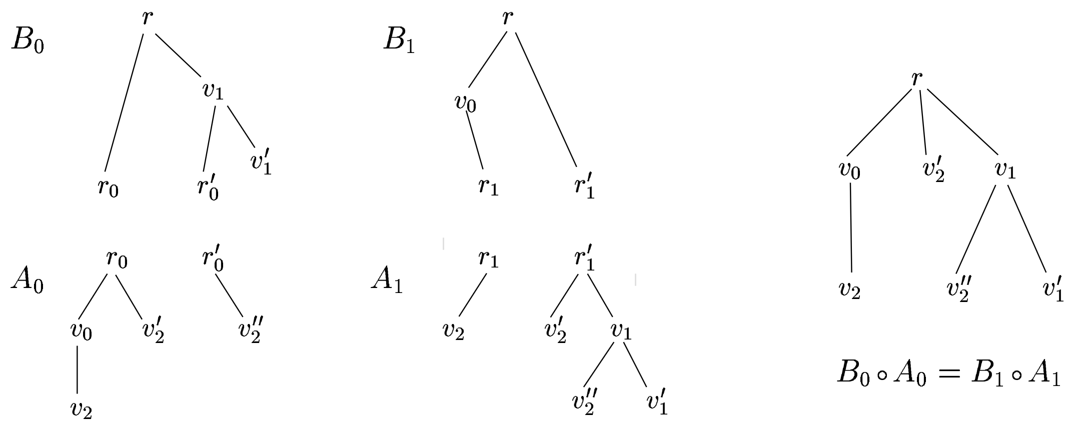


Figure 8: A consistent pair $\vec{A}$ of topographs, with IPO $\vec{B}$

Now the main content of Appendix A is to give, as Construction 96, a complete description of the set of IPOs for every pair $\vec{A}$ of topographs with domain 0; it then justifies the construction. The set is non-empty iff $\vec{A}$ is consistent. Here we repeat the construction less formally.

Construction 31 (IPOs in topographs) (This is a less formal version of Construction 96 from Appendix A.) First we construct the unique IPO for a pair $\vec{A}$ of epis satisfying the consistency conditions; then we show how this may be extended, non-uniquely, when the pair is non-epi – i.e. has barren roots.

Stage 1 To construct a unique IPO $\vec{B} : \vec{m} \to n$ for a pair $\vec{A} : 0 \to \vec{m}$ of epis. (Remarks in italics indicate how the consistency conditions validate the construction.)

roots The shared codomain n of $\vec{B}$ is constructed as follows: For each site $r_0 \in m_0$, unless $r_0 = A_0(v_2)$ for some v_2 deep in A_1, create a root $\widehat{r}_0 \in n$; similarly for each $r_1 \in m_1$. Equate $\widehat{r}_0$ and $\widehat{r}_1$ whenever $r_0 = A_0(v_2)$ and $r_1 = A_1(v_2)$ for some v_2.

nodes Take V_1 to be the nodes of B_0; similarly for B_1.

parents For sites: If $\widehat{r}_0$ exists, define $B_0(r_0) \triangleq \widehat{r}_0$; otherwise find v_2 deep in A_1 with $A_0(v_2) = r_0$ and define $B_0(r_0) \triangleq A_1(v_2)$. (*This will be in V_1 and independent of choice of v_2.*) For nodes: If v_1 is shallow in A_1, say $A_1(v_1) = r_1$, then define $B_0(v_1) \triangleq \widehat{r}_1$. (*This will exist.*) Otherwise define $B_0(v_1) \triangleq A_1(v_1)$. (*This will be in V_1.*) Similarly for B_1.

Stage 2 To extend the construction for each barren root in an arbitrary pair $\vec{A} : 0 \to \vec{m}$ of topographs.

barren roots If r_0 is barren in A_0 then *either* create a new root $\widehat{r}_0 \in n$ and define $B_0(r_0) \triangleq \widehat{r}_0$ (leaving $\widehat{r}_0$ barren in B_1), *or* define $B_0(r_0) \triangleq v_1$ for any non-atomic $v_1 \in V_1$. Similarly for B_1, if any r_1 is barren in A_1. ■

Example 7 (topograph IPO) Figure 8 shows an example of an IPO $\vec{B}$, for a pair $\vec{A}$ which we have already seen to be consistent in Example 6. The pair are epis, so Stage 1 of the construction applies. In forming the codomain of the IPO only one root $r = \widehat{r}_0 = \widehat{r}'_1$ is created. It is informative to check why $\widehat{r}'_0$ and $\widehat{r}_1$ are not created, and also to walk through the rest of the construction. Later we shall continue this example by adding an IPO for a pair of monographs with the same support, and then combining these into an IPO for a pair of bigraphs. ■

We now restate Theorem 97 from Appendix A:

Corollary 32 (IPOs in topographs) *A pair $\vec{B} : \vec{m} \to n$ is an IPO for $\vec{A} : h \to \vec{m}$ iff it is generated by Construction 31.*

Since the construction generates at least one IPO when $\vec{A}$ satisfies (C0) – (C3), the theorem also shows that these conditions are sufficient for consistency. As we shall see later, if G and F are bigraphs whose topograph components are $G^{\mathsf{T}} = A_0$ and $F^{\mathsf{T}} = B_0$, then any transition $G \xrightarrow{F}$ is defined in terms of a topograph IPO $\vec{B}$ for $\vec{A}$.

Note that if A_0 has a barren root then, provided B_0 has any non-atomic node, Stage 2 yields at least two distinct IPOs for the pair $\vec{A}$. This shows that in bigraphs IPOs –hence RPOs– exist where a pushout cannot exist (else it would be the unique IPO). This justifies the remark following Definition 6.

In the second alternative of Stage 2 we can think of the construction as 'losing' the barren root of A_0 inside a node of B_0. (By contrast, in the first alternative the root is preserved.) We shall return to this phenomenon, which we call *elision*, in Definition 78.

Finally, to ensure that our derived behavioural relations over bigraphs are congruential, we need the following corollary of Theorem 95, proved in Appendix A:

Corollary 33 (topographs always have RPOs) *Every pair $\vec{A}$ of topographs with a bound $\vec{E}$ in* **Top** *has an RPO w.r.t. $\vec{E}$.*

4.2 Monographs

In this subsection we develop the theory of monographs far enough to define their contribution to dynamics. Since equivalence relations play an important role we review them briefly here.

Equivalence relations: terminology and facts We write $R^{\equiv}$ for the smallest equivalence containing a relation R, and $\equiv_0 \sqcup \equiv_1$ for the smallest equivalence including $\equiv_0$ and $\equiv_1$. An equivalence $\equiv$ is *discrete* if $x \equiv y \Rightarrow x = y$. If $\equiv$ is over $X \uplus Y$ then its *restriction to* X is written $\equiv \restriction X$, meaning $\equiv \cap X^2$; its *restriction from* X is written $\equiv \setminus X$, meaning $\equiv \restriction Y$.

$\equiv_0 \sqcup \equiv_1$ is the least upper bound (lub) of $\equiv_0$ and $\equiv_1$ in the lattice of equivalence relations over a set X. We have that $(x, x') \in \equiv_0 \sqcup \equiv_1$ iff there is a sequence

$$x = x_0 \equiv_{i_1} x_1 \cdots \equiv_{i_n} x_n = x' \qquad (x_{j-1} \neq x_j)$$

where $n \geq 0$, $i_j \in \{0, 1\}$ and $i_{j-1} \neq i_j$ $(1 < j \leq n)$. More generally, (x, x') is an instance of $\equiv_1 \sqcup \cdots \sqcup \equiv_k$ iff there is a chain

$$x = x_0 \equiv_{i_1} x_1 \cdots \equiv_{i_n} x_n = x' \qquad (x_{j-1} \neq x_j)$$

where $i_j \in \{1, \ldots k\}$ and $i_{j-1} \neq i_j$ $(1 < j \leq n)$. We call this a *lub chain for* $\{\equiv_1 \mid 1 \leq i \leq k\}$. (If the conditions $x_{j-1} \neq x_j$ and $i_{j-1} \neq i_j$ do not hold, the chain can readily be collapsed to a shorter one where they do hold.)

We shall often have to combine equivalences over distinct but overlapping sets. For example if $\equiv_0$ is over $X_0 \uplus Y$ and $\equiv_1$ is over $X_1 \uplus Y$, where X_0 and X_1 are disjoint, then $\equiv_0 \sqcup \equiv_1$ is over $X_0 \uplus Y \uplus X_1$.

Formally, we may regard this as first extending $\equiv_0$ to an equivalence over $X_0 \uplus Y \uplus X_1$ by making $x \equiv_0 x$ for each $x \in X_1$, similarly extending $\equiv_1$, and then combining them. But it is perfectly correct to make the combination (by the above prescription) without this extension. We can use the notion of *support* of an equivalence to reduce its domain before combination; if $\equiv$ is over X, then its support consists of those members $x \in X$ whose equivalence class $[x]_\equiv$ is not a singleton.

A common case of restriction is $(\equiv_0 \sqcup \equiv_1)\setminus Y$, where $\equiv_0$ is over $X_0 \uplus Y$ and $\equiv_1$ is over $X_1 \uplus Y$. A common manipulation is that if the support of $\equiv_2$ is disjoint from Y then

$$(\equiv_0 \sqcup \equiv_1 \sqcup \equiv_2)\setminus Y \;=\; (\equiv_0 \sqcup \equiv_1)\setminus Y \sqcup \equiv_2 \;\;.$$

We are now ready to define monographs.

Definition 34 (monograph) A *monograph* $A = (V, ctrl, \equiv) : X \to Y$ has finite[6] sets X of *conames* and Y of *names*; a set V of nodes and a control function $ctrl$; and an equivalence $\equiv$ upon the set $X + P + Y$ of *ports*, where the set P of *inner* ports is the disjoint sum of node arities, i.e. $P \triangleq \sum_{v \in V} ar(ctrl(v))$. ■

We shall consistently use corresponding lower case letters to denote members for the port-sets, $x \in X$, $p \in P$ and so on, to avoid repeated qualifications such as '$x \in X$'. This convention will also allow us to write, for example, $x \equiv_A p$ instead of $\langle 0, x\rangle \equiv_A \langle 1, p\rangle$. Since the conames X and the names Y may not be disjoint we shall often write $\underline{x}$ for a coname x. We shall often abuse notation even further by writing A instead of $\equiv_A$ for the equivalence relation, and $[r]_A$ for the $\equiv_A$-equivalence class of any port r. So for example xAp means $\langle 0, x\rangle \equiv_A \langle 1, p\rangle$. Furthermore if $B = (Q, \equiv_B) : Y \to Z$ is another monograph sharing an interface Y with A, we may write e.g. $pAxBq$ to mean pAx & xBq.

Definition 35 (precategory of monographs) The precategory **Mog** has name sets as objects and monographs as arrows. The composition $B \circ A : X \to Z$ of two monographs $A = (V_A, ctrl_A, \equiv_A) : X \to Y$ and $B = (V_B, ctrl_B, \equiv_B) : Y \to Z$ is defined when their node sets are disjoint; its node set and control function are then $V_A \uplus V_B$ and $ctrl_A \uplus ctrl_B$, and its equivalence[7] is $(\equiv_B \sqcup \equiv_A)\setminus Y$. The identity on X is $\mathsf{id}_X : X \to X$ with no nodes and the equivalence $\{(\underline{x}, x) \mid x \in X\}^{\equiv}$. ■

It is routine to establish the identity and associativity properties of composition.

From the last two definitions it is quite clear that the structural character of monographs, and of their precategory, lies almost entirely in the equivalences over ports. Since the theory of monographs – especially their RPO theory – is technically complex, we wish to present it in the simplest possible medium. This will be in a precategory which is the image of a forgetful functor on **Mog**, as follows:

Definition 36 (edge net) An *edge net* $E = (P, \equiv) : X \to Y$ has finite sets X of *conames* and Y of *names*; a set P of inner ports; and an equivalence $\equiv$ upon the set $X + P + Y$ of *ports*. ■

[6] The reader may ask why we constrain the interface sets to be finite, since this may be inconvenient when representing (say) the π-calculus. The answer is simple; RPOs do not exist for monographs without this constraint. But the inconvenience is not serious. The outer name set of an agent a is understood to contain all the names which it 'uses' – i.e. which are not idle in a (see Definition 39 below), and we are only interested in finitely presented agents, which can therefore only 'use' finitely many names. Moreover, in Section 6 we shall show that two agents with a given name set are bisimilar iff they are so when provided with a larger name set.

[7] This definition is slightly inaccurate, for the sake of brevity. Since the name sets may not always be disjoint, the fully accurate definition of $\equiv_{B \circ A}$ is as follows: Let P_A and P_B be the two sets of inner ports; take the lub of the two equivalences on $X + P_A + Y + P_B + Z$ induced by $\equiv_A$ and $\equiv_B$; then take the equivalence thus induced on $X + (P_A \uplus P_B) + Z$.

Definition 37 (precategory of edge nets) The precategory **Edg** has name sets as objects and edge nets as arrows. The composition $F \circ E : X \to Z$ of two edge nets $E : X \to Y$ and $F : Y \to Z$ is defined when their sets P and Q of inner ports are disjoint; its inner port set is then $P \uplus Q$ and its equivalence is $(\equiv_F \sqcup \equiv_E)\backslash Y$. The identity on X is $\mathsf{id}_X : X \to X$ with no inner ports and the equivalence $\{(\underline{x}, x) \mid x \in X\}^{\equiv}$.

The forgetful functor $\mathcal{E} : \mathbf{Mog} \to \mathbf{Edg}$ is the identity on interfaces; on a context $A = (V, ctrl, \equiv) : X \to Y$ it yields the edge net $E = (P, \equiv) : X \to Y$, where P is the set of inner ports of A. ■

It is a routine matter to check that $\mathcal{E}$ is indeed a functor. In what follows we shall establish a few properties of both monographs and edge nets, but the main work of establishing RPOs and characterising IPOs will be done for edge nets and then (easily) lifted through $\mathcal{E}$ to monographs. With this approach, we have succeeded in basing the theory of bigraphs upon two well-studied mathematical structures; trees (for topographs) and equivalences (for edge nets). Technically we could avoid monographs in much of this development, letting edge nets play their role; but in applying the theory it is important to be able to use the topograph-monograph combination.

The following few properties apply equally to monographs and to their edge nets. We use the term 'support' to mean the node set in the former case and the set of inner ports in the latter case.

Proposition 38 (isomorphisms in monographs) *A context $\iota : X \to Y$ in* **Mog** *or* **Edg** *is an isomorphism iff it has empty support and its equivalence is a bijection from X to Y.*

Definition 39 (properties of ports) A name $x \in X$, the name set of an monograph or edge net A, is an *alias for $x' \in X$ in A* if $x \neq x'$ and xAx'. x is *idle* in A if $rAx \Rightarrow r \in X$ for any port r in A. A port r is *open* in A if rAx for some $x \in X$; otherwise it is *closed* in A. There are obvious dual concepts *co-alias*, *co-idle* etc. for the conames of A. ■

We have met aliases in the examples of Section 2; they played a role in the examples on arithmetic and security, where in each case the reaction rule creates an alias in the reactum. They have no analogue for topographs. Although they appear to be necessary for certain calculi, we shall here find it useful to study bigraphs without aliases (though we keep co-aliases, which play a role akin to substitutions). To this end, we now define

Definition 40 (alias-free) Call a monograph *alias-free* if it contains no aliases. Denote by $\mathbf{Mog}_{\bar{\mathsf{a}}}$ and $\mathbf{Edg}_{\bar{\mathsf{a}}}$ the subprecategories of **Mog** and **Edg** respectively which include only *alias-free* contexts. It is easily shown that alias-freedom is preserved by composition. ■

We now give a simple characterisation of epimorphisms in both these precategories.

Proposition 41 (epimorphisms in monographs) *A context $A : W \to X$ in* **Mog** *or* **Edg** *is an epimorphism (epi) iff it has no idle names and no aliases. Similarly a context in* $\mathbf{Mog}_{\bar{\mathsf{a}}}$ *or* $\mathbf{Edg}_{\bar{\mathsf{a}}}$ *is epi iff it has no idle names.*

Proof We need only prove the first; the argument applies equally to monographs and edge nets.

($\Rightarrow$) Suppose A has an idle name $x \in X$. Then pick two contexts with empty node sets: $B_0 : X \to \{y\}$ with the identity equivalence, and $B_1 : X \to \{y\}$ with the equivalence $\{(x, y)\}^{\equiv}$; clearly $B_0 \neq B_1$ but $B_0 \circ A =$

$B_1 \circ A$. On the other hand, suppose $x, x' \in X$ are distinct, with xAx'; then pick $B_i := X \to \emptyset)$ $(i = 0, 1)$ with empty node sets, where B_0 has the identity equivalence and B_1 has the equivalence $\{(x, x')\}^{\equiv}$.

($\Leftarrow$) Assume that A, with inner port set P, has no idle names and no aliases, and let $B_0 \neq B_1$. We shall prove $B_0 \circ A \neq B_1 \circ A$. Clearly if B_0, B_1 have different node sets or interfaces then also $B_0 \circ A \neq B_1 \circ A$. Otherwise we have $B_i : X \to Y$ $(i = 0, 1)$ both with inner port set Q say, with three possibilities:

1. For some $x, x' \in X$, we have, say, xB_0x' but not xB_1x'. Since x, x' are not idle, there are $p, p' \in W \uplus P$ with $pAx, p'Ax'$. Then $p(B_0 \circ A)p'$, but not $p(B_1 \circ A)p'$ since A has no aliases.
2. For some $x \in X$ and $q \in Q \uplus Y$, say xB_0q but not xB_1q. Then similarly for some pAx we have $p(B_0 \circ A)q$ but not $p(B_1 \circ A)q$ (again, the no-alias condition is needed here).
3. Otherwise, there must be $q, q' \in Q \uplus Y$ co-closed (i.e. not related to X) in B_0 and B_1, with say qB_0q' but not qB_1q'. Then clearly $q(B_0 \circ A)q'$ but not $q(B_1 \circ A)q'$. ■

We now turn to the definition of tensor product in monographs. Unlike the case of topographs, where $A \otimes B$ is defined to exist iff A and B have disjoint support, in monographs we also have to require disjointness of the name sets representing their domains and codomains respectively.

Definition 42 (tensor product) For two monographs $A_i : X_i \to Y_i$ $(i = 0, 1)$ in **Mog** we take their *tensor product* $A_0 \otimes A_1 : X_0 \uplus X_1 \to Y_0 \uplus Y_1$ to be defined when X_0 and X_1 are disjoint, Y_0 and Y_1 are disjoint, and A_0 and A_1 have disjoint node sets. The equivalence on the ports of the product is then just the union of the equivalences for A_0 and A_1 respectively (recalling that the port sets are disjoint). ■

Again, as for topographs, it is routine to check that this definition satisfies the requirements of Definition 3.

For the remainder of this subsection we shall consider an arbitrary pair $\vec{A} : W \to \vec{X}$ of monographs or edge nets. We first make a simplifying assumption. In this all the work of this subsection, the conames W which are the common domain of A_0 and A_1 are treated exactly as inner ports common to both contexts; so w.l.o.g. we shall assume W to be empty. Now the inner ports of $\vec{A}$ may be partitioned into three sets: P_i occurring in A_i alone $(i = 0, 1)$ and P_2 shared between A_0 and A_1. (Of course, for monographs this partition is induced by a partition of the node sets.) We shall regularly use $p_i, p'_i, p''_i, \ldots$ $(i = 0, 1, 2)$ to range over these sets, and x_i over X_i $(i = 0, 1)$. Since we are considering the domain W to be empty, the reader may like to note that it comes to the same thing if we consider $p_2, p'_2, \ldots$ as ranging over $P_2 \uplus W$.

As with **Top**, we shall be concerned with consistency. The following pair of conditions will turn out to be necessary and sufficient for consistency of a pair of monographs or edge nets:

Definition 43 (consistency conditions) We define two conditions on a pair $\vec{A} : \emptyset \to \vec{X}$ of monographs or edge nets (recalling that $\bar{\imath}$ means $1 - i$):

$$\text{(O1)} \quad p_iA_ip_2 \Rightarrow p_2 \text{ open in } A_{\bar{\imath}}$$
$$\text{(O2)} \quad p_2A_ip'_2 \Rightarrow (p_2A_{\bar{\imath}}p'_2 \text{ or } p_2 \text{ open in } A_{\bar{\imath}}) \; .$$
■

These conditions are somewhat easier than those for topographs, but let us state them informally, in words rather than in symbols. Talk of A_0, A_1 as 'left' and 'right', and of equivalent ports as 'peers'; so for example a port q is a left peer of a shared port p_2 if qA_0p_2, and so on. The conditions declare the following (and also with 'left' and 'right' exchanged) for any shared port p_2:

(O1) If p_2 has an unshared inner port as a left peer then it is open on the right.

(O2) If p_2 has a shared left peer, then on the right either they are peers or they are both open.

Example 8 (consistent monographs) Consider the pair $\vec{A} : \emptyset \to \vec{X}$ of monographs in Figure 9, where $X_0 = \{x_0, y_0, z_0\}$ and $X_1 = \{x_0, y_0\}$. As in Example 6, controls are not shown and the nodes with subscript 2 are shared. The pair is consistent. But if z_0 is removed from A_0, or x_1 from A_1 (i.e. their edges are closed) then (O1) is violated; similarly if x_0 or y_0 is removed then (O2) is violated. ■

We now wish to characterise the IPOs in $\mathbf{Mog}_{\bar{a}}$, and also state conditions under which RPOs exist. In doing so, we shall summarise the results of Appendices B and C, which are proved for edge nets; we therefore rely on the following, which is easy to prove by standard categorical reasoning:

Proposition 44 (moving RPOs from edge nets to topographs) *In* $\mathbf{Mog}_{\bar{a}}$, *let* $(\vec{B}, B)$ *be an RPO candidate for* $\vec{A}$ *w.r.t.* $\vec{E}$. *Then in* $\mathbf{Edg}_{\bar{a}}$, $\mathcal{E}(\vec{B}, B)$ *is an RPO candidate for* $\mathcal{E}(\vec{A})$ *w.r.t.* $\mathcal{E}(\vec{E})$. *Moreover, the former is an RPO in* $\mathbf{Mog}_{\bar{a}}$ *iff the latter is an RPO in* $\mathbf{Edg}_{\bar{a}}$.

The main content of Appendix B is to give, in Constructions 98 and 105, a complete description of a non-empty set of IPOs for every consistent pair $\vec{A}$; it then justifies the construction. As for topographs, we shall repeat the construction informally.

Construction 45 (IPOs in monographs) (This is a less formal version of Constructions 98 and 105 from Appendix B.)

For any pair $\vec{A}$ of monographs satisfying (O1) – (O2), we characterise their IPOs in two stages. First we construct a bound, which also yields the unique IPO when $\vec{A}$ are epis. We then describe how to extend the construction, non-uniquely, to non-epis.

Stage 1 To construct a unique IPO $\vec{B} : \vec{X} \to Y$ for a pair $\vec{A} : \emptyset \to \vec{X}$ of epi monographs.

- **codomain** The shared codomain Y is constructed as follows: For each coname $x_0 \in X_0$, unless $x_0 A_0 p_2$ for some p_2 closed in A_1, create a name $\widehat{x_0} \in Y$; similarly for each $x_1 \in X_1$. Equate $\widehat{x}_0$ and $\widehat{x_1}$ whenever $x_0 A_0 p_2 A_1 x_1$ for some p_2 .
- **inner ports** Take the inner ports of B_0 to be P_1 (similarly for B_1). This determines the nodes of B_0.
- **edges** Generate the equivalence B_0 from the following (similarly for B_1):

$$\begin{array}{lll} x_0 B_0 x_0' & \text{if } x_0 A_0 p_2 A_1 p_2' A_0 x_0' & \qquad p_1 B_0 p_1' \quad \text{if } p_1 A_1 p_1' \\ x_0 B_0 p_1 & \text{if } x_0 A_0 p_2 A_1 p_1 & \qquad p_1 B_0 \widehat{x}_1 \quad \text{if } p_1 A_1 x_1 \; . \\ x_0 B_0 \widehat{x}_0 & & \end{array}$$

Stage 2 To extend the construction for each idle name in an arbitrary pair $\vec{A} : \emptyset \to \vec{X}$ of monographs.

- **idle names** If x_0 is idle in A_0 then *either* create a new name $\widehat{x}_0 \in Y$ and set $x_0 B_0 \widehat{x_0}$ (leaving $\widehat{x}_0$ idle in B_1), *or* set $x_0 B_0 q$ for any port q closed in B_0. Similarly for B_1. (Note that q may be either an inner port p_1 or another $x_0' \in X_0$.) ■

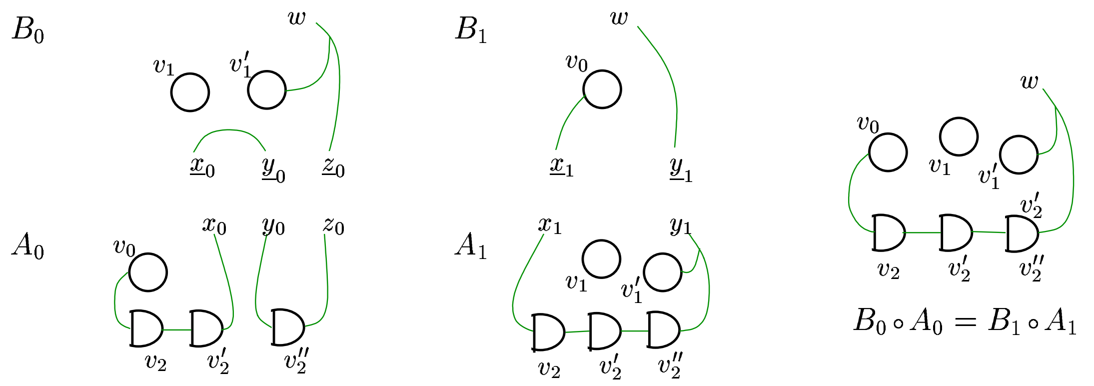


Figure 9: A consistent pair $\vec{A}$ of monographs, with IPO $\vec{B}$

Example 9 (monograph IPO) Figure 9 shows an IPO $\vec{B} : \vec{X} \to W$ for pair $\vec{A}$ of monographs which we already found to be consistent in Example 8. As with the topograph example if Figure 8, the pair $\vec{A}$ are epis, so Stage 1 of the construction applies. In forming the codomain W of the IPO we find that only $\widehat{z}_0$ and $\widehat{y}_1$ are created; moreover they are equated (as w) because they are linked to the same port of a shared node. Thus $W = \{w\}$, a singleton. Note particularly the co-alias $\underline{x}_0 = \underline{y}_0$ in B_0; it arises because x_0 and y_0 are connected to shared inner ports which are linked in A_1. It is a good exercise to check how the construction generates the other edges in B_0 and B_1. ∎

The construction is justified by Theorem 108 from Appendix B, which we restate here as a corollary:

Corollary 46 (consistency and IPOs in monographs) *The IPOs for any pair of monographs in* $\mathbf{Mog}_{\bar{\mathsf{a}}}$ *are, up to isomorphism, exactly those produced by Construction 45.*

Since the construction generates at least one IPO when $\vec{A}$ satisfies (O1) – (O2), the theorem also shows that these conditions are sufficient for consistency. Just as for topographs, we shall later find that if G and F are bigraphs whose monograph components are $G^{\mathsf{M}} = A_0$ and $F^{\mathsf{M}} = B_0$, then any transition $G \xrightarrow{F} \triangleright$ is defined in terms of an monograph IPO $\vec{B}$ for $\vec{A}$.

Note that if A_0 has an idle name then, provided B_0 has any closed port, Stage 2 yields at least two distinct IPOs for the pair $\vec{A}$. Like the IPO construction for topographs, this shows that in bigraphs IPOs –hence RPOs– exist where a pushout cannot exist (else it would be the unique IPO), justifying the remark following Definition 6.

As for topographs, the second alternative in Stage 2 represents a kind of *elision*, to be defined in Definition 78. We may think of the construction as 'losing' an idle name of A_0 somewhere within B_0.

As for topographs, we need to know that RPOs exist sufficiently often for monographs. The fact that every consistent pair $\vec{A}$ of monographs has at least one IPO does not imply that an RPO always exists for the pair w.r.t. any bound $\vec{E}$. And indeed, it does not always exist! A counter-example appears in Appendix C. But we can show that it exists often enough for our purposes. The proof – for edge nets – is given as Theorem 126 in the same appendix; with the help of Proposition 44 we can deduce the corresponding result for monographs:

Corollary 47 (a monograph pair with one epi always has an RPO) *In* $\mathbf{Mog}_{\bar{\mathsf{a}}}$*, every pair* $\vec{A}$ *of monographs of which at least one is an epimorphism has an RPO w.r.t. any bound* $\vec{E}$*.*

We shall derive a corresponding result for bigraphical reactive systems in the next subsection.

4.3 Bigraphs

We now resume the discussion of bigraphs.

Definition 48 (precategory of bigraphs) The precategories **Big**($\mathcal{K}$) and **Big**$_{\bar{\mathsf{a}}}$($\mathcal{K}$) of bigraphs over a signature $\mathcal{K}$ have pairs $I = \langle m, X \rangle$ as objects (*interfaces*) and bigraphs $G : (U, ctrl_G, G^{\mathsf{T}}, G^{\mathsf{M}}) : I \to J$ as arrows (*contexts*), where G^{M} has no aliases in the case of **Big**$_{\bar{\mathsf{a}}}$. If $H : J \to K$ is another bigraph with node set disjoint from U, then their composition is defined directly in terms of the composition of the components as follows:

$$H \circ G \ \triangleq\ \langle H^{\mathsf{T}} \circ G^{\mathsf{T}}, H^{\mathsf{M}} \circ G^{\mathsf{M}} \rangle : I \to K\ .$$

The identities are $\langle \mathsf{id}_m, \mathsf{id}_X \rangle : I \to I$, where $I = \langle m, X \rangle$. ■

We shall continue to omit the signature $\mathcal{K}$ except when we are making a formal definition. We now combine some familiar topograph and monograph structures to yield bigraph structures.

Proposition 49 (isomorphisms in bigraphs) *The isomorphisms in* **Big** *and* **Big**$_{\bar{\mathsf{a}}}$ *are all combinations* $\iota = \langle \iota^{\mathsf{T}}, \iota^{\mathsf{M}} \rangle$ *of a topograph isomorphism and a monograph isomorphism.*

Definition 50 (tensor product) The *tensor product* of two bigraph interfaces is defined by $\langle m, X \rangle \otimes \langle n, Y \rangle \triangleq \langle m + n, X \uplus Y \rangle$ when X and Y are disjoint. The *tensor product* of two bigraphs $G_i : I_i \to J_i$ $(i = 0, 1)$ is defined by

$$G_0 \otimes G_1 \triangleq \langle G_0^{\mathsf{T}} \otimes G_1^{\mathsf{T}}, G_0^{\mathsf{M}} \otimes G_1^{\mathsf{M}} \rangle : I_0 \otimes I_1 \to J_0 \otimes J_1$$

when the interfaces exist and the node sets are disjoint. This combination is well-formed, since its components share the same node set. ■

We can now assert that

Proposition 51 (wide precategories of bigraphs) **Big** *and* **Big**$_{\bar{\mathsf{a}}}$ *are monoidal wide precategories.*

Proof First, they are well-supported, with node sets as support. Second, they are monoidal (Definition 3) with a naming function yielding the name set of each interface. Third, to make them wide precategories (Definition 10) we identify the required ingredients as follows:

(1) The origin is $\epsilon \triangleq \langle 0, \emptyset \rangle$.

(2) The width functor is given on objects by $\mathsf{width}\langle m, X \rangle \triangleq m$. On each arrow $C : \langle m, X \rangle \to \langle n, Y \rangle$, for each site $s \in m$ we define $\mathsf{width}(C)(s)$ to be the unique root $r \in n$ is such that $r >_C s$.

(3) The isomorphisms from $I = \langle m, X \rangle$ to itself include $\pi_I \triangleq \langle \pi, \mathsf{id}_X \rangle$, where $\pi : m \to m$ is any isomorphism on m. ■

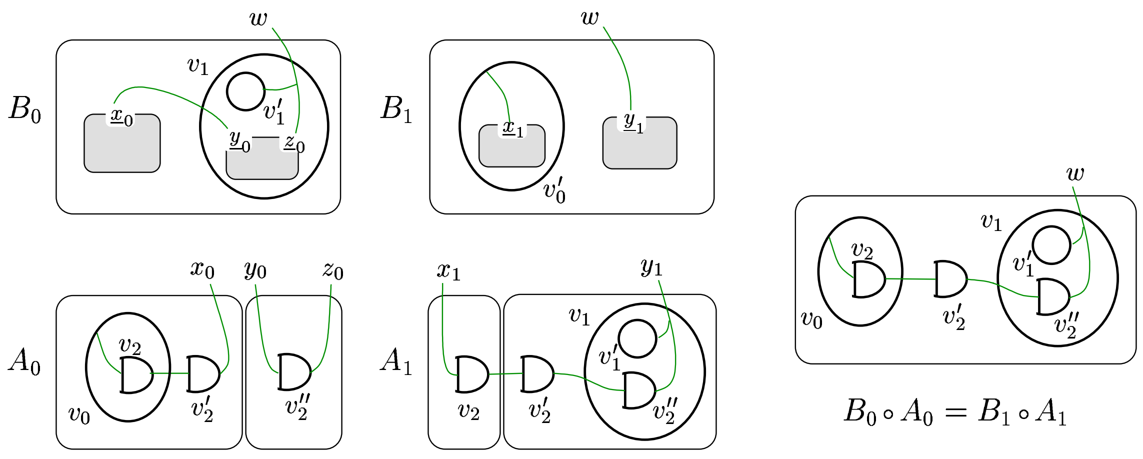


Figure 10: A consistent pair $\vec{A}$ of bigraphs, with IPO $\vec{B}$

The origin ϵ is neither an initial nor a terminal object. However id_ϵ is the unique arrow from the origin to itself, and we shall denote it also by ϵ.

Proposition 52 (epimorphisms in bigraphs) *A bigraph G in* $\mathbf{Big}_{\bar{\mathsf{a}}}$ *is an epimorphism iff its components* G^{T} *and* G^{M} *are epimorphisms in* **Top** *and* $\mathbf{Mog}_{\bar{\mathsf{a}}}$ *respectively.*

For RPOs and IPOs, the corresponding result depends upon the fact that the RPO and IPO constructions treat the node sets of **Top** and $\mathbf{Mog}_{\bar{\mathsf{a}}}$ consistently.

Proposition 53 (RPOs in bigraphs) *A triple* $(\vec{B}, B)$ *is an RPO for* $\vec{A}$ *w.r.t.* $\vec{E}$ *in* $\mathbf{Big}_{\bar{\mathsf{a}}}$ *iff both*

- $(\vec{B}^{\mathsf{T}}, B^{\mathsf{T}})$ *is an RPO for* $\vec{A}^{\mathsf{T}}$ *w.r.t.* $\vec{E}^{\mathsf{T}}$ *in* **Top**, *and*
- $(\vec{B}^{\mathsf{M}}, B^{\mathsf{M}})$ *is an RPO for* $\vec{A}^{\mathsf{M}}$ *w.r.t.* $\vec{E}^{\mathsf{M}}$ *in* $\mathbf{Mog}_{\bar{\mathsf{a}}}$.

Thus, in particular, we get every IPO in bigraphs by combining a topograph IPO and a monograph IPO.

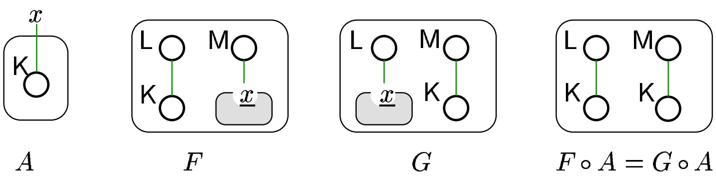
Example 10 (Bigraph IPO) To illustrate this, we can now combine the examples of IPOs in Example 7 for topographs and Example 9 for monographs, since they have the same support. They appear in Figures 8 and 9, and we show their combination in Figure 10. Comparing the three figures, it seems that the separation of a bigraph into its parts is a valuable aid to understanding; on the other hand the composite bigraph diagrams seem better for the presentation of reaction rules, in the way that we used them in Section 2. ■

We now have everything we need for the definition of bigraphical reactive systems in the next section. To end the present section, we shall now show why we have to base that definition on bigraphs in $\mathbf{Big}_{\bar{\mathsf{a}}}$, rather than upon the *support quotient* of $\mathbf{Big}_{\bar{\mathsf{a}}}$, as defined in Definition 4.

The difference between a well-supported precategory and its support quotient category is that, in the former, operations such as composition keep track of support elements, while in the latter they do not. For bigraphs, this means that in $\mathbf{Big}_{\bar{\mathsf{a}}}$ we keep track of the identity of nodes, while in its support quotient we do not; in a familiar terminology, we deal with *concrete graphs* in $\mathbf{Big}_{\bar{\mathsf{a}}}$ and *abstract graphs* in its quotient. As promised in Subsection 3.3, we shall now show by examples that in passing to the support quotient of $\mathbf{Big}_{\bar{\mathsf{a}}}$ we lose structural properties that are essential for deriving transition systems.

In these examples we shall work with diagrams. The only difference between a faithful diagram of a concrete bigraph G and the abstract bigraph $[G]$ is that the former should make explicit the names $(u, v, \ldots)$ of nodes, while the latter should not. In all the examples of IPOs in this section we have respected this convention in our diagrams; indeed, the notions of consistency and IPO have depended crucially on node identity. However, this does not rule out the possibility that RPOs and IPOs, defined in some other way, may exist in the support quotient category. We shall now show definitively that they do not exist, in very simple cases.

Our first example shows that epis are not preserved by support quotient.



Example 11 (support quotient loses epis) The diagram shows an abstract bigraph A, the image of an epi in $\mathbf{Big}_{\bar{\mathsf{a}}}$. Two distinct (abstract) contexts F and G are shown, such that the abstract bigraphs $F \circ A$ and $G \circ A$ coincide (because this bigraph, being abstract, does not record where each K-node came from). Thus A is not an epi in the support quotient of $\mathbf{Big}_{\bar{\mathsf{a}}}$. ∎

A more crucial loss of structure occurs with RPOs, as follows.

Example 12 (support quotient loses RPOs) Let A be as in Example 11. Since any concrete representative of A is an epi in $\mathbf{Big}_{\bar{\mathsf{a}}}$, any concrete representative of the pair A, A will have an RPO in $\mathbf{Big}_{\bar{\mathsf{a}}}$ w.r.t. any bound. Moreover, this RPO will differ according to whether the two K-nodes in the representatives of A have the same or different identities. Figure 11 shows two candidate RPOs for the abstract pair A, A w.r.t. the abstract pair E, E; these two candidates are in fact the images of the two different concrete RPOs mentioned above. The first candidate is the triple $(\mathsf{id}, \mathsf{id}, E)$, while the second is (C_0, C_1, C) as shown.

Now if an RPO (D_0, D_1, D) exists for A, A w.r.t. E, E then there must exist abstract bigraphs F and G that mediate to the two candidates, making the diagram commute. But this leads to a contradiction, as follows. First, D_0 and D_1 have empty support, since for example $F \circ D_0 = \mathsf{id}$. From $D_0 \circ A = D_1 \circ A$ it can then be deduced that $D_0 = D_1$.[8] It follows that $C_0 = G \circ D_0 = G \circ D_1 = C_1$, a contradiction. ∎

Thus, to derive transition systems we must use concrete bigraphs. However, once derived, these systems and their congruence theorems can be transferred to abstract bigraphs as shown in Propositions 16 and 18.

[8]This would be easy if A were epi, but from Example 11 it is not! However, one can use the following lemma about support quotients which can be verified from the definitions in Subsection 3.1: Let f be epi, with singleton support, and let g, h have empty support. Then $g \circ f \simeq h \circ f$ implies $g = h$.

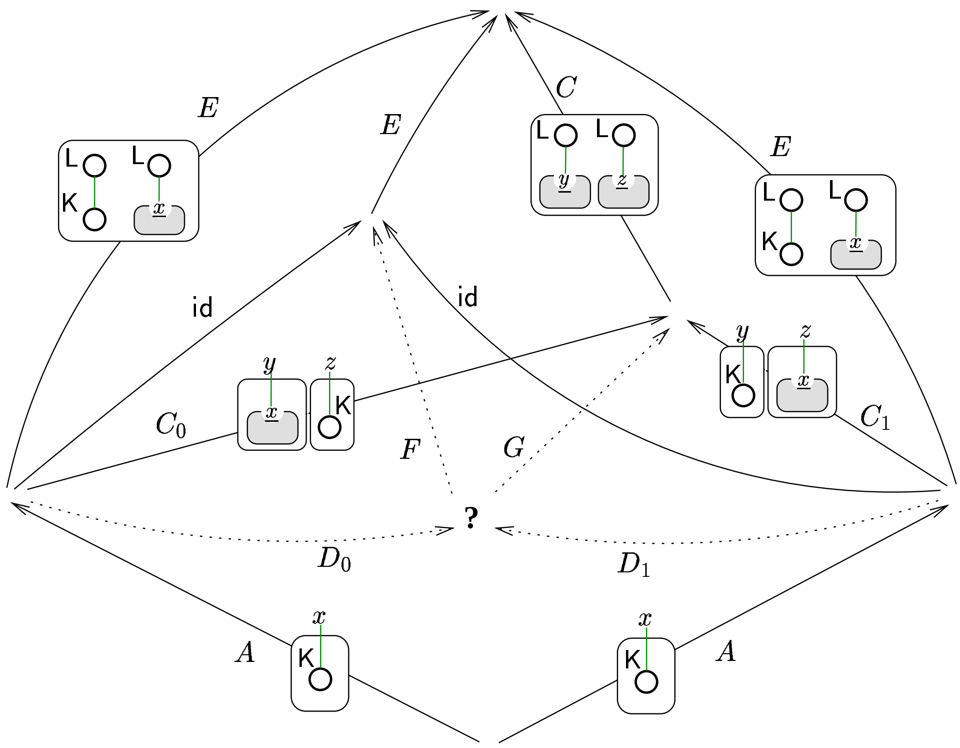


Figure 11: Two abstract bigraphs may lack an RPO

5 Bigraphical reactive systems

We now give the central definition of the paper.

Definition 54 (bigraphical reactive system (BRS)) In **Big**($\mathcal{K}$) let reacts be a set of triples (r, r', λ), where r, r' are agents with interface I and $\lambda \in \mathsf{plac}(I)$. Then the *bigraphical reactive system (BRS)* **Big**($\mathcal{K}, \mathsf{reacts}$) consists of the monoidal wide reactive system (Definition 12) over **Big**($\mathcal{K}$) whose reaction rules are reacts, and whose activity map act is defined for each bigraph $G : I \to J$ by the activity of its topograph; that is,

$$\mathsf{act}(G) \triangleq \{i \in \mathsf{width}(I) \mid G^{\mathsf{T}} \text{ active at } i\} \; .$$

Big$_{\bar{\mathsf{a}}}$($\mathcal{K}, \mathsf{reacts}$) is defined similarly. ■

This definition is justified as follows:

Proposition 55 (a bigraphical system is a WRS) *The definition of bigraphical systems satisfies the conditions required for a monoidal WRS.*

Proof We already know from Proposition 51 that **Big** and **Big**$_{\bar{\mathsf{a}}}$ are wide precategories. It remains to show that the activity maps defined above satisfy the two conditions of Definition 12. This only concerns the topograph of a bigraph, and the details are routine. ■

5.1 Wide bisimilarity

With this definition, we are now able to apply the derivation of placed transitions (Subsection 3.4) to any BRS. So we can deduce from Theorem 20 a congruence property for wide bisimilarity in a BRS with sufficient RPOs. The existence of RPOs in **Big** is still an open question, but in $\mathbf{Big}_{\bar{\mathsf{a}}}$ we are better off; we have them for any pair $\vec{A}$ provided the monograph A_1^{M} is epi – i.e. has no idle names – as a consequence of the work of the previous section on RPOs for topographs, monographs and bigraphs (Corollaries 33,47 and Proposition 53). So we can assert:

Theorem 56 (congruence of bisimilarity in BRSs) *Let* reacts *be a set of reaction rules for* $\mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K})$ *in which every redex is an epimorphism. Then* $\mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K},\mathsf{reacts})$ *has all redex-RPOs. Hence wide bisimilarity of agents in* $\mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K},\mathsf{reacts})$ *is a congruence.*

How awkward is it to require redexes to be epis? It can be argued (see Chapter 7 in Leifer [18]) that a reaction rule whose redex has an idle name leads to rather strange behaviour, unlikely to be met in applications; we tend to regard such rules as unreasonable. The same can be argued for redexes with barren roots. So, even though RPOs exist always for topographs, we are happy to limit our attention to those cases of $\mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K},\mathsf{reacts})$ in which every redex is epi – i.e. has neither barren roots nor idle names.

This result –together with our expectation of similar congruence properties for other equivalences and preorders– tells us that our derived transition systems are worthy of attention. Of course, it does not tell us how easy they are to work with; nor does it tell us how large the congruence relation is. For these purposes we must rely on the characterisation of IPOs given in the preceding section, to yield in turn a characterisation of transitions and hence of bisimilarity itself. We begin this work in Section 6. Meanwhile, in the rest of the present section we investigate how this theory –hitherto specific to bigraphs– can be transferred to other WRSs.

5.2 WRS functors and sub-WRSs

When working in a WRS over **A** we may often be interested only in agents and contexts which lie in a subprecategory of **A**. More generally, we may be interested in $\mathbf{A}'$ for which a functor $\mathcal{F} : \mathbf{A}' \to \mathbf{A}$ exists. So we extend the notion of functor to WRSs as follows:

Definition 57 (WRS functor, sub-WRS) Let $\mathrm{WR}_{\mathbf{A}}$ and $\mathrm{WR}_{\mathbf{A}'}$ be wide reactive systems, with components respectively $\epsilon, \mathsf{width}, \mathsf{reacts}, \mathsf{act}$ and $\epsilon', \mathsf{width}', \mathsf{reacts}', \mathsf{act}'$. A functor $\mathcal{F} : \mathbf{A}' \to \mathbf{A}$ of precategories is then a *WRS functor* from $\mathrm{WR}_{\mathbf{A}'}$ to $\mathrm{WR}_{\mathbf{A}}$ if it preserves all these components; that is:

$$\begin{aligned}&\mathcal{F}(\epsilon') = \epsilon \; ;\\ &\mathsf{width} \circ \mathcal{F} = \mathsf{width}' \; ;\\ &(r, r', \lambda) \in \mathsf{reacts}' \Rightarrow (\mathcal{F}(r), \mathcal{F}(r'), \lambda) \in \mathsf{reacts} \; ; \text{ and}\\ &\mathsf{act}'(C) \subseteq \mathsf{act}(\mathcal{F}(C)) \; .\end{aligned}$$

Call $\mathcal{F}$ *monoidal* if both $\mathbf{A}'$ and $\mathbf{A}$ are monoidal and $\mathcal{F}$ preserves tensor product. If $\mathcal{F}$ is a (monoidal) inclusion functor then we call $\mathbf{A}'$ a (*monoidal*) *sub-WRS* of **A**. ■

We shall be interested in WRS functors which treat redex-RPOs with respect, in the following sense:

Definition 58 (functor creating redex-RPOs) Let $\mathcal{F} : \mathrm{WR}_{\mathbf{A}'} \to \mathrm{WR}_{\mathbf{A}}$ be a WRS functor. Then $\mathcal{F}$ is said to *create redex-RPOs (in* $\mathrm{WR}_{\mathbf{A}'}$*)* if the following holds for every reaction rule in $\mathrm{WR}_{\mathbf{A}'}$ with redex r' and place λ': Let the pair a', r' have bound G', E' in $\mathrm{WR}_{\mathbf{A}'}$, where E' is active at λ'; let (F, D, C) be an RPO for $\mathcal{F}(a', r')$ w.r.t. $\mathcal{F}(G', E')$ in $\mathrm{WR}_{\mathbf{A}}$; then there exists an RPO (F', D', C') for a', r' w.r.t. G', E' in $\mathrm{WR}_{\mathbf{A}'}$ such that $\mathcal{F}(F', D', C') = (F, D, C)$. ■

Note that the conditions of Definition 57 say that the image of every reaction rule of $\mathbf{A}'$ is a rule of $\mathbf{A}$, and the image of a context E' of $\mathbf{A}'$ is at least as active as E'; so we deduce the following:

Proposition 59 (functors which create sufficient RPOs) *Let* $\mathrm{WR}_{\mathbf{A}}$ *have all redex-RPOs, and let the functor* $\mathcal{F} : \mathrm{WR}_{\mathbf{A}'} \to \mathrm{WR}_{\mathbf{A}}$ *create redex-RPOs. Then* $\mathrm{WR}_{\mathbf{A}'}$ *has all redex-RPOs.*

We shall meet some very interesting functors of this kind, and they will allow us to work in models which are refinements of BRSs as well as directly in BRSs themselves. There is always work to do in proving that a specific functor creates RPOs; so it is worth finding some general conditions that ensure this and are easy to check. We shall now look at one such set of sufficient conditions, which we shall be able to apply a few times. For the remainder of this subsection we revert to our notation for arbitrary precategories.

Definition 60 (safe functor) Let $\mathcal{F} : \mathbf{A}' \to \mathbf{A}$ be a functor of precategories. Say that $\mathcal{F}$ is *safe* if it has the following properties:

(1) Let f be an epi in $\mathbf{A}$, and let $\mathcal{F}(f') = f$ and $\mathcal{F}(h') = g \circ f$. Then there exists unique g' such that $\mathcal{F}(g') = g$ and $h' = g' \circ f'$.

(2) Let $\vec{h}$ be an IPO for $\vec{f}$ in $\mathbf{A}$, and let $\mathcal{F}(\vec{f'}) = \vec{f}$. Then there exist a bound $\vec{h}'$ for $\vec{f'}$ in $\mathbf{A}'$ such that $\mathcal{F}(\vec{h}') = \vec{h}$.

If $\mathcal{F}$ is an inclusion, call $\mathbf{A}'$ a *safe subprecategory* of $\mathbf{A}$. ■

Note that condition (2) does not require $\vec{h}'$ to be an IPO in $\mathbf{A}'$; it is therefore not too hard to verify in particular cases. An easy consequence of the definition is:

Proposition 61 (safe functors create epis) *Let* $\mathcal{F} : \mathbf{A}' \to \mathbf{A}$ *be safe, with* $\mathcal{F}(f')$ *an epi of* $\mathbf{A}$*. Then* f' *is an epi of* $\mathbf{A}'$*.*

When the functor is an inclusion then the safe condition can be simplified:

Proposition 62 (safe subprecategories) *A subprecategory* $\mathbf{A}'$ *of* $\mathbf{A}$ *is safe iff the following hold:*

(1) If f *is an epi in* $\mathbf{A}$*, and both* $g \circ f$ *and* f *are in* $\mathbf{A}'$*, then* g *is in* $\mathbf{A}'$*.*

(2) If $\vec{h}$ *is an IPO for* $\vec{f}$ *in* $\mathbf{A}$*, and* $\vec{f}$ *is in* $\mathbf{A}'$*, then* $\vec{h}$ *is in* $\mathbf{A}'$*.*

We now show that, with one extra condition (which will always hold for BRSs) if $\mathcal{F}$ is safe then it creates enough RPOs.

Proposition 63 (safe WRS functors create RPOs) *Let* $\mathcal{F} : \mathbf{A}' \to \mathbf{A}$ *be safe. Let* $\vec{g}'$ *be a bound for* $\vec{f}'$ *in* $\mathbf{A}'$*, with* $\mathcal{F}(\vec{f}') = \vec{f}$ *and* $\mathcal{F}(\vec{g}') = \vec{g}$*. Let* $(\vec{h}, h)$ *be an RPO in* $\mathbf{A}$ *for* $\vec{f}$ *w.r.t.* $\vec{g}$*, such that both* f_1 *and* h_0 *are epis. Then there exists an RPO* $(\vec{h}', h')$ *in* $\mathbf{A}'$ *for* $\vec{f}'$ *w.r.t.* $\vec{g}'$*, with* $\mathcal{F}(\vec{h}', h') = (\vec{h}, h)$*.*

Proof (outline) The first step is to create in $\mathbf{A}'$ a bound $\vec{h}'$ for $\vec{f}'$ such that $\mathcal{F}(\vec{h}') = \vec{h}$, using condition (2). By Proposition 61 h'_0 is epi; therefore by condition (1) we find h' in $\mathbf{A}'$ such that $h' \circ h'_0 = g'_0$. We now wish to prove that $h' \circ h'_1 = g'_1$. For this, since f'_1 is epi by Proposition 61, it is enough to prove that $h' \circ h'_1 \circ f'_1 = g'_1 \circ f'_1$; this follows from equations already known. We have therefore established that $(\vec{h}', h')$ is an RPO candidate.

Now let $(\vec{k}', k')$ be any other candidate. Then its $\mathcal{F}$-image $(\vec{k}, k)$ in $\mathbf{A}$ is a candidate; hence, since $(\vec{h}, h)$ is an RPO in $\mathbf{A}$, there is a unique arrow j mediating it to $(\vec{k}, k)$. Create the preimage j' such that $j' \circ h'_0 = k'_0$, using condition (1) for h_0 epi. Now deduce that $j' \circ h'_1 = k'_1$ and $k' \circ j' = h'$ from known equations, using the facts that f'_1 and h'_0 respectively are epis.

Thus j' is a mediator from $(\vec{h}', h')$ to $(\vec{k}', k')$. Its uniqueness can be deduced from the uniqueness of the mediator j in $\mathbf{A}$. This completes the proof that $(\vec{h}', h')$ is an RPO. ■

We shall see the consequence of this proposition for BRSs in the next subsection.

5.3 BRS functors and sub-BRSs

In Definition 54 we defined a class of WRSs directly on **Big** and $\mathbf{Big}_{\bar{\mathsf{a}}}$, simply by equipping bigraphs with a signature and reaction rules. But there are many variants of bigraphs we may want to consider; for example linear bigraphs (those where all edges have exactly two ports), and oriented bigraphs with – say – exactly one positive port in every edge. WRS functors allow us to derive such variants, and if they create redex-RPOs then the derived BRS will have all redex-RPOs.

Definition 64 (derived BRS) A WRS functor $\mathcal{F} : \mathrm{WR}_{\mathbf{A}'} \to \mathbf{Big}(\mathcal{K}, \mathsf{reacts})$ is called a *bigraphical reactive system (BRS) derived from* $\mathbf{Big}(\mathcal{K}, \mathsf{reacts})$. If $\mathcal{F}$ is an inclusion then call $\mathrm{WR}_{\mathbf{A}'}$ a *sub-BRS* of $\mathbf{Big}(\mathcal{K}, \mathsf{reacts})$. Make the same definitions also with $\mathbf{Big}_{\bar{\mathsf{a}}}$. ■

The following can then be deduced from Proposition 59 together with Theorem 20:

Corollary 65 (bisimilarity for derived BRSs) *Let* $\mathrm{WR}_{\mathbf{A}} = \mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K}, \mathsf{reacts})$*, with every redex in* reacts *an epimorphism. Let the WRS functor* $\mathcal{F} : \mathrm{WR}_{\mathbf{A}'} \to \mathrm{WR}_{\mathbf{A}}$ *create redex-RPOs. Then* $\mathrm{WR}_{\mathbf{A}'}$ *has all redex-RPOs. Consequently wide bisimilarity in* $\mathrm{WR}_{\mathbf{A}'}$ *is a congruence.*

Now we deduce a further corollary concerning safe functors, with the help of Proposition 63:

Corollary 66 (bisimilarity for safe derived BRSs) *Let* $\mathrm{WR}_{\mathbf{A}} = \mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K}, \mathsf{reacts})$*, with every redex in* reacts *an epimorphism. Let the WRS functor* $\mathcal{F} : \mathrm{WR}_{\mathbf{A}'} \to \mathrm{WR}_{\mathbf{A}}$ *be safe. Then* $\mathcal{F}$ *creates redex RPOs. Consequently* $\mathrm{WR}_{\mathbf{A}'}$ *has all redex-RPOs, and wide bisimilarity in* $\mathrm{WR}_{\mathbf{A}'}$ *is a congruence.*

Proof We need only note that in a BRS, whenever a redex r is epi, and a, r has an RPO with legs F, D, then F is also epi; consequently the conditions of Proposition 63 are met. ■

We shall now look at a few examples of safe BRS functors or sub-BRSs that create redex-RPOs. In each case the construction seems a rather natural one, and it is not surprising to find that the RPO condition can be transferred to them. Nonetheless, the proof of safety in each case needs a little care. We omit these proofs; the reader interested in getting to grips with the theory will find them to be excellent exercises. In some cases, but not all, the no-alias condition of $\mathbf{Big}_{\bar{\mathsf{a}}}$ is essential.

Example 13 (linear bigraphs) A very simple sub-BRS of $\mathbf{Big}_{\bar{\mathsf{a}}}$ consists of the *linear* bigraphs; those in which every edge has exactly two ports. Linearity (in this sense) is a purely monographic property. This sub-BRS is safe. The proof of the IPO condition –part of the safety condition– depends upon the characterisation of IPOs in monographs in Construction 45; it is made easier by the observation that all linear monographs are epis. ■

Example 14 (epimorphic bigraphs) It is very attractive to deal only with epimorphic bigraphs, since they have a simple dynamic theory; this is because there are no elisive transitions (see Section 6), so the characterisation of transitions is easier. This again is a safe sub-BRS.

A slightly larger sub-BRS –again safe– consists of the bigraphs whose monographs (but not necessarily the topographs) are epi. This appears to be practically significant. Note that Example 2 in Section 2 could not be handled in this sub-BRS, because the epi property is not preserved by reaction. This is clear from the reaction rule; the name x is unused by the reactum of the rule. But it can be shown that if the reactum of every reaction rule has no idle names then the epi property for monographs will be preserved both by reaction and by transition; thus the behaviour of systems can be analysed entirely within this sub-BRS. Clearly this is not possible for the π-calculus, but it may be true for certain variant calculi. ■

Example 15 (oriented bigraphs) We may consider a kind of bigraph with extra structure, by which a *sign* is ascribed to every port (including names and co-names). To achieve this we must extend signatures to assign a sign to each port of a control; we must also extend interfaces to assign a sign to every name in the name set. Then in an oriented bigraph $G : \langle m, X^{\mathsf{sg}}\rangle \to \langle n, Y^{\mathsf{sg}}\rangle$ –where the superscript $(\cdot)^{\mathsf{sg}}$ indicates a sign assignment– each name will have the sign in the outer interface, each coname will have sign opposite to that in the inner interface, and each inner port will have sign dictated by the signature.

To make the signs do some work we can *orient* each graph by requiring every edge to contain –say– exactly one positive port. For example, if the left port of the control `get` in Example 2 is declared positive in the signature, then every message is addressed to a unique receiver. This is true for example in the Join calculus [8], but not in the π-calculus, where it is an important source of non-determinism. Thus the orientation of graphs has significant practical implications.

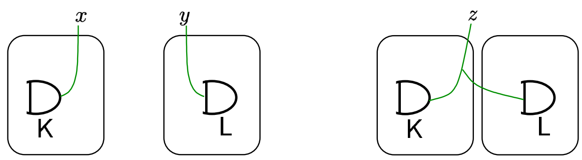
The WRS functor in this case is the obvious forgetful functor from signed BRSs to $\mathbf{Big}_{\bar{\mathsf{a}}}$. It is not an inclusion, so we do not have a sub-BRS. But it is a safe functor. This is one case where the reader will discover the no-alias condition useful; I do not know whether the functor is safe if aliases are admitted. ■

Example 16 (typed bigraphs) Analogous to signing ports, we can type them. We leave the details to the reader. The practical significance is obvious. Again we get a safe forgetful functor from typed to untyped BRSs. However, the situation is only simple of we demand exact matching of types: no polymorphism and no subtyping. It is an intriguing question how to deal with these richer type phenomena in bigraphical systems. ■

At least two of the above examples – linear BRSs and oriented BRSs – correspond to cases where colleagues and I had previously, with considerable difficulty, established the existence of RPOS [4, 18]. (The linear case has not been published). The correspondence is not exact, because we have adopted the no-alias condition here. However, these examples provide evidence that bigraphs as defined here are somehow central in this kind of theory.

5.4 BRSs for scoping and binding

We now move to an example of a different flavour, to do with the scoping of names. As observed at the outset, we have adopted the orthogonal treatment of topographs and monographs –with no association between a name and any particular topographic region– as the best way to get a tractable behavioural theory. But there is another reason to avoid having *every* name of a bigraph G associated with a particular region of G. Consider the agents a and b in the diagram, where K and L have arity one.



$a = \mathsf{K}(x)$ $b = \mathsf{L}(y)$ $c = \mathsf{K}(z) \mid \mathsf{L}(z)$

Clearly the agent c as shown is given by $c = ({}^{z}/\underline{x} \circ a) \mid ({}^{z}/\underline{y} \circ b)$, or equivalently by $c = ({}^{z}/\underline{x}, \underline{y}) \circ (a \otimes b)$. (We are anticipating the notation and conventions for substitution contexts, to be introduced in Subsection 6.1.) Here z is localised to neither region of c, or at least it is unnatural to make it so. The reader may note that this phenomenon also arises in our example of a remote reaction rule for the π-calculus, Example 4. It turns out that in the refined BRS presented below, which will permit but not enforce name-localisation, the only context C such that $c = C \circ (a \otimes b)$ will have z local to neither region. Thus to enforce the localisation of each name to a particular region reduces expressive power.

We shall use *scoping* to mean the association of a name with a region. We have just seen the advantage of having some names with unlimited scope, but we need also to reflect the kind of scoping represented by the π-calculus prefix construction $x(z).P$, where the scope of z actually coincides with the term guarded by the prefix. In our π-calculus examples this becomes the region inside a get-node.

In this subsection we propose a refined notion of bigraph that provides scoping, and we shall take advantage of it in Subsection 6.5 where we propose a definition of parametric reaction rules. In both of these subsections we give less detail than in the rest of the report; the consequences of these definitions are under study, to be reported at a future date.

If we designate the scope of a name x to be a particular region r in a bigraph G, then we want to ensure that (1) any node with a port linked to x also lies in that region, and (2) $G \circ H$ will have this property for any H that is also subject to the scoping discipline. Hence the following definition:

Definition 67 (scoping bigraph) Let $\mathcal{K}$ be a signature as for bigraphs. Then the precategory $\mathbf{Big}^{\mathsf{s}}_{\bar{\mathsf{a}}}(\mathcal{K})$ of *scoping bigraphs over* $\mathcal{K}$ is as follows: An object takes the form $\langle m, \vec{X}, X\rangle$ where $\langle m, X\rangle$ is an object of $\mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K})$ and $\vec{X} = X_0, \ldots, X_{m-1}$ a sequence of disjoint subsets of X. An arrow $G : \langle m, \vec{X}, X\rangle \to \langle n, \vec{Y}, Y\rangle$ is a bigraph $G : \langle m, X\rangle \to \langle n, Y\rangle$ of $\mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K})$ such that, whenever $y \in Y_r$ for $r \in n$ a root of G, then two *scoping* conditions hold:

(S1) If xGy for $x \in X$ then there exists a site $s \in m$ with $x \in X_s$ and $s <_G r$.

(S2) If pGy where p is any inner port of node u in G, then $u <_G r$.

The objects and arrows are called *scoped interfaces* and *scoping bigraphs*. ■

It is straightforward to check that, with composition and identities as in $\mathbf{Big}_{\bar{\mathsf{a}}}$, this yields a precategory.[9] With an obvious definition of tensor product it also yields a wide monoidal precategory.

There is a natural forgetful functor from scoping bigraphs to bigraphs. This functor preserves epis, and can be proved to create RPOs. Omitting the details of the definition of a *scoping bigraphical reactive system*, which faithfully matches the standard BRS definition (Definition 54), we have the following:

Theorem 68 (redex RPOs and bisimilarity congruence in scoping BRSs) *For any signature $\mathcal{K}$ and set* reacts *of reaction rules, the forgetful WRS functor from $\mathbf{Big}^{\mathsf{s}}_{\bar{\mathsf{a}}}(\mathcal{K},\mathsf{reacts})$ to a BRS creates redex-RPOs. Therefore if every redex in* reacts *is an epimorphism, then $\mathbf{Big}^{\mathsf{s}}_{\bar{\mathsf{a}}}(\mathcal{K},\mathsf{reacts})$ has all redex-RPOs. Hence wide bisimilarity of agents in $\mathbf{Big}^{\mathsf{s}}_{\bar{\mathsf{a}}}(\mathcal{K},\mathsf{reacts})$ is a congruence.*

Proof (sketch) The only new requirement is to prove the first part. The functor turns out not to be safe, and therefore a specific argument is needed. This depends upon identifying the correct mediating interface in constructing an RPO in the scoped BRS. The details are omitted here. ■

Scoping pertains to the names of a bigraph; we now turn to a related notion pertaining to control ports, which we shall call *binding*.[10] A *control port* of a control K is any member of its arity $ar(K)$; it also refers to the corresponding port of any K-node. We want to designate certain control ports –such as the right-hand port of a get node in our π-calculus examples– as binding, and to ensure that binding is propagated downwards by composition. The following refinement of scoping bigraphs arises naturally.

Definition 69 (binding bigraphs) Let $\mathcal{K}^{\mathsf{b}}$ be a *binding* signature, i.e. a signature $\mathcal{K}$ as for bigraphs with certain control ports designated as *binding*. Then the precategory $\mathbf{Big}^{\mathsf{b}}_{\bar{\mathsf{a}}}(\mathcal{K}^{\mathsf{b}})$ of *binding bigraphs over* $\mathcal{K}^{\mathsf{b}}$ is the subprecategory of $\mathbf{Big}^{\mathsf{s}}_{\bar{\mathsf{a}}}(\mathcal{K})$ whose arrows $G : \langle m, \vec{X}, X\rangle \to \langle n, \vec{Y}, Y\rangle$ satisfy three further conditions for every binding port q of a control node v in G:

(B1) If xGq for $x \in X$ then there exists a site $s \in m$ with $x \in X_s$ and $s <_G v$.

(B2) If pGq where p is any inner port of node u in G, then $u <_G v$.

(B3) q is closed in G. ■

Note that the first two conditions correspond closely to the scoping conditions. It is a routine matter to check that this is a monoidal subprecategory. Furthermore, concerning the natural forgetful functor to bigraphs we have the following as a corollary of Theorem 68:

Corollary 70 (redex RPOs and bisimilarity congruence in binding BRSs) *For a binding signature $\mathcal{K}^{\mathsf{b}}$ and set* reacts *of reaction rules, the forgetful WRS functor from $\mathbf{Big}^{\mathsf{b}}_{\bar{\mathsf{a}}}(\mathcal{K}^{\mathsf{b}},\mathsf{reacts})$ to a BRS creates redex-RPOs. Therefore if every redex in* reacts *is an epimorphism, then $\mathbf{Big}^{\mathsf{b}}_{\bar{\mathsf{a}}}(\mathcal{K}^{\mathsf{b}},\mathsf{reacts})$ has all redex-RPOs. Hence wide bisimilarity of agents in $\mathbf{Big}^{\mathsf{b}}_{\bar{\mathsf{a}}}(\mathcal{K}^{\mathsf{b}},\mathsf{reacts})$ is a congruence.*

[9] This depends upon the no-alias assumption; an extra condition would be needed for **Big**.

[10] This is consistent with the meaning of "binding" for action structures.

Binding BRSs deserve further study. For example, the stronger constraint that *every* name should be scoped may have some interest. For the present paper, we return to these derived BRSs in Subsection 6.5.

6 Towards applications

In this section we prepare for detailed exploration of the behaviour of particular bigraphical systems. In Subsection 6.1 we define some useful forms of bigraph and some important operations, including all the constructions of the *term language* introduced in Subsection 2.2. In Subsection 6.2 we look at two manipulations of interfaces, *elision* and *extension*. In Subsection 6.3 we find certain transitions to be superfluous in establishing bisimilarity, introduce a notion of *adequate set* of transitions, and enumerate all transitions for a specific BRS. In Subsection 6.4 we define the notion of *engaged* transition, which depends essentially on the fact that transitions are placed. It turns out that the engaged transitions are adequate under certain conditions. Finally, in Subsection 6.5 we begin to examine *parametric reactions* and the π-calculus, leaving their detailed study for future work.

Unless otherwise stated we are working in $\mathbf{Big}_{\bar{\mathsf{a}}}(\mathcal{K},\mathsf{reacts})$ for unspecified $\mathcal{K}$ and reacts.

Terminology We shall call an agent or context *narrow* if its interfaces have width one, *shallow* if all its sites are shallow, *atomic* if all its nodes are atomic, and *open* if all its ports are open.

6.1 Operations on bigraphs

Definition 71 (special graphs) Notation for several entities with empty support is introduced (with alternative notation in the second column):

Topographs:	0	$: 0 \to 0$	id_0	no sites and no roots
	1_m	$: m \to 1$		m sites and one root
	1	$: 0 \to 1$	1_0	one barren root
	n	$: 0 \to n$	$1 \otimes \cdots \otimes 1$	no sites and n roots
Monographs:	$\emptyset$	$: \emptyset \to \emptyset$	$\mathsf{id}_\emptyset$	no names or conames
	X	$: \emptyset \to X$		discrete set of names
	$\underline{X}$	$: X \to \emptyset$		discrete set of conames
	νX	$: X \to \emptyset$	$\underline{X}$	restrict X
	$\underline{x} = y$	$: \{x\} \to \{y\}$	$^{y}/_{\underline{x}}$	substitution
	$x = y$	$: \emptyset \to \{x, y\}$		alias
	$\underline{x} = \underline{y}$	$: \{x, y\} \to \emptyset$		co-alias

For any interface $I = \langle n, X\rangle$ we use I also to denote the bigraph combination $\langle n, X\rangle : \epsilon \to I$, having n barren roots and idle names X. ∎

In the monograph operations indexed by the set X we may use x, $\underline{x}$, νx for the case when $X = \{x\}$, a singleton. Note that aliases occur in **Big** but not in $\mathbf{Big}_{\bar{\mathsf{a}}}$.

Definition 72 (promotion) If a topograph $A : m \to n$ has empty support we may *promote* it to a bigraph $\langle A, \mathsf{id}_X\rangle : \langle m, X\rangle \to \langle n, X\rangle$; we abbreviate $\langle A, \emptyset\rangle$ to A. Similarly, if a monograph $B : X \to Y$ has empty support we may promote it to $\langle \mathsf{id}_m, B\rangle : \langle m, X\rangle \to \langle m, Y\rangle$, abbreviating $\langle 0, B\rangle$ to B.

More generally, an operation f on topographs which does not change the support set may be *promoted* to an operation on bigraphs by $f(G) \triangleq \langle f(G^{\mathsf{T}}), G^{\mathsf{M}}\rangle$. Similarly, for an operation g on monographs which preserves support, $g(G) \triangleq \langle G^{\mathsf{T}}, g(G^{\mathsf{M}})\rangle$. ∎

To illustrate promotion, for the bigraph $G : \epsilon \to \langle m, \{x\} \uplus Z\rangle$ we may substitute y for x by writing $(^y/_{\underline{x}} \otimes \mathsf{id}_Z) \circ G$, with codomain $\langle m, \{y\} \uplus Z\rangle$; it abbreviates $\langle \mathsf{id}_m, {}^y/_{\underline{x}} \otimes \mathsf{id}_Z\rangle \circ G$.

We shall also allow ourselves to abbreviate any bigraph $F \otimes \mathsf{id}_Z$ to F, when the context makes Z clear. This is convenient when F is a restriction, substitution or co-alias; for example we can further abbreviate $(^y/_{\underline{x}} \otimes \mathsf{id}_Z) \circ G$ to $(^y/_{\underline{x}}) \circ G$.

We now introduce a derived form of tensor product, useful in applications:

Definition 73 (parallel product) The *parallel product* of two interfaces $I_i = \langle m_i, X_i\rangle$ $(i = 0, 1)$ is always defined; it is $I_0 \mid I_1 \triangleq \langle m_0 + m_1, X_0 \cup X_1\rangle$.

The *parallel product* $A_0 \mid A_1$ of two topographs $\vec{A}$ with disjoint node sets is just their tensor product. If two monographs $\vec{B} : \vec{X} \to \vec{Y}$ have disjoint node sets $\vec{V}$ then their *parallel product* $B_0 \mid B_1 : X_0 \cup X_1 \to Y_0 \cup Y_1$ has node set $V_0 \uplus V_1$, control function $ctrl_0 \uplus ctrl_1$ and equivalence $\equiv_{B_0} \sqcup \equiv_{B_1}$. We extend it to two bigraphs $G_i : I_i \to J_i$ with disjoint node sets by

$$G_0 \mid G_1 \;\triangleq\; \langle G_0^{\mathsf{T}} \mid G_1^{\mathsf{T}}, G_0^{\mathsf{M}} \mid G_1^{\mathsf{M}}\rangle : I_0 \mid I_1 \to J_0 \mid J_1 \ .$$ ∎

Parallel product resembles the tensor product $\otimes$, but takes the *union* of domains and codomains; thus it may coalesce edges from G_0 and G_1. It is associative on topographs, monographs and bigraphs, with units 0, $\emptyset$ and ϵ respectively. It is commutative on monographs.[11]

Later we shall need the following simple property that an agent can be factorised uniquely into a parallel product of narrow factors:

Proposition 74 (unique parallel factorisation) *An agent $d : \epsilon \to \langle m, X\rangle$ in $\mathbf{Big}_{\bar{\mathsf{a}}}$ can be expressed uniquely (up to a renaming of Z) in the form $\nu Z \circ (d_0 \mid \cdots \mid d_{m-1})$, where $d_i : \epsilon \to \langle 1, X\rangle$ $(i \in m)$ and each $z \in Z$ is non-idle in at least two of the factors d_i.*

Note that $\mathbf{Big}_{\bar{\mathsf{a}}}$ is closed under parallel product of bigraphs with disjoint sets of conames; this condition will hold when we use it in examples.

Using parallel product we can define non-injective substitutions such as $(^z/_{\underline{x},\underline{y}}) \triangleq {}^z/_{\underline{x}} \mid {}^z/_{\underline{y}}$. The co-alias $\underline{x} = \underline{y}$ is then expressible as $\underline{z} \circ (^z/_{\underline{x},\underline{y}})$. Dually, we can define co-substitutions such as $(^{x,y}/_{\underline{z}}) \triangleq {}^x/_{\underline{z}} \mid {}^y/_{\underline{z}}$; thence the alias $x = y$ is expressible as $(^{x,y}/_{\underline{z}}) \circ z$. This shows how $\mathbf{Big}_{\bar{\mathsf{a}}}$ is *not* closed under arbitrary parallel product.

Together with parallel product, the next two definitions explain most of the term language introduced in Subsection 2.2:

Definition 75 (merge) Let G have codomain $\langle m, X\rangle$. Define $[G] \triangleq \langle 1_m, \mathsf{id}_X\rangle \circ G$, the *merge* of G; it coalesces the m regions of G into one, so has codomain $\langle 1, X\rangle$. ∎

[11] In [23] parallel product was wrongly stated to be commutative on bigraphs, though no results there were affected.

Finally, the basic building blocks for all bigraphs are as follows:

Definition 76 (element, molecule, atom) Let K have arity k, let $\vec{x} = x_0, \ldots, x_{k-1}$ be a sequence of names (not necessarily distinct) and let $X = \{\vec{x}\}$. Then the *elementary* bigraph or *element* $K_v(\vec{x}) : \langle 1, \emptyset \rangle \to \langle 1, X \rangle$ has one root, one site, and a single node v which is the site's parent. For $0 \leq i < k$ the i^{th} inner port of v is linked to the name x_i.

For any bigraph $G : I \to \langle m, Y \rangle$ we call $(K_v(\vec{x}) \mid \mathsf{id}_Y) \circ [G] : I \to \langle 1, X \cup Y \rangle$ a *molecule*[12] ; using our abbreviation convention we may write it $K_v(\vec{x}) \circ [G]$. When $G = \epsilon$ then the molecule is $K_v(\vec{x}) \circ 1_0$, which we call an *atom*. ■

Note that, in the molecular construction, the names Y of G may not be disjoint from those of the element; thus elements nested inside one another may have their ports linked.

We may now complete the semantic definition of the term language of Subsection 2.2. Given the constructions defined here, all that remains is to make minor syntactic adjustments (since the term language is oriented to programming rather than algebra), and to take care in handling the sites.

To illustrate, consider the bigraph $G : \langle 3, X \rangle \to \langle 2, Y \rangle$ of Figure 7. Suppose that the nodes v_0, v_1 have control K with arity one, and v_2, v_3 have control L with arity three. Then, omitting node subscripts on controls, a term for G is

$$\begin{aligned} \underline{x}_0 = y_2 \,.\; & [\, K(y_0)[\, K(y_0)[-_0]\,] \mid L(y_0, y_1, y_2)\,] \\ & \mid\; [\, -_1 \mid L(y_2, \underline{x}_1, y_2)[-_2]\,] \,. \end{aligned}$$

Note that since $\underline{x}_0$ and y_2 are equivalent, $\underline{x}_0$ could appear in place of y_2 anywhere in the body of the term. To represent the term algebraically we transform it as follows: (1) render the equation prefix by parallel product; (2) make explicit the composition ($\circ$) in each molecule; (3) replace any conames (here only $\underline{x}_1$) in each molecule by a new restricted name; and (4) replace the sites by identities. The new name in step (3) is needed because X and Y are not disjoint in general. So, assuming that composition binds more tightly than parallel product, we get the expression

$$\begin{aligned} \underline{x}_0 = y_2 \mid \nu z \circ (A \mid {}^{z}/_{\underline{x}_1}), \text{ where } A = \; & [\, K(y_0) \circ [\, K(y_0) \circ [\mathsf{id}_1]\,] \mid L(y_0, y_1, y_2) \circ 1_0\,] \\ & \mid\; [\mathsf{id}_1 \mid L(y_2, z, y_2) \circ [\mathsf{id}_1]\,] \,. \end{aligned}$$

Note the identities for the sites. Of course '$[\mathsf{id}_1]$' can be removed, since $[\mathsf{id}_1] = \mathsf{id}_1$.

We now give the general construction. Note that the sites in a term will not always occur in numerical order (as they do in the above example), so a permutation is needed in general.

Construction 77 (Meaning of the term language) Let $g = e.\,\nu L.\, b$ be a linear term of the language, representing a bigraph $G : \langle m, X \rangle \to \langle n, Y \rangle$. Here e contains (besides singletons) equations representing aliases, co-aliases and substitutions, L are the local names (disjoint from X and Y), and the body b is a sequence of n regions. Then, in terms of our operations over bigraphs, we represent g by the expression

$$G \;\triangleq\; E \mid (\nu L \circ B) \,,$$

where E is just the parallel product of the members of e, and the expression $B : \langle m, X \rangle \to \langle n, L \uplus Y \rangle$ is obtained from the term b as follows. The overall structure of b needs only slight change; for each molecule

[12] This definition replaces that of molecular bigraph in [23], which was inconvenient for defining the term language.

$K(\vec{z})[\cdots]$ we write $K(\vec{z})\circ[\cdots]$, and for a molecule $K(\vec{z})$ with no contents we write $K(\vec{z})\circ 1_0$, an atom. It remains to deal with the sites $-_0, -_1, \ldots, -_{m-1}$ (each of which occurs just once since g is linear), and the occurrences of conames $\underline{x}$ in atoms and molecules. To this end:

- let $\pi : m \to m$ be the topograph isomorphism whose parent function is the bijection $\{i \mapsto s_i \mid i \in m\}$, where $-_{s_0}, \ldots, -_{s_{m-1}}$ is the left-to-right order of sites in the term g;
- if $X = \{x_1, \ldots, x_k\}$, let $X' = \{x'_1, \ldots, x'_k\}$ be distinct new names disjoint from Y and L.

Then, in the body b, replace each site by id_1 and each coname $\underline{x}_i$ by x'_i, leading to an expression A. Finally, complete the construction by setting

$$B \triangleq \nu X' \circ (A \circ \pi \mid x'_1/\underline{x}_1 \mid \cdots \mid x'_k/\underline{x}_k)\ . \qquad \blacksquare$$

So we have shown that the term language is 'just' algebra, with a little sugar added. We have, in effect, established the term language as the core of a programming language for bigraphical systems. The rest of such a language should allow a user to specify a signature and a set of reaction rules. It may also take advantage of some of the BRS refinements described in Section 5, in particular the ability to provide names with scope.

6.2 Elision and extension

We now look at two manipulations of interfaces. One of them, *extension*, consists in adding extra idle names or barren roots to bigraph A, arriving at $H \otimes A$ (H an interface). The other, with which we begin, is the notion of *elision* which arose in characterising IPOs for both topographs and monographs in Section 4. As we saw there, for a consistent pair $\vec{A}$ of epis there is exactly one IPO. But we are interested in pairs a, r where r is epi (being a redex) while a may not be. We begin by looking at the extra IPOs which may then arise; these are elisions. Let us define them more exactly:

Definition 78 (elision) For a topograph $B : m \to n$ and finite ordinal ℓ, let η_ℓ be a map from ℓ to the non-atomic nodes of B. Form $\eta_\ell B : \ell + m \to n$ from B by incrementing its existing sites by ℓ, and defining the parent of each new site $i \in \ell$ to be $\eta_\ell(i)$. Then $\eta_\ell B$ is an *elision of* ℓ *in* B.

Similarly, let $B : Y \to Z$ be a monograph and X disjoint from Y. For any map η_X from X to the closed ports of B, form $\eta_X B : X \uplus Y \to Z$ by adding each new coname x to the equivalence class of $\eta_X(x)$. Then $\eta_X B$ is an *elision of* X *in* B.

For an interface $H = \langle \ell, X\rangle$, an *elision of* H *in* a bigraph $B : I \to J$ takes the form $\eta_H B \triangleq \eta_\ell(\eta_X B)$, with interfaces $H \otimes I \to J$. A transition is called *an elision*, or *elisive*, if it takes the form $H \otimes a \xrightarrow{\eta_H F}_\lambda a'$ up to isomorphism. $\blacksquare$

A key property of an elision is that $(\eta_H B) \circ H = B$.

Now Proposition 53 expresses the RPOs of $\mathbf{Big}_{\bar{\mathsf{a}}}$ in terms of those for topographs and monographs, while Corollaries 32 and 46 characterise the IPOs of **Top** and $\mathbf{Mog}_{\bar{\mathsf{a}}}$. For non-epi agents in $\mathbf{Big}_{\bar{\mathsf{a}}}$ this yields:

Proposition 79 (IPOs for non-epi agents) *Let r be an epi. Then*

(1) A pair F, D of contexts is an IPO for the agents $x \otimes A, r$ iff it takes the form either $\mathsf{id}_x \otimes F', x \otimes D'$ *or* $\eta_x F', D'$*, where* F', D' *is an IPO for* a, r.

(2) A pair F, D of contexts is an IPO for the agents $1 \otimes A, r$ iff it takes the form either $\mathsf{id}_1 \otimes F', 1 \otimes D'$ *or* $\eta_1 F', D'$*, where* F', D' *is an IPO for* a, r.

A first consequence of this is that bisimilarity is strictly preserved when we *extend* the interface of an agent a by adding idle names and/or barren roots. For example, when considering whether a and b are equivalent, it makes no difference how large we make the name set in their outer interface, provided it includes the names they actually use. To be precise:

Proposition 80 (interface extension strictly preserves bisimilarity) *Let H be an interface. Then*

$$a \sim b \text{ iff } H \otimes a \sim H \otimes b\,.$$

Proof ($\Rightarrow$) In the forward direction, simply recall that $\sim$ is a congruence (Theorem 20), so is preserved by a context of the form $H \otimes \mathsf{id}$.

($\Leftarrow$) By induction it will be enough to treat only the two cases $H = 1$ and $H = x$; we distinguish these cases below. In each case we shall show that $\mathcal{S} = \{(a, b) \mid H \otimes a \sim H \otimes b\}$ is a wide bisimulation, where $H = \langle \ell, X \rangle$. Assume $a, b : \epsilon \to I$ with H disjoint from I. Let $a \xrightarrow{F}_{\lambda} a'$, choosing $F : I \to J$ so that $F \circ b$ is defined. We may also assume H disjoint from J, since IPOs are preserved by isomorphism. Then $H \otimes a \xrightarrow{G}_{\ell \cdot \lambda} H \otimes a'$ by Proposition 79, where $G = \mathsf{id}_H \otimes F$.

Now $F \circ b$ is defined, so $G \circ (H \otimes b)$ is defined; hence from $H \otimes a \sim H \otimes b$ we deduce $H \otimes b \xrightarrow{G}_{\ell \cdot \lambda} d'$ with $H \otimes a' \sim d'$. Now, if $H = 1$ (resp. $H = x$) this transition cannot be isomorphic to any transition of the form $H \otimes b \xrightarrow{\eta'_H F'}_{\mu}$, since such an elision closes the site 0 (resp. the coname $\underline{x}$), which is open in G. Hence by Proposition 79 there is a transition $b \xrightarrow{F}_{\lambda} b'$ with $d' = H \otimes b'$. So $(a', b') \in \mathcal{S}$, and we are done. ■

6.3 Adequate transitions

How essential are all our derived transitions? We know that if we define bisimilarity over the entire set we get a congruence. What happens if we define it only over a subclass? Every subclass of the transitions gives rise to a bisimilarity equivalence, and as the class is reduced this equivalence is increased in general. But it may be that certain reduced classes give rise to the same relation as the full class – which is not only an equivalence but even a congruence. Let us make this precise:

Definition 81 (adequate transitions) Let T be a class of transitions, i.e. a subclass of the quadruples a, F, λ, a' such that $a \xrightarrow{F}_{\lambda} a'$. Then $\mathcal{S}$ is a *T-bisimulation* (or *bisimulation for T*) if, whenever $a\mathcal{S}b$, then for every transition $a \xrightarrow{F}_{\lambda} a'$ in T there exists b' such that $b \xrightarrow{F}_{\lambda} b'$ and $a'\mathcal{S}b'$. let $\sim_T$ be the largest T-bisimulation. Call T *adequate* if $\sim_T = \sim$. ■

(Note especially that the transition of b, which matches the T-transition of a, need not itself be in T. This means that $\sim_T$ is in general not transitive.)

We may loosely say that a transition is *superfluous* if there exists an adequate class not containing it. Now, we may expect a transition $a \xrightarrow{F}_{\lambda} a'$ to be superfluous if 'a contributes nothing' to the redex on which the transition is based; for then the entire redex is 'provided by F'. The tensor product allows us to make this intuition precise and justify it.

Definition 82 (tensorial transition) A transition from $a : \epsilon \to I$ is *tensorial* if, up to isomorphism, it takes the form $a \xrightarrow{\mathsf{id}_I \otimes r}_\lambda a \otimes r'$, where (r, r') is a reaction rule. ■

Note that the form of the transition implies that a and r have disjoint support; if this were not so, we could hardly expect a tensorial transitions to be superfluous. In fact, they are so:

Proposition 83 (non-tensorial transitions are adequate) *The class* NT *of non-tensorial transitions is adequate.*

Proof (outline) It is enough to show that the following is a bisimulation for all placed transitions:

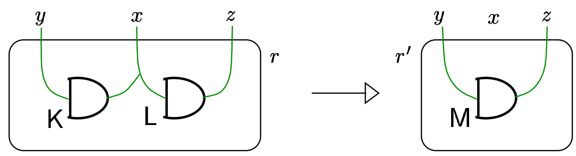
$$\mathcal{S} \triangleq \{(C \circ a, C \circ b) \mid C \text{ active},\ a \sim_{\mathsf{NT}} b\} ;$$

for then, taking C as the identity context, we find $\sim_{\mathsf{NT}} \subseteq \mathcal{S} \subseteq \sim$.

The proof follows the lines Theorem 20. We take an arbitrary transition for $C \circ a$, and derive from it the underlying transition for a by taking an RPO. In case the transition for a is non-tensorial, from $a \sim_{\mathsf{NT}} b$ we deduce a matching transition for b, thence one for $C \circ b$ as in the cited proof. On the other hand suppose that a's transition is tensorial, say $a \xrightarrow{\mathsf{id}_I \otimes r}_\lambda a \otimes r'$, where the reaction rule (r, r') has interface J. then – independently of the fact that $a \sim_{\mathsf{NT}} b$ – it is easy to construct a matching tensorial transition $b \xrightarrow{\mathsf{id}_I \otimes r}_\lambda b \otimes r'$, and to check that the context $b \otimes \mathsf{id}_J$ which occurs in its IPO is appropriately active. The proof then proceeds as in the previous case. ■

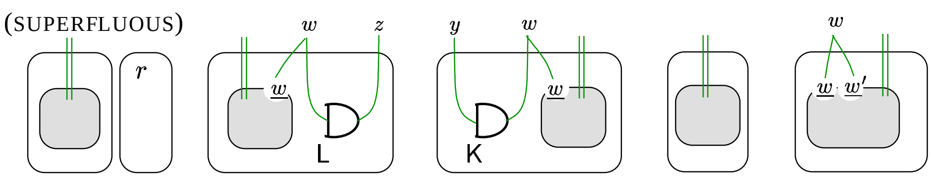
We shall now examine the transitions derived for a simple specific BRS, motivated by the π-calculus.

Example 17 (transitions for a simple BRS) Take the signature $\mathcal{K} \triangleq \{\mathsf{K}, \mathsf{L}, \mathsf{M}\}$, all atomic and non-reactive with arity two. Let there be a single narrow reaction rule:



$$[\,\mathsf{K}(y, x) \mid \mathsf{L}(x, z)\,] \longrightarrow x.\,\mathsf{M}(y, z)$$

We enumerate the kinds of transition possible for a narrow agent $a : \epsilon \to \langle 1, W \rangle$; without loss of generality we assume $x, y, z \notin W$. The enumeration is based upon the extent of node-sharing between a and r. If a has non-empty node-set we obtain five kinds of transitions for a, four of which are non-tensorial. The transition labels are shown in the following diagrams. Any names in W not shown in the diagrams are exported to the outer interface; we have indicated this by a double line.



$[-] \mid r$ $\quad$ $[\, - \mid \mathsf{L}(w, z)\,]$ $\quad$ $[\,\mathsf{K}(y, w) \mid -\,]$ $\quad$ $[\,-\,]$ $\quad$ $w/\underline{w}, \underline{w}'$

The transitions arise as follows:

No nodes shared There is one transition: $a \xrightarrow{[-]|r}_{\{1\}} a \mid r'$. It may also be written $a \xrightarrow{\mathsf{id}_I \otimes r}_{\{1\}} a \otimes r'$, which we recognise as tensorial – hence superfluous. Note especially that this is the only label with two regions; the reaction $r \longrightarrow r'$ occurs 'alongside' a without merging with it.

One node shared If the K-node is shared, then its right-hand port must be open in a, at w say, and the unique IPO of the pair a, r equates w with x and contains just an L-atom. The transition takes the form $a \xrightarrow{[-|\mathsf{L}(w,z)]}_{\{0\}} a'$. Similarly, if the L-node is shared there is a transition $a \xrightarrow{[\mathsf{K}(y,w)|-]}_{\{0\}} a'$.

Both nodes shared There are two cases. If the K- and L-nodes are already linked in a, then there is a transition of the form $a \xrightarrow{[-]}_{\{0\}} a'$; this corresponds to a τ transition in CCS. Otherwise the two nodes are open in a, say at w and w', and there is a transition of form $a \xrightarrow{w/\underline{w},\underline{w'}}_{\{0\}} a'$; in explicit fusions [11] this is a fusion supplied by the context.

If a has an empty node-set we have $a = \langle 1, W \rangle$, and it has only a (superfluous) tensorial transition.

If a is not narrow then there is one other type of transition, when both nodes are shared but are in different regions of a; these regions are coalesced by the 'label' context. We omit the details. ■

The transitions for this example are simple because we excluded all elisions by choosing a redex r to be open and atomic – for then every transition label F is also open and atomic.

Another simplicity in the present example is that in a non-superfluous transition $a \xrightarrow{F}_{\lambda} a'$, if a is narrow then so is F, hence also a'; thus bisimulations for narrow agents need involve only narrow agents. This, in turn, means that placed transitions degenerate to pure transitions; so in this case the addition of places has not increased the discriminating power of transitions.

6.4 Engaged transitions

Example 17 also provides insight for our next topic. All its non-superfluous transitions are such that the underlying reaction rule is *engaged* by the agent a, in the sense that one or more nodes are shared between a and r. The place component of transitions allows us to make this phenomenon precise:

Definition 84 (engaged) A transition $a \xrightarrow{F}_{\lambda} a'$, where $a : \epsilon \to \langle \ell, X \rangle$, is *engaged* if $\mathsf{width}(F)(\ell) \cap \lambda \neq \emptyset$; otherwise it is *disengaged*. ■

It is not immediately clear that, if a transition of a is engaged, then a shares with the redex r of the underlying reaction rule. Indeed it is misleading to speak of 'the' underlying rule; for in general a single transition may arise from different rules. In spite of this, however, we can show that if a transition of a is engaged then the redex of *any* underlying rule will share with a. In fact we get a stronger result, providing another characterisation of 'engaged':

Proposition 85 (engaged means sharing nodes) *Let F, D be any IPO for a, r underlying the transition $a \xrightarrow{F}_{\lambda} a'$. Then the transition is engaged iff $|a| \cap |r| \neq \emptyset$.*

Proof We proceed by induction on the size of the interface $I = \langle \ell, Y \rangle$ of a, taking the basis of the induction to be when a is epi. In that case it is easy to show, from the construction of IPOs, that $\mathsf{width}(F)(\ell) \cap \lambda \neq \emptyset$ iff (the transition is tensorial) iff $|a| \cap |r| \neq \emptyset$.

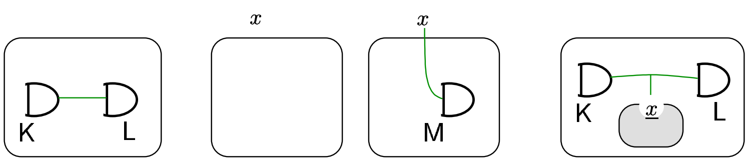
For the inductive step, we must consider transitions of $b = 1 \otimes a$ and of $b = x \otimes a$. We shall treat only the former, since the second is simpler. So suppose $b = 1 \otimes a \xrightarrow{G}_{\mu} b'$, with underlying IPO G, E for b, r, where b and r have widths $1 + \ell$ and m. From Proposition 79 there are two cases of how this arises from an IPO for a, r:

Case $G, E = \mathsf{id}_1 \otimes F,\ 1 \otimes D$, where F, D is an IPO for a, r: Then the latter underlies a transition $a \xrightarrow{F}_{\lambda} a'$ for which we may assume the proposition as inductive hypothesis. We need to show that $\mathsf{width}(G)(1+\ell) \cap \mu \neq \emptyset$ iff $|1 \otimes a| \cap |r| \neq \emptyset$. Now $\mu = \mathsf{width}((1 \otimes D))(m)$, so $\mathsf{width}(G)(1+\ell) \cap \mu = 1 \cdot (\mathsf{width}(F)(\ell) \cap \lambda)$; also $|1 \otimes a| = |a|$. The required result then follows from the inductive hypothesis.

Case $G, E = \eta_1 F, D$, where F, D is an IPO for a, r: Again this IPO underlies a transition of a for which we may assume the proposition. The inductive argument is similar to the previous case, using the fact that $\mathsf{width}(\eta_1 F)(1 \otimes \ell) = \mathsf{width}(F)(\ell)$. ■

Engaged transitions are a smaller class than non-tensorial transitions; for they exclude not only tensorial transitions but also elided tensorial transitions. It is therefore useful to know when the class of engaged transitions is adequate. Here is a simple example of a BRS in which it is not adequate:

Example 18 (need for disengaged transitions) Take the signature $\mathcal{K} = \{\mathsf{K}, \mathsf{L}, \mathsf{M}\}$, each atomic with arity one, and a single reaction rule $(r, r', 1)$ with $r = \nu x.[\,\mathsf{K}(x) \mid \mathsf{L}(x)\,]$ and $r' = 1$. Then the class E of engaged transitions is not adequate. To see this, let $a = x \mid 1$ and $b = M(x)$.



$r = \nu x.[\,\mathsf{K}(x) \mid \mathsf{L}(x)\,]$ $\quad a = x \mid 1$ $\quad b = \mathsf{M}(x)$ $\quad F = [\,\mathsf{K}(\underline{x}) \mid \mathsf{L}(\underline{x}) \mid -\,]$

First, it is easy to check that $a \otimes n$ and $b \otimes n$ have no E-transitions for any $n \geq 0$, except tensorial ones. Then one may show that $\{(a \otimes n, b \otimes n) \mid n \geq 0\}$ is a E-bisimulation, hence $a \sim_{\mathsf{E}} b$.

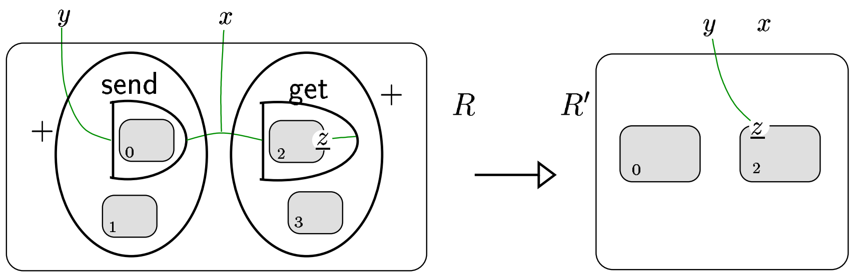
Next, take $F = [\,\mathsf{K}(\underline{x}) \mid \mathsf{L}(\underline{x}) \mid -\,]$. Then a has the disengaged transition $a \xrightarrow{F}_{\{0\}} 1$, while b has no F transition; hence $a \not\sim b$.

Note that this transition is also elisive; so in this BRS any adequate set must contain elisive transitions. ■

The key feature of this example is that the redex has a closed port. Now it is easy to show that if a redex is open and atomic then so is the label of any transition based upon it; we therefore deduce the following immediately from Proposition 83:

Proposition 86 (engaged transitions are adequate for open atomic redexes) *In a BRS in which every redex is open and atomic, the class of engaged transitions is adequate.*

The advantage of being able to limit attention to engaged transitions is considerable. We conjecture that this will also be possible for certain BRSs with parametric reaction rules, as defined in the next subsection.



$$[\,+[\mathsf{send}(y,x)[-_0] \mid -_1] \mid +[\mathsf{get}(x,\underline{z})[-_2] \mid -_3]\,] \longrightarrow {}^{y}\!/\underline{z}.\,x.\,[-_0 \mid -_2]$$

Figure 12: Communication in the synchronous π-calculus

6.5 Parametric reaction and binding

Hitherto we have worked only with *ground* reaction rules (r, r', λ), composed of agents. We have given examples (Examples 2, 3 and 4) of *parametric* rules, but we have avoided defining their meaning. In this final subsection we propose a general definition for a parametric rule (R, R', λ), where R, R' are contexts. The definition is conjectural, because further investigation is needed to see how well it behaves in practice. We finish with a brief discussion of binding, in preparation for further research.

Let us first look at two more parametric rules pertaining to the π-calculus: communication in the synchronous calculus, and replicated input in the asynchronous calculus.

Example 19 (summation and replication in π-calculus) The rule for communication in the (original) synchronous π-calculus [24]

$$(M + \overline{x}\langle y\rangle.P) \mid (N + x(z).Q) \longrightarrow P \mid \{{}^{y}\!/z\}Q$$

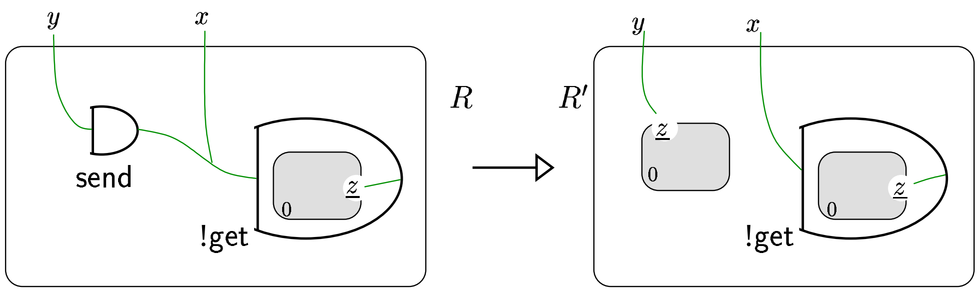
employs summation, and has four parameters of which two are discarded by the reaction and one is subjected to a substitution. To represent this rule, we add $+$ as a non-atomic inactive control to the signature of Example 2. The rule is shown in Figure 12, both as a diagram and in the term language.

Anyone familiar with the π-calculus may be puzzled to see the parallel product, '$|$', appear in the rule under summation. It is not really surprising; the operator is best understood as 'place the processes side-by-side, sharing any names free in both'. In the absence of any surrounding inactive control, this implies '…and let the processes run concurrently', and that is what happens, both here and in the π-calculus. Indeed, many derivatives of the π-calculus – especially the asynchronous calculus – dispense with summation, since its power can usually be simulated by other constructions.

Turning to replication, for simplicity we revert to the asynchronous calculus. Traditionally, replication has been dealt with in two ways: Either the structural congruence rule $P \mid !P \equiv !P$ is introduced, avoiding the need for a special reaction rule for replication, or an inductive transition rule such as

$$\frac{P \mid !P \xrightarrow{\ell} P'}{!P \xrightarrow{\ell} P'}$$

is used, where ℓ is drawn from a given family of labels (including τ to represent reaction). The latter approach is not open to us, since we mean to *derive* labels; the former approach involves imposing a structural



$$[\,\mathsf{send}(y,x) \mid \mathsf{!get}(x,\underline{z})[-_0]\,] \longrightarrow [\,{}^{y}\!/\!_{\underline{z}}\,.\,[-_0] \mid \mathsf{!get}(x,\underline{z})[-_0]\,]$$

Figure 13: Replication in the asynchronous π-calculus

congruence upon bigraphs, which has not been studied. For the present, we shall confine ourselves to replicated input, denoted by an extra control !get, non-atomic and inactive with arity one. A rule for this is shown in Figure 13. This is a simple case of duplicating a factor in the reactum R' of a parametric rule, where the two copies are subjected to different substitutions. The substitution ${}^{y}\!/\!_{\underline{z}}$ in the term for the reactum qualifies only one copy of the site, so cannot be lifted to the top level, since the effect would be to broaden its scope to include two copies. Thus, though our 'normal-form' terms suffice to describe linear contexts, they need slight extension to describe non-linear ones. ■

Now let us look at two general issues to do with parametric rules. The first is to do with linearity. By mentioning a context R with inner width m as a parametric redex we intend that it has a m distinct 'formal parameters', i.e. m holes in which to place the distinct factors of any m-ary agent d with correct interface. In other words, our parametric rules are *left-linear*. But a rule need not be *linear*; we may wish each factor of d to be placed at zero, one or more of the sites of R'. Our three examples in Section 2 are all linear, each factor appearing exactly once in the reactum; but the example of synchronous communication (Figure 12) discards two of its four factors, while the replication example (Figure 13) duplicates its only factor.

The second issue is to do with the *scope* of conames. We can assume that each coname of R is to be used by exactly one factor of the parameter; for if sharing is required, R can use a co-alias. (See the end of this subsection for the connection with the binding BRS $\mathbf{Big}^{\mathsf{b}}_{\mathsf{a}}$ introduced in Subsection 5.4.) having required this discipline, the rule must also cater for of different instantiations of such conames in the reactum.

In fact, a parametric rule must deal with much of the same detail as an analogous rule in term-rewriting, corresponding to the definition of a second-order (recursive) function F. This takes a form such as

$$F(f_0,\ldots,f_{m-1}) \longrightarrow \cdots f_2(E_0,E_1)\cdots f_0(E_2)\cdots f_0(E_3)\cdots ,$$

where the right-hand side is an arbitrary expression which applies several of the function parameters f_i of F, some of them several times, each time with different subexpressions E_j as arguments. Much of the complexity of the following definition is present in such functional rules, and not due to our graphical formalism!

Conjectured definition (parametric reaction rule) A *parametric reaction rule* for $\mathbf{Big}_{\bar{\mathsf{a}}}$ is a triple of the form $(\,R\!:\!I\to J,\ R'\!:\!I'\to J,\ \lambda\,)$ with $\lambda\in\mathsf{plac}(J)$. It generates a set of *ground (reaction) rules*, defined as follows. Let $I = \langle m, X\rangle$ and $I' = \langle m', X'\rangle$; then the rule is equipped with

$\{X_i \mid i \in m\}$	a partition of names, i.e. $X = \biguplus\{X_i \mid i \in m\}$;
$\mathsf{inst} : m' \to m$	an *instantiation map*;
$\sigma_j : X_{\mathsf{inst}(j)} \to X'$	an *instance substitution* for each $j \in m'$.

In terms of these, each ground rule takes the form

$$((\mathsf{id}_H \otimes R) \circ d,\ (\mathsf{id}_H \otimes R') \circ d',\ \lambda)$$

for some interface $H = \langle \ell, W\rangle$. The redex is required to be epi, and the agents $d : \epsilon \to H \otimes I$ and $d' : \epsilon \to H \otimes I'$ take the following form for some agent $e : \epsilon \to Z \otimes H$:

$$\begin{aligned} d &= \nu Z \circ (e \mid d_0 \mid \cdots \mid d_{m-1}) \\ d' &= \nu Z \circ (e \mid d'_0 \mid \cdots \mid d'_{m'-1}) . \end{aligned}$$

Here d is factorised uniquely according to Proposition 74, with ℓ factors comprised in e and factors $d_i : \epsilon \to Z \otimes W \otimes \langle 1, X_i\rangle$ for each $i \in m$; in place of the latter d' has factors $d'_j \simeq \sigma_j \circ d_{\mathsf{inst}(j)}$ for each $j \in m'$. ■

The reader may like to work out how the definition specialises to our proposed rules for the π-calculus (Example 19). In understanding the definition, note these points:

1. In the factorisation of the parameter d, for each $i \in m$ only the factor d_i may use the 'bound' names X_i; thus the intended scoping is respected.
2. For given i there may be more than one copy of d_i used by R'; each has a distinct node set (by $\simeq$), and each is subject to a distinct substitution σ_j ($i = \mathsf{inst}(j)$). This corresponds to the distinct applications of the function parameter f_i in our functional rewrite rule.
3. However, each factor d'_j in the reactum will share both the restricted names Z and the 'free' names W shared by all factors in the redex.
4. The factor e also shares these names; it is included for technical reasons, but can safely be ignored for the present paper.

Let us now look briefly at the relationship between these parametric rules and the binding BRS $\mathbf{Big}^{\mathsf{b}}_{\mathsf{a}}$ introduced in the previous section (Definition 67). Take for example the rule for communication in the synchronous π-calculus (Example 19); it is only applicable when the binding port of the redex's get-node is indeed linked only to ports within the region of that node. But if we work with this rule in $\mathbf{Big}_{\bar{\mathsf{a}}}$ then nothing prevents other get-nodes, apart from those in redexes, from violating this discipline. Therefore we expect a closer correspondence with the behavioural theory of the original π-calculus if we work in $\mathbf{Big}^{\mathsf{b}}_{\mathsf{a}}$. In that case we would naturally take the reaction rules to be the preimages of our rules in $\mathbf{Big}_{\bar{\mathsf{a}}}$ under the forgetful functor from $\mathbf{Big}^{\mathsf{b}}_{\mathsf{a}}$. We leave this question for further research; this will provide a test for our conjectured definition of parametric rules, as well as for the importance of our derived BRS for binding.

7 Related topics and further work

In this concluding section we first discuss some topics connecting this work with other research; we then indicate the most immediate lines along which the present theory should develop in its own terms. Finally we speculate whether and how the theory of concurrency may stabilise.

We begin by commenting on four topics that have been the focus of considerable research, which should be explored in relation to bigraphs: causal dependency, locality, graph rewriting and hierarchy.

Causal dependency In this work we have taken bisimilarity, as originally defined for CCS or the π-calculus, to be typical of the behavioural equivalences and preorders that are guaranteed by our theory to be congruences. Others were treated in Leifer's PhD Dissertation [18]. Hitherto we have only considered behavioural relations which are not sensitive to causal dependency. We would like to extend our study to a wider class of relations; a good account of their variety and taxonomy may be found in van Glabbeek and Plotkin [12].[13]

How can we treat causal dependency in bigraphs, or more abstractly in wide reactive systems? In the π-calculus it can be done by keeping track of the *occurrences* of redexes, noting whether they are disjoint. In WRSs there are two possibilities. First, we may represent the concurrency of two independent transitions by their tensor product; this corresponds closely to allowing concurrent firings in a Petri net, and indeed –as was first pointed out by Mesegeur and Montanari [21]– a Petri net can be well modelled by a monoidal category. Second, we can keep track of successive redex occurrences by means of their *supports*, i.e. their node-sets. A small but important change to Definition 15 would be needed. There, transitions $a \xrightarrow{F}_{\lambda} a'$ are closed under support translation (this was in order to derive transitions for the support quotient WRS in Definition 16); but if we replace $a' \simeq D \circ r'$ by $a' = D \circ r'$ in Definition 15, then we may keep track of the supports of successive reactions and determine which are disjoint. This deserves further study.

Locality In this paper I have avoided giving a technical meaning to the word *location*, because it already has a rich variety of meanings in concurrency. These are thoroughly surveyed by Castellani [3] in the recently published Handbook of Process Algebras. Her broad classification is into *abstract locations* which –crudely simplified– keep track of causal independency (discussed above), and *concrete locations* which form an essential ingredient of behaviour (e.g. they can be manipulated as data). Castellani calls a model *distributed* if it provides concrete locality.

Bigraphs are distributed in this sense. Their concrete locality, just as in action calculi, consists simply of the control nodes and their nesting structure. In Castellani's terms, these nodes are *units of distribution*. When bigraphs are specialised to a particular model by a signature and some reaction rules, nodes with a given control K may be made to function variously as –in Castellani's terms– units of *communication*, of *mobility*, of *security* or of *failure*; it is the reaction rules which determine the role of K-nodes, for each control K. Take failure as an example. Recalling the conjectural definition of parametric reaction in Subsection 6.5, one can simply define an active control failzone (say) with suitable arity, and define reaction rules which model failure by letting a failzone-node discard its entire contents and report the occurrence; alternatively, it may mutate into an inactive node with control stopped (say), which is susceptible to later resuscitation. Turning to mobility, we saw in the ambient reaction rule (Example 3) how an amb-node is made to function as a unit of mobility.

Bigraphs also provide what we may call the *composition* of locality. Sites and places (Definitions 23, 11) are not themselves units of mobility (etc), but they provide means by which the concrete locality of a complex

[13] I avoid here any judgement on the relative merit of equivalences which are causality-sensitive versus those which are not. It is clear that both are important, for different reasons. In particular, the CCS bisimilarity was adopted to allow a concurrent process to be semantically equivalent to a sequential one; thus one can improve the efficiency of a program by making it more concurrent, without changing its semantic equivalence class. This has strong practical and theoretical significance, yet it is precluded by equivalences that preserve causal dependency. There are equally cogent, but different, reasons to study causality-sensitive equivalences.

system can be built up incrementally by categorical composition and tensor product. The notion of *width* is an essential part of this toolkit; for example, it allows us to compose a whole distributed system from parts, each of which may be distributed over that whole.

Graph rewriting There is a well-developed algebraic theory of graph rewriting, especially in relation to concurrency. The third volume of a recent Handbook [7] is devoted to this topic. A prominent part is played by the *double pushout* (DPO) construction introduced by Ehrig [6] for representing rewriting rules, called here reaction rules. This construction works in what we may call an *embedding* category, with graphs as objects and graph morphisms –or embeddings– as arrows. In our terminology, a reaction rule consists of a pair $(\ell : K \to L,\ r : K \to R)$ of embeddings, where K is the common part (often a discrete graph) shared by the redex L and reactum R. A reaction $G \longrightarrow\!\!\triangleright H$ is then defined by a pair of pushouts, as shown.

$$\begin{array}{ccccc} L & \xleftarrow{\ \ell\ } & K & \xrightarrow{\ r\ } & R \\ \downarrow & & \downarrow & & \downarrow \\ G & \longleftarrow & D & \longrightarrow & H \end{array}$$

On the other hand our monographs **Mog** are *contextual* precategories, with sets (of names) as objects and graphs –or graphical contexts– as arrows. A link between these two views of graph rewriting is given by Gadducci and Heckel [9]. Essentially, each arrow (monograph) $G : X \to Y$ of **Mog** is considered as a pair of embeddings of X and Y (considered as discrete graphs) respectively into G; then for $H : Y \to Z$ the composition $H \circ G : X \to Z$ is formed by a pushout of the two embeddings of Y. Another link between the views, at least when embeddings are injective on control nodes, was employed by Cattani, Leifer and Milner [4]; it was shown that a category whose arrows are embeddings is isomorphic to (essentially) the *coslice* category $\epsilon/$**Mog**. (The contextual category differed slightly, but a similar result should hold for **Mog**.) These correspondences suggest linking bigraphs more deeply with the graph-rewriting tradition.

There are considerable differences of purpose, however. A main concern of this work is to derive labelled transition systems, with graph-contexts as labels, in order to study behavioural congruences. For this purpose *relative* pushouts (RPOs) are of prime importance, and we seek these in a contextual –not an embedding– precategory. At first sight, our use of RPOs has little to do with the DPO construction. But our contextual RPOs are shown in [4] to arise from *relative coproducts* in the embedding category, via the coslice isomorphism already mentioned.

Hierarchy Nesting of control nodes was discussed above in connection with locality. There are several notions of structured and/or hierarchical graphs, some of which (with applications) can be found in the already cited Handbook [7]. A very different application from ours, but still in concurrency, is Harel's *statecharts* [13]. A statechart is a generalisation of the transition diagram of an automaton, in which the states have hierarchical structure, represented by the regions of the chart. The difference from bigraphs is fundamental: a bigraph represents a single distributed state, while a statechart represents the dynamic behaviour of a whole system. But we may still ask: How alike are the graphs? A pleasant feature of statecharts is that regions can intersect; for example one structured state of a wristwatch may correspond to the alarm being set, and another to the day registered being Monday; each of these states has many substates, and these two sets of substates have a proper intersection – i.e. a (sub)state does not have a unique immediate superstate.

Then why have we adopted tree-like topographs, by insisting that each point has a unique parent? The meaning is different from in statecharts, but not wholly different. For example, in the special case of ambients, it would make good sense for two ambients to have a proper intersection; one might represent the city of Heraklion, while the other represents the University of Crete (locally known as the 'longest university in the world' and having many campuses in other cities). In other applications of bigraphs, e.g. the π-calculus, it makes less sense for control nodes to intersect. But since bigraphs are a general model, we would like to know if our constraint on topography –adopted to avoid tackling too many problems at once– is necessary. The question may not be hard to answer, since the requirements are quite sharp; for example, are the conditions of a WRS (Definition 12) satisfied if topographs are generalised to directed acyclic graphs?

Variations are possible within the tree-like topography. Drewes, Hoffmann and Plump [5] generalise DPO-based graph rewriting to hypergraphs in which a subgraph may be assigned to each hyperedge.[14] However, no links can exist between the subgraph and its parent; instead, such links may be formed when the subgraph is 'promoted' to top level as a result of a transformation. With this scheme it is not easy to see how to represent the ambient calculus, for example. The same limitation holds for the scheme of Hirsch and Montanari for hyperedge replacement [15]; however, as they show, the scheme permits an encoding of the π-calculus.

Returning to bigraphs, the present work suggests three lines for immediate development:

Matching known calculi The first task is to test bigraph theory against existing calculi, such as the π-calculus and the ambient calculus, to see how our derived transition systems and equivalences match those already known. The first step in each case is to enumerate the derived transitions, and then to look for a subclass of them –such as the engaged transitions (Definition 84)– that constitutes an adequate set (Definition 81). We would hope to find a correspondence between these and the familiar transition systems. If this correspondence is close, then the induced behavioural relations (e.g. bisimilarity) are likely to agree; if not, we should try to characterise these relations in terms of the original definition of the calculus.

For the π-calculus, it is reasonable to expect that we recover the original bisimilarity congruence [24]. The reader may recall that the original bisimilarity *equivalence* is not quite a congruence, but that congruence is obtained by closing up under substitutions. Now we can already observe (see Example 17) that our derived transitions include some whose labels are just substitutions; so the bigraph framework begins to explain what is needed to obtain a congruence *directly*.

Understanding RPOs better Much can be done to improve the RPO theory. At present we only know about the existence of RPOs in $\mathbf{Big}_{\overline{a}}$, which has no aliases. I conjecture that RPOs exist under manageable conditions, in the general case; but I found the proofs –though not easy– to be easier without aliases, and this restriction also makes examples more tractable. With or without aliases, the RPO theory ought to be simpler, but at present I do not see how to make it so. One possibility is a more gradual theory, expressing RPOs and IPOs for compound agents in terms of those for their components – i.e. an inductive presentation, which is markedly absent at present. This could link nicely with the traditional inductive presentation of transition systems.

The reader will have noticed that much of the complexity of IPOs comes from elisions (Definition 78), which in turn arise from non-epi agents – those with idle names or barren roots. We saw that elisions

[14] In standard hypergraph terminology our control nodes and names correspond roughly to hyperedges and nodes respectively; I have preferred the terminology used here partly because control nodes are also nodes of a topograph.

are needed for congruence in these cases, i.e. the non-elisive transitions are not an adequate set (Definition 81, Example 18). But it is worth asking why we *need* non-epi agents; for example, if we could arrange that the epi property is preserved by reaction then we could work entirely in a the sub-BRS with only epis (Example 14). We leave this question open.[15]

Exploring refined BRSs The definition of bigraphs represents a deliberate choice to keep the structures as simple as possible, consistent with the challenge to find a model for mobile interaction that is both rigorous and useful. In bigraphs themselves there are no restrictions upon wiring in relation to locality, arcs are undirected and may branch without constraint, ports are untyped, and so on. In Section 5 we began to show how these constraints may be incorporated into *derived* BRSs, paying special attention to the scoping of names. It remains to be discovered in future work whether this has the effect of recovering the original bisimilarity of the π-calculus in this new setting. Beyond that, derived BRSs, in conjunction with specific signatures and reaction rules, seem to offer a very general way of customising the bigraph model for particular applications while keeping its core theory intact.

Conclusion It is worth reflecting on the timeliness, or otherwise, of trying to find a common conceptual basis for concurrent calculi. The central concepts underlying bigraphs, *connectivity* and *locality*, have become important largely because of technological progress. Computer scientists would not be building these ideas into languages and calculi if, for some reason, we had remained fixed in the era of the freestanding stored-program computer. So we must ask: What comes next? Will some new concept –enabled by technology– burst upon us and render our theoretical efforts obsolete? Do we have any right to believe that the quest for a stable science of concurrent processes will converge?

There are three reasons for optimism. First, both connectivity and locality[16] have been studied for a long time, and their technological importance has increased during that time. So whatever other concepts become important, these two are likely to persist. Second, they are both yielding to mathematical and logical treatment via calculi of one kind or another, and tractable mathematical models lead to stable theories. Finally, because tractable models are good for thinking in (which is their purpose), industrial designers have an incentive to adopt these models in structuring their designs and are beginning to do so; in other words, they begin to use the models to mould the new technology into shape. We can hope for a science of concurrency to stabilise, step by step, through this dialectic between theory and implementation.

Acknowledgements James Leifer has contributed much of the basis for this work in his PhD Dissertation. Philippa Gardner and Peter Sewell have had a strong influence on different parts of the work. I thank them, as well as Tony Hoare, Ole Jensen, Alistair Turnbull and Lucian Wischik, for valuable discussions.

References

[1] Berry, G. and Boudol, G. (1992), The chemical abstract machine. Journal of Theoretical Computer Science, Vol 96, pp217–248.

[2] Cardelli, L. and Gordon, A.D. (2000), Mobile ambients. Foundations of System Specification and Computational Structures, LNCS 1378, pp140–155.

[15] Looking at Example 2 for the π-calculus, we see that the name x, previously idle, becomes idle after the reaction; perhaps, instead, it could be linked to a special kind of dead node? Exploring in this direction, one discovers the danger of an accretion of dead nodes, requiring some sort of quotient to remove them. But does such a quotient preserve RPOs? The plot thickens … .

[16] Some would add causality to this list, though it may be reducible to the other two.

[3] Castellani, I. (2001), Process algebras with localities. *Handbook of Process Algebra* (eds. Bergstra, Ponse and Smolka), Elsevier, Amsterdam, pp947–1045.

[4] Cattani, G.L., Leifer, J.J. and Milner, R. (2000), Contexts and Embeddings for closed shallow action graphs. University of Cambridge Computer Laboratory, Technical Report 496. [Submitted for publication.] Available at `http://pauillac.inria.fr/~leifer`.

[5] Drewes, F., Hoffmann, B. and Plump, D. (2000), Hierarchical graph transformation. In Proc. Conference on Foundations of Software Science and Computation Structures (FoSSACS'00), LNCS 1784, Springer Verlag, 33pp.

[6] Ehrig, H. (1979), Introduction to the algebraic theory of graph grammars. Proc. first International Workshop on Graph Grammars and their Application to Computer Science and Biology, LNCS Vol 73, pp1–69.

[7] Ehrig, H., Kreowski, H-J., Montanari, U. and Rozenberg, G. eds (1999), *Handbook of Graph Grammars and Computing by Graph Transformation, Volume 3: Concurrency, Parallelism and Distribution*, World Scientific.

[8] Fournet, C. and Gonthier, G. (1996), The reflexive Cham and the join calculus. Proc. 23rd Annual ACM Symposium on Principles of Programming Languages, Florida, pp372–385.

[9] Gadducci, F. and Heckel, R. (1998) An inductive view of graph transformation. Proc. 8th Conference on Category Theory in Computer Science (CTCS'99), Vol 29 of Electronic Notes in TCS, Elsevier Science.

[10] Gardner, P.A. (2000), From process calculi to process frameworks. Proc. CONCUR 2000, 11th International Conference on Concurrency theory, pp69–88.

[11] Gardner, P.A. and Wischik, L. (2000), Explicit fusions. Proc. MFCS 2000. LNCS 1893, pp373–383.

[12] van Glabbeek, R. and Plotkin, G. (1995). Configuration Structures. Proceedings 10th Annual IEEE Symposium on Logic in Computer Science (LICS 95), IEEE Computer Society Press, pp199–209.

[13] Harel, D. (1987) Statecharts: a visual formalism for complex systems. Science of Computer Programming, 8:231–274.

[14] Hasegawa, M. (1999) *Models of sharing graphs.* Distinguished Dissertation Series, Springer-Verlag.

[15] Hirsch, D. and Montanari, U. (2001), Synchronised hyperedge replacement with name mobility. Proc. 12th International Conference on Concurrency Theory (CONCUR 2001), LNCS 2154, pp121–136.

[16] Honda, K. (1995), Process Structures. In Proc. TPPP'94, Proceedings of Theory and Practice of Parallel Programming, LNCS 907, Springer-Verlag, pp25–44.

[17] Lafont, Y. (1990), Interaction nets. Proc. 17th ACM Symposium on Principles of Programming Languages (POPL 90), pp95–108.

[18] Leifer, J.J. (2001), Operational congruences for reactive systems. PhD Dissertation, University of Cambridge Computer Laboratory. Distributed in revised form as Technical Report 521.
Available from `http://pauillac.inria.fr/~leifer`.

[19] Leifer, J.J. and Milner, R. (2000), Deriving bisimulation congruences for reactive systems. Proc. CONCUR 2000, 11th International Conference on Concurrency theory, pp243–258. Available at `http://pauillac.inria.fr/~leifer`.

[20] Leifer, J.J. and Milner, R. (2000), Shallow linear action graphs and their embeddings. University of Cambridge Computer Laboratory, Technical Report 508. To appear in a festscrift in honour of Rod Burstall, 2001.

[21] Meseguer, J. and Montanari, U. (1990), Petri nets are monoids. Journal of Information and Computation, Vol 88, pp105–155.

[22] Milner, R. (1996), Calculi for interaction. Acta Informatica 33, pp707–737.

[23] Milner, R. (2001), Bigraphical reactive systems. Proc. 12th International Conference on Concurrency Theory (CONCUR 2001), LNCS 2154, pp16–35.

[24] Milner, R., Parrow, J. and Walker D. (1992), A calculus of mobile processes, Parts I and II. Journal of Information and Computation, Vol 100, pp1–40 and pp41–77.

[25] Parrow, J. and Victor, B. (1998), The fusion calculus: expressiveness and symmetry in mobile processes. Proc. LICS'98, IEEE Computer Society Press.

[26] Petri, C.A. (1962), Fundamentals of a theory of asynchronous information flow. Proc. IFIP Congress '62, North Holland, pp386–390.

[27] Sewell, P. (1998), From rewrite rules to bisimulation congruences. Proc CONCUR'98, LNCS 1466, pp269–284. Full version to appear in Theoretical Computer Science, Vol 272/1–2.

[28] Sewell, P. and Unyapoth, A. (2001), Nomadic Pict: Correct communication infrastructure for mobile computation. Proc. 28th ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages, pp128–141.

[29] Wojciechowski, P.T. and Sewell, P. (1999), Nomadic Pict: Language and infrastructure design for mobile agents. Proc. ASA/MA '99, Palm Springs, California.

[30] Yoshida, N. (1995) Graph Notation for Concurrent Combinators. In Proc. TPPP'94, Proceedings of Theory and Practice of Parallel Programming, LNCS 907, Springer-Verlag, pp393–412.

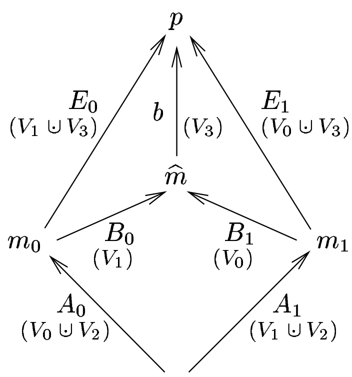
APPENDICES

A RPOs and IPOs in topographs

In this appendix we first prove that RPOs always exist in the category **Top** of topographs, as required for Subsection 4.1. Then we enumerate all the IPOs for any pair $\vec{A}$ of topographs with common domain.

Let $\vec{A} : m \to \vec{m}$ and $\vec{E} : \vec{m} \to p$ be two pairs of topographs, with $\vec{E}$ a bound for $\vec{A}$. Then the nodes may be partitioned into four sets: V_i common to A_i and $E_{\bar{\imath}}$ $(i = 0, 1)$, V_2 common to A_0 and A_1 and V_3 common to E_0 and E_1. From now on we shall consistently use $v_i, v'_i, \ldots$ to range over V_i $(i = 0, 1, 2, 3)$. We shall also take a simplifying liberty; we shall just treat the case in which $m = 0$, since the sites in m are treated by the construction exactly as members of V_2. It comes to the same thing if we think of $v_2, v'_2 \ldots$ ranging over $m \uplus V_2$.

We now seek an RPO $(\vec{B}, B)$ for $\vec{A}$ w.r.t. $\vec{E}$. We begin by constructing $B_i : m_i \to \widehat{m}$ $(i = 0, 1)$.



First we fix the node set of B_i to be $V_{\bar{\imath}}$, as shown. Next, we have to define the multiplicity $\widehat{m}$, and we do this in two stages. First, we select those roots in m_i which must be "exported" via $\widehat{m}$, because their parents in E_i lie outside $V_{\bar{\imath}}$:

$$m'_i \triangleq \{r \in m_i \mid E_i(r) \in V_3 \uplus p\} .$$

These will provide the members of $\widehat{m}$; furthermore, in $\widehat{m}$ we shall equate any pair of them which are parents of the same node in V_2. So our second stage is to define an equivalence on $m'_0 + m'_1$ and divide by it to obtain $\widehat{m}$:

$$\begin{aligned} \sim \; &\triangleq \; \{(\langle 0, r_0\rangle, \langle 1, r_1\rangle) \mid A_i(v_2) = r_i \in m'_i \ (i = 0, 1) \text{ for some } v_2\}^{\equiv} \\ \widehat{m} \; &\triangleq \; (m'_0 + m'_1)/\sim \ . \end{aligned}$$

We are now ready to define the legs $\vec{B}$ of our RPO. It helps to think of B_0, for example, as follows; on nodes in V_1 it imitates A_1, while on sites in m_0 it imitates E_0 as far as it can.

Construction 87 (RPO legs) We define B_0 (B_1 is similar):

$$\begin{aligned} B_0(v_1) \; &\triangleq \; \left\{ \begin{array}{ll} w & \text{if } w \in V_1 \\ [\langle 1, w\rangle]_\sim & \text{if } w \in m_1 \end{array} \right\} \quad \text{where } w = A_1(v_1) \\ B_0(r_0) \; &\triangleq \; \left\{ \begin{array}{ll} w & \text{if } w \in V_1 \\ [\langle 0, r_0\rangle]_\sim & \text{if } w \in V_3 \uplus p \end{array} \right\} \quad \text{where } w = E_0(r_0) \ . \end{aligned}$$ ■

Note that this definition is good. For in the first case the equation $E_0 \circ A_0 = E_1 \circ A_1$ ensures that if $w = A_1(v_1)$ is in m_1 then $E_1(w) \in V_3 \uplus p$, hence $w \in m'_1$; in the second case the condition $w \in V_3 \uplus p$ similarly ensures $r_0 \in m'_0$. It is also easy to check that no atomic node has become a parent in B_0, assuming that both A_1 and E_0 have this property.

We now prove the first property required of an RPO.

Proposition 88 (RPO legs are a bound) $B_0 \circ A_0 = B_1 \circ A_1$.

Proof Let $E \triangleq E_0 \circ A_0 = E_1 \circ A_1$. The possible arguments v for the parent function of $B_i \circ A_i$ lie in $v \in V_0 \uplus V_1 \uplus V_2$. Consider cases (omitting symmetric ones) for the pair (v, w), where $w = E(v)$:

Case $v \in V_0$, $w \in V_0$. Then $w = A_0(v_0)$, so $(B_1 \circ A_1)(v) = B_1(v) = w = A_0(v) = (B_0 \circ A_0)(v)$.

Case $v \in V_0$, $w \in V_1 \uplus V_2$. Contradicts $E_0 \circ A_0 = E_1 \circ A_1$.

Case $v \in V_0$, $w \in V_3 \uplus p$. Then for some $r_0 \in m_0$ we have $A_0(v) = r_0$ and $E_0(r_0) = w$, whence $(B_0 \circ A_0)(v) = B_0(r_0) = [\langle 0, r_0\rangle]_\sim = B_1(v) = (B_1 \circ A_1)(v)$.

Case $v \in V_2$, $w \in V_0$. Then $(B_0 \circ A_0)(v) = A_0(v) = w$; also for some $r_1 \in m_1$ we have $A_1(v) = r_1$ and $E_1(r_1) = w$, whence $(B_1 \circ A_1)(v) = B_1(r_1) = w$.

Case $v \in V_2$, $w \in V_2$. Then $A_0(v) = w = A_1(v)$, whence $(B_0 \circ A_0)(v) = w = (B_1 \circ A_1)(v)$.

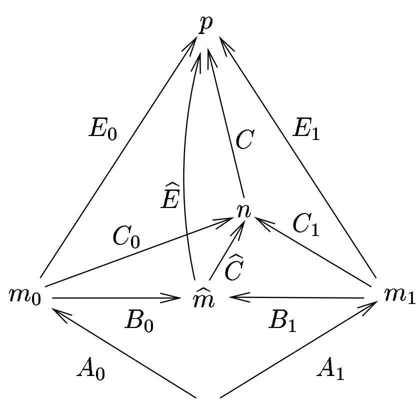
Case $v \in V_2$, $w \in V_3 \uplus p$. Then $A_i(v) = r_i$ $(i = 0, 1)$ where $\langle 0, r_0\rangle \sim \langle 1, r_1\rangle$, so $(B_0 \circ A_0)(v) = (B_1 \circ A_1)(v) = [\langle i, r_i\rangle]_\sim$. ■

For later use we record an important property of $\vec{B}$:

Proposition 89 (RPO legs keep an epi) *If A_1 is epi then B_0 is epi.*

Proof Consider $r \in \widehat{m}$; there exists either $s_0 \in m'_0$ with $\langle 0, s_0\rangle \in r$ and $r = B_0(s_0)$, or $s_1 \in m'_1$ with $\langle 1, s_1\rangle \in r$. In the latter case, as A_1 has no barren roots, either $s_1 = A_1(v_1)$ for some $v_1 \in V_1$, in which case by construction $r = B_0(v_1)$, or $s_1 = A_1(v_2)$ for some $v_2 \in V_2$. In the latter case, since $s_1 \in m'_1$, there exists $s_0 = A_0(v_2) \in m'_0$ with $\langle 0, s_0\rangle \in r$, and again $r = B_0(s_0)$. Thus x is not barren in B_0. ■

Now, instead of directly defining the head B of our RPO, we shall consider any candidate $(\vec{C}, C)$, i.e. a triple of topographs such that $C_0 \circ A_0 = C_1 \circ A_1$ and $C \circ C_i = E_i$ $(i = 0, 1)$, and define a mediating topograph $\widehat{C}$ such that $\widehat{C} \circ B_i = C_i$ $(i = 0, 1)$.



Construction 90 (RPO mediator) Let the triple $(\vec{C}, C)$ be such that $C_0 \circ A_0 = C_1 \circ A_1$ and $C \circ C_i = E_i$ $(i = 0, 1)$, with $C_i : m_i \to n$ $(i = 0, 1)$ and $C : n \to p$. Let $V_3 = V_4 \uplus V_5$, where $V_{\bar{\imath}} \uplus V_4$ is the node set of C_i $(i = 0, 1)$ and V_5 the node set of C. Then the *mediator* $\widehat{C} : \widehat{m} \to n$ is defined as follows, noting that its parent function is $\widehat{C} : \widehat{m} \uplus V_4 \to V_4 \uplus n$:

$$\begin{array}{rcll} \widehat{C}(v_4) & \triangleq & C_i(v_4) & \text{for } v_4 \in V_4 \ (i = 0, 1) \\ \widehat{C}(s) & \triangleq & C_i(r_i) & \text{for } \langle i, r_i \rangle \in s \in \widehat{m} \ (i = 0, 1) \ . \end{array}$$

This definition is good, for we have $C_0(v_4) = (C_0 \circ A_0)(v_4) = (C_1 \circ A_1)(v_4) = C_1(v_4)$, and also if $\langle 0, r_0 \rangle \sim \langle 1, r_1 \rangle$ with $r_i = A_i(v_2)$ $(i = 0, 1)$ then $C_0(r_0) = (C_0 \circ A_0)(v_2) = (C_1 \circ A_1)(v_2) = C_1(r_1)$. It is also easy to check that no atomic node has become a parent in $\widehat{C}$. ■

Proposition 91 (RPO mediator respects bounds) *If $\widehat{C}$ is a mediator for the candidate $(\vec{C}, C)$, as defined in Construction 90, then $\widehat{C}$ is the unique topograph F for which $F \circ B_i = C_i$ $(i = 0, 1)$.*

Proof Unicity is simple: the definition of $\widehat{C}$ is determined by the condition $F \circ B_i = C_i$ since $(F \circ B_i)(v_4) = \widehat{C}(v_4)$, and also if $\langle i, r_i \rangle \in s \in \widehat{m}$ then $(F \circ B_i)(r_i) = F(s)$.

We now check that $(\widehat{C} \circ B_0)(w) = C_0(w)$ for all $w \in m_0 \uplus V_1 \uplus V_4$. We have already proved the equation when $w \in m'_0 \uplus V_4$, so two cases remain:

Case $w = r_0 \in m_0 \setminus m'_0$. Then $E_0(r_0) \in V_1$, so using $C \circ C_0 = E_0$ we have $(\widehat{C} \circ B_0)(r_0) = B_0(r_0) = E_0(r_0) = C_0(r_0)$.

Case $w = v_1 \in V_1$. If $E_0(v_1) \in V_1$ we argue as in the previous case. Otherwise $E_0(v_1) \in V_3 \uplus p$, so $B_0(v_1) = \langle 1, A_1(v_1) \rangle$ and we have $(\widehat{C} \circ B_0)(v_1) = \widehat{C}\langle 1, A_1(v_1) \rangle = C_1(A_1(v_1)) = (C_1 \circ A_1)(v_1) = (C_0 \circ A_0)(v_1) = C_0(v_1)$. ■

Now we observe that, trivially, $(\vec{E}, \mathsf{id}_p)$ is a candidate. Hence

Corollary 92 (RPO head respects bounds) *Let $\widehat{E} : \widehat{m} \to p$ be the special case of $\widehat{C}$, as defined in Construction 90, in which $n = p$, $C_i = E_i$ and $C = \mathsf{id}_p$, so that*

$$\begin{array}{rcll} \widehat{E}(v_3) & \triangleq & E_i(v_3) & \textit{for } v_3 \in V_3 \ (i = 0, 1) \\ \widehat{E}(s) & \triangleq & E_i(r_i) & \textit{for } \langle i, r_i \rangle \in s \in \widehat{m} \ (i = 0, 1) \ . \end{array}$$

Then $\widehat{E} \circ B_i = E_i$ $(i = 0, 1)$.

(Of course we can also claim that $\widehat{E}$ is unique with this property, but we do not need to do so.) We are now ready to complete the definition of our RPO:

Construction 93 (RPO head) An RPO $(\vec{B}, B)$ for $\vec{A}$ w.r.t. $\vec{E}$ is defined by taking $\vec{B}$ as in Construction 87 and setting $B \triangleq \widehat{E}$. ■

It remains to complete the proof of the following:

Proposition 94 (unique RPO mediator) *For any candidate RPO $(\vec{C}, C)$, the mediator $\widehat{C}$ is the unique F such that $F \circ B_i = E_i$ $(i = 0, 1)$ and $C \circ F = B$.*

Proof We know already that $\widehat{C}$ is unique such that $\widehat{C} \circ B_i = E_i$; recalling our definition of B, it only remains to prove that $C \circ \widehat{C} = \widehat{E}$. Consider the possible arguments $w \in \widehat{m} \uplus V_3$:

Case $w = s \in \widehat{m}$. Then say $\langle 0, r_0 \rangle \in s$ with $r_0 \in m_0'$, so $E_0(r_0) \in V_3 \uplus p$; moreover $\widehat{C}(s) = C_0(r_0)$. We now distinguish two cases, and use $E_0 = C \circ C_0$. If $E_0(r_0) \in V_4$ then $E_0(r_0) = (C \circ C_0)(r_0) = C_0(r_0) = \widehat{C}(s)$, so $(C \circ \widehat{C})(s) = \widehat{C}(s) = E_0(r_0) = \widehat{E}(s)$. On the other hand if $E_0(r_0) \in V_5 \uplus p$ then $E_0(r_0) = C(C_0(r_0))$ where $C_0(r_0) = \widehat{C}(s) \in n$, so $(C \circ \widehat{C})(s) = C(\widehat{C}(s)) = E_0(r_0) = \widehat{E}(s)$.

Case $w = v_4 \in V_4$. Then $(C \circ \widehat{C})(v_4) = \widehat{C}(v_4) = C_0(v_4) = (C \circ C_0)(v_4) = E_0(v_4) = \widehat{E}(v_4)$.

Case $w = v_5 \in V_5$. Then $(C \circ \widehat{C})(v_5) = C(v_5) = (C \circ C_0)(v_5) = E_0(v_5) = \widehat{E}(v_5)$. ■

We have therefore established

Theorem 95 (topographs always have RPOs) *Let $\vec{A} : m \to \vec{m}$ and $\vec{E} : \vec{m} \to p$ be two pairs of topographs, with $\vec{E}$ a bound for $\vec{A}$. Then the triple $(\vec{B}, B)$ defined in Constructions 87 and 93 is an RPO for $\vec{A}$ w.r.t. $\vec{E}$.*

For practical purposes we need to go further. Given a pair $\vec{A}$ with common domain, we would like a construction which yields all its IPOs; then when A_1 is a redex, we shall know all the labelled transitions of A_0. It turns out that there is a finite, usually small, number of IPOs for given consistent pair $\vec{A}$. In fact, if both components are epi – i.e. they have no barren roots – then there is exactly one, which is therefore not only an IPO but a pushout. In general though, at least one of the pair may have barren roots, and we find that these alone lead to the existence of a larger family of IPOs.

Now, recall the three consistency conditions defined in Definition 29. Under these conditions, we now construct one or more IPOs for $\vec{A}$. The construction rests on (and justifies) the claim that the IPO $\vec{D}$ is fully determined by $\vec{A}$, except for the value to which D_i maps the barren roots of A_i; indeed, it may map a barren root to any $v_{\bar{\imath}} \in V_{\bar{\imath}}$, or to a root of D_i.

Construction 96 (IPOs in topographs) Assume the consistency conditions (C1) - (C3) for the pair of topographs $\vec{A} : h \to \vec{m}$. We define a family of IPOs for $\vec{A}$ as follows.

For $i = 0, 1$ choose any subset $\ell_i \subseteq m_i$ whose members are all barren in A_i. Set $k_i = m_i \backslash \ell_i$; then define k_i' and an equivalence over $k_0' + k_1'$ by

$$\begin{array}{rcl} k_i' & \triangleq & \{ r \in k_i \mid \forall v_2.\ A_i(v_2) = r \Rightarrow A_{\bar{\imath}}(v_2) \in m_{\bar{\imath}} \} \\ \approx & \triangleq & \{ (\langle 0, r_0 \rangle, \langle 1, r_1 \rangle) \mid \exists v_2.\ A_i(v_2) = r_i \in k_i' \text{ for } i = 0, 1 \}^{\equiv} . \end{array}$$

Take $n \triangleq (k_0' + k_1')/\approx$. Choose two functions $\eta_i : \ell_i \to V_{\bar{\imath}}$, arbitrary except that $\eta_i(r)$ is non-atomic for all $r \in \ell_i$. Then define topographs $D_i : m_i \to n$ with node sets $V_{\bar{\imath}}$, whose parent functions extend the chosen functions η_i as follows (we give D_0; D_1 is similar):

$$\begin{array}{rcl} D_0(v_1) & \triangleq & \left\{ \begin{array}{ll} w & \text{if } w \in V_1 \\ [\langle 1, w \rangle]_{\approx} & \text{if } w \in m_1 \end{array} \right\} \text{ where } w = A_1(v_1) \\ D_0(r_0) & \triangleq & \left\{ \begin{array}{ll} A_1(v_2) & \text{if } r_0 \in k_0 \backslash k_0', \text{ for some } v_2 \text{ with } A_0(v_2) = r_0 \\ [\langle 0, r_0 \rangle]_{\approx} & \text{if } r_0 \in k_0' \\ \eta_0(r_0) & \text{if } r_0 \in \ell_0 . \end{array} \right. \end{array}$$

This defines an IPO $\vec{D} : \vec{m} \to n$ for every chosen quadruple $(\ell_i,\ \eta_i : \ell_i \to V_{\bar{\imath}})$ $(i = 0, 1)$. ■

The functions η_i are called *elisions* in Definition 78. This refers to the fact that the barren roots l_i in A_i are not exported in the IPO interface n, but instead mapped into the body of B_i.

To validate this definition we first have to show that the definition of D_0 is good and unambiguous. Thus in the second clause for $D_0(v_1)$ we must ensure that $w \in k'_1$, and in the first clause for $D_0(r_0)$ we must ensure that a v_2 exists such that $A_0(v_2) = r_0$, and that each such v_2 yields the same value $A_1(v_2)$ in V_1. The consistency conditions do ensure this; they also ensure that $D_0 \circ A_0 = D_1 \circ A_1$, so they are indeed sufficient for consistency. The main validation lies in the following theorem.

Theorem 97 (characterizing IPOs for topographs) *A pair $\vec{D} : \vec{m} \to n$ is an IPO for $\vec{A} : h \to \vec{m}$ iff it is generated by Construction 96.*

Proof (outline) ($\Rightarrow$) Take $\vec{E} = \vec{D}$ in Construction 87 which constructs the legs of an RPO $\vec{B}$ for $\vec{A}$ w.r.t. $\vec{E}$; then show that $\vec{B} = \vec{D}$, up to isomorphism of their codomains.

($\Leftarrow$) Take the legs $\vec{B}$ of an arbitrary RPO for $\vec{A}$ w.r.t. $\vec{E}$, as constructed in Construction 87; then apply Construction 96 with the choice

$$
\begin{aligned}
\ell_i &\triangleq \{r \in m_i \mid r \text{ barren in } A_i,\ E_i(r) \in V_{\bar{i}}\} \\
\eta_i : \ell_i \to V_{\bar{i}} &\triangleq E_i \restriction \ell_i \ ;
\end{aligned}
$$

then show that the construction produces $\vec{D}$ equal to $\vec{B}$ up to isomorphism. ∎

B IPOs in edge nets

In this appendix we characterise all the IPOs in $\mathbf{Edg}_{\bar{\mathsf{a}}}$, as required for Subsection 4.2. Section B.1 contains general definitions and properties; Section B.2 derives a pushout for any consistent pair $\vec{A}$ of epimorphisms; Section B.3 characterises the IPOs for any consistent pair.

B.1 Basic properties

We recall from Section 4 the definitions of edge nets and their precategory, Definitions 36 and 37. As in Subsection 4.2 we examine an arbitrary pair $\vec{A} : W \to \vec{X}$, and we adopt the conventions stated there; in particular we assume w.l.o.g that W is empty, since the constructions here treat any conames in W exactly as if they were shared inner ports.

In Subsection 4.2 we asserted that the consistency conditions (O1) and (O2) of Definition 43 are necessary for $\vec{A}$ to have a bound; we now give a construction which yields a bound under those conditions.

Construction 98 (bound) Given $\vec{A} : \emptyset \to \vec{X}$, we define a bound $\vec{D} : \vec{X} \to Y$ as follows:

The inner port set of D_i will be $P_{\bar{\imath}}$.

The set of names in X_i which are be exported via the bound's co-domain Y is defined as follows:

$$\widehat{X}_i \triangleq \{x_i \mid x_i A_i p_2 \Rightarrow p_2 \text{ open in } A_{\bar{\imath}}\} .$$

We define an equivalence relation over $\widehat{X}_0 + \widehat{X}_1$, and take Y to be its quotient:

$$\begin{aligned} \sim \; &\triangleq \; \{(x_0, x_1) \mid x_0 A_0 p_2 A_1 x_1 \text{ for some } p_2\}^{\equiv} \\ Y \; &\triangleq \; (\widehat{X}_0 + \widehat{X}_1)/\sim \; . \end{aligned}$$

Then we extend $\sim$ to $\widehat{X}_0 + \widehat{X}_1 + Y$ by declaring that $x_i \sim [x_i]_\sim$.

Finally the parents functions of $\vec{D}$ are defined by

$$D_i \triangleq (A_0 \sqcup A_1 \sqcup \sim) \backslash X_{\bar{\imath}} \backslash P_i \backslash P_2 \; . \qquad \blacksquare$$

Before justifying this construction we prove an important lemma about it:

Lemma 99 *Assume $\vec{A}$ satisfies (O1) and (O2). If $x_0 A_0 p_2 A_1 x_1$, then $x_0 \in \widehat{X}_0$, $x_1 \in \widehat{X}_1$ and $x_0 \sim x_1$.*

Proof Note that p_2 open in A_1. Suppose $x_0 \notin \widehat{X}_0$; then for some p_2' we have $x_0 A_0 p_2'$ and p_2' closed in A_1. Then $p_2 A_0 p_2'$, so (O2) implies that p_2, p_2' are either both open or both closed in A_1, yielding a contradiction. Hence $x_0 \in \widehat{X}_0$; similarly $x_1 \in \widehat{X}_1$, and so $x_0 \sim x_1$ by definition. ■

We now validate our construction under the consistency conditions:

Proposition 100 (consistency in edge nets) *In both* **Edg** *and* $\mathbf{Edg}_{\bar{\mathsf{a}}}$*, the pair $\vec{A}$ is consistent iff the conditions (O1) and (O2) hold.*

Proof The necessity of the conditions is straightforward; they can easily be deduced using the equation $E_0 \circ A_0 = E_1 \circ A_1$ where $\vec{E}$ is any bound for $\vec{A}$. For sufficiency, we shall show that the conditions imply that $D_0 \circ A_0 = D_1 \circ A_1$ for the pair $\vec{D}$ defined in Construction 98.

By symmetry it will be enough to prove $D_0 \circ A_0 = \widehat{A}$, where $\widehat{A} \triangleq (A_0 \sqcup A_1 \sqcup \sim)\setminus X_0 \setminus X_1$. Now from definitions and a simple manipulation of restriction we have

$$D_0 \circ A_0 = (A_0 \sqcup (A_0 \sqcup A_1)\setminus P_0 \setminus P_2 \sqcup \sim)\setminus X_0 \setminus X_1 \ .$$

Abbreviate $(A_0 \sqcup A_1)\setminus P_0 \setminus P_2$ to $\overline{A}_1$. Noting that $\overline{A} \subseteq A_0 \sqcup A_1$, it immediately follows that $D_0 \circ A_0 \subseteq \widehat{A}$.

It remains to show the converse, that $(A_0 \sqcup A_1 \sqcup \sim)\setminus X_0 \setminus X_1 \subseteq (A_0 \sqcup \overline{A}_1 \sqcup \sim)\setminus X_0 \setminus X_1$. So, from a lub chain for the former, we must eliminate all A_1-links in favour of instances of A_0, $\overline{A}_1$ and $\sim$.

Case $x_1 A_1 p_1$. Immediately $x_1 \overline{A}_1 p_1$.

Case $x_1 A_1 p_2$. Then there must be a left adjacent instance of $\sim$, hence $x_1 \in \widehat{X}_1$. So p_2 is open in A_0, say $p_2 A_0 x_0$, and by Lemma 99 we deduce a replacement $x_1 \sim x_0 A_0 p_2$.

Case $p_1 A_1 p_1'$. Immediately $p_1 \overline{A}_1 p_1$.

Case $p_1 A_1 p_2$. By (O1) we have p_2 open in A_0, say $p_2 A_0 x_0$, and we deduce $p_1 \overline{A}_1 x_0 A_0 p_2$.

Case $p_2 A_1 p_2'$. By (O2) we have p_2, p_2' open in A_0, say $p_2 A_0 x_0$ and $p_2' A_0 x_0$; hence $p_2 A_0 x_0 \overline{A}_1 x_0' A_0 p_2'$.

This completes the proof that $\vec{D}$ is a bound for $\vec{A}$, at least in **Edg**. But it is straightforward to prove that if $\vec{A}$ has no aliases then $\vec{D}$ has no aliases; hence it is a bound in $\mathbf{Edg}_{\bar{a}}$ also. ∎

So far we have looked at both **Edg** and $\mathbf{Edg}_{\bar{a}}$. We end this section with a useful property of composition in $\mathbf{Edg}_{\bar{a}}$.

Lemma 101 *Let $A : W \to X$ and $B : X \to Y$ be edge nets in $\mathbf{Edg}_{\bar{a}}$ with disjoint port-sets P and Q. Let p, p' range over $W + P$, x, x' over X and q, q' over $Q + Y$. Then*

$$\begin{array}{lrcl}
(1) & p(B \circ A)q & \Rightarrow & pAxBq \text{ for some } x \in X \\
(2) & p(B \circ A)p' & \Rightarrow & pAp' \text{ or } pAxBx'Ap' \text{ for some } x, x' \in X \\
(3) & q(B \circ A)q' & \Rightarrow & qBq' \ .
\end{array}$$

B.2 IPOs for epis in $\mathbf{Edg}_{\bar{a}}$

In this section we shall deal with the case of a consistent pair $\vec{A}$ in which both components are epis; in fact we shall find that it possesses a unique IPO, i.e. a pushout.

Theorem 102 (pushouts for epis in edge nets) *If $\vec{A}$ is consistent pair of epis in $\mathbf{Edg}_{\bar{a}}$, then the pair $\vec{D}$ defined in Construction 98 is a pushout for $\vec{A}$.*

Proof Let $\vec{E} : \vec{X} \to Z$ be any bound for $\vec{A}$ in $\mathbf{Edg}_{\bar{a}}$. We must show that there exists a unique arrow $E : Y \to Z$ such that $E \circ D_i = E_i$ $(i = 0, 1)$.

Suppose that E_i has inner ports $P_{\bar{i}} \uplus P_3$; then we define E with inner ports P_3 as follows:

$$\begin{array}{rcl}
\widehat{E}_i & \triangleq & E_0 \upharpoonright (P_3 \uplus \widehat{X}_0) \text{ (i=0,1)} \\
\widehat{E}_2 & \triangleq & \{(x_0, x_1) \in \widehat{X}_0 \times \widehat{X}_1 \mid \exists p_i.\, x_i A_i p_i E_{\bar{i}} x_{\bar{i}} \text{ for some } i \in \{0,1\}\}^{\equiv} \\
E & \triangleq & (\widehat{E}_0 \sqcup \widehat{E}_1 \sqcup \widehat{E}_2 \sqcup \sim)\setminus X_0 \setminus X_1 \ .
\end{array}$$

We first record a few useful facts, some of which depend on the fact that no x_i is idle in A_i:

$$\begin{array}{ll} (1) & p_i A_i x_i \Rightarrow x_i \in \widehat{X}_i \\ (2) & p_3 E_i x_i \Rightarrow x_i \in \widehat{X}_i \\ (3) & p_3 E_i p'_3 \Rightarrow p_3 E_{\bar{\imath}} p'_3 \\ (4) & x_i \sim x'_i \Rightarrow x_i E_i x'_i \\ (5) & x_i E_i x'_i \Rightarrow (x_i \in \widehat{X}_i \Leftrightarrow x'_i \in \widehat{X}_i) \ . \end{array}$$

To prove $E \circ D_0 \subseteq E_0$, first note that by manipulation of restrictions

$$E \circ D_0 \ = \ (\widehat{E}_0 \sqcup \widehat{E}_1 \sqcup \widehat{E}_2 \sqcup \widehat{A}_0 \sqcup \widehat{A}_1 \sqcup \sim) \backslash P_0 \backslash P_2 \backslash X_1 \backslash Y \ .$$

So from a lub chain for this relation we must transform all links except instances of $\widehat{E}_0$. Most transformations are simple. We indicate whenever they use one of the above facts, or non-idleness, but we do not flag uses of the equation $E_0 \circ A_0 = E_1 \circ A_1$.

We begin by eliminating $\widehat{E}_2$-links. It is enough to consider the two generating instances, as follows:

Case $x_0 A_0 p_0 E_1 x_1$. Since x_1 is not idle, either $x_1 A_1 p_1$ and using Lemma 99 we deduce $\boldsymbol{x_0 E_0 x'_0 \sim x_1}$, or $x_1 A_1 p_2$ and we deduce $\boldsymbol{x_0 E_0 p_1 A_1 x_1}$.

Case $x_0 E_0 p_1 A_1 x_1$. No replacement needed.

We now have a lub chain in $\{E_0, \widehat{E}_1, A_0, A_1, \sim\}$. Eliminating $\widehat{E}_1$-links:

Case $p_3 \widehat{E}_1 p'_3$. From (3) deduce $\boldsymbol{p_3 E_0 p'_3}$.

Case $p_3 \widehat{E}_1 x_1$. Since x_1 is not idle, for some $p \in P_0 \uplus P_2$ deduce $\boldsymbol{p_3 E_0 x_0 A_0 p A_1 x_1}$.

Case $x_1 \widehat{E}_1 x'_1$. Since x_1, x'_1 are not idle, we have $x_1 A_1 p$ and $x'_1 A_1 p'$ for some $p, p' \in P_1 \uplus P_2$. There are several cases; with the help of Lemma 99 deduce $\boldsymbol{x_1 A_1 p_1 E_0 p'_1 A_1 x'_1}$, or $\boldsymbol{x_1 A_1 p_1 E_0 x_0 \sim x'_1}$, or $\boldsymbol{x_1 \sim x_0 E_0 x_0 \sim x'_1}$, or $\boldsymbol{x_1 A_1 p_2 A_0 p'_2 A_1 x'_1}$.

We now have an lub chain for $(E_0 \sqcup A_0 \sqcup A_1 \sqcup \sim) \backslash P_0 \backslash P_2 \backslash X_1 \backslash Y$. In eliminating A_1-links, we have to consider possible adjacent links:

Case $p_2 A_1 p'_2$. Deduce $\boldsymbol{p_2 A_0 p'_2}$ or $\boldsymbol{p_2 A_0 x_0 E_0 x'_0 A_0 p'_2}$.

Case $p_2 A_1 p_1$. Deduce $\boldsymbol{p_2 A_0 x_0 E_0 p_1}$.

Case $p_2 A_1 x_1$. There must be a left adjacent A_0-link; if it is $x_0 A_0 p_2$ then by Lemma 99 replace $x_0 A_0 p_2 A_1 x_1$ by $\boldsymbol{x_0 \sim x_1}$, and if it is $p'_2 A_0 p_2$ then replace $p'_2 A_0 p_2 A_1 x_1$ by $\boldsymbol{p'_2 A_0 x_0 \sim x_1}$.

Case $p_1 A_1 p'_1$. Deduce $\boldsymbol{p_1 E_0 p'_1}$.

Case $p_1 A_1 x_1$. There must be a right adjacent link $x_1 \sim q$, where also $x_1 A_1 p_2 A_0 x_0 \sim q$. Then replace $p_1 A_1 x_1 \sim q$ by $\boldsymbol{p_1 E_0 x_0 \sim q}$.

We now have a lub chain for $(E_0 \sqcup A_0 \sqcup \sim) \backslash P_0 \backslash P_2 \backslash X_1 \backslash Y$. But no A_0-links can remain; for it cannot be an alias, and otherwise it involves P_0 or P_2 so cannot be an end link, and no neighbouring A_0- or $\sim$-link can involve P_0 or P_2. We therefore have a lub chain for $(E_0 \sqcup \sim) \backslash X_1 \backslash Y$. But the support of $\sim$ is $X_0 + X_1 + Y$, and E_0 has support disjoint from X_1 and Y; so the only possible $\sim$-link has the form $x_0 \sim x'_0$, from which by (4) we deduce $\boldsymbol{x_0 E_0 x'_0}$. This completes the proof that $E \circ D_0 \subseteq E_0$.

To prove $E_0 \subseteq E \circ D_0$, we have to deduce from each instance of E_0 a lub chain for $(\widehat{E}_0 \sqcup \widehat{E}_1 \sqcup \widehat{E}_2 \sqcup \widehat{A}_0 \sqcup \widehat{A}_1 \sqcup \sim) \backslash P_0 \backslash P_2 \backslash X_1 \backslash Y$.

Case $p_3 E_0 p'_3$. Then $\boldsymbol{p_3 \widehat{E}_0 p'_3}$.

Case $p_3 E_0 p_1$. Then $p_3 E_1 x_1 A_1 P_1$, and $x_1 \in \widehat{X}_1$ by (2), so $\boldsymbol{p_3 \widehat{E}_1 x_1 A_1 P_1}$.

Case $p_3 E_0 x_0$. Then $x_0 \in \widehat{X}_0$ by (2), so $\boldsymbol{p_3 \widehat{E}_0 x_0}$.

Case $p_1 E_0 p'_1$ Then $p_1(E_1 \circ A_1)p'_1$, so either $\boldsymbol{p_1 A_1 p'_1}$ or $p_1 A_1 x_1 E_1 x'_1 A_1 p'_1$; in the latter case $x_1, x'_1 \in \widehat{X}_1$ by (1) and (5), hence $\boldsymbol{p_1 A_1 x_1 E_1 x'_1 p'_1}$.

Case $p_1 E_0 x_0$, $x_0 \notin \widehat{X}_0$. Then $x_0 A_0 p_2$ for some p_2 closed in A_1; so $p_2(E_1 \circ A_1)p'_2$, hence $p_2 A_1 p'_2$, hence $\boldsymbol{p_1 A_1 p_2 A_0 x_0}$.

Case $p_1 E_0 x_0$, $x_0 \in \widehat{X}_0$. Since x_0 is not idle in A_0 there are two cases: If $x_0 A_0 p_0$ then $p_1(E_1 \circ A_1)p_0$ so $p_1 A_1 x_1 E_1 p_0$, with $x_1 \in \widehat{X}_1$ by (1), so $\boldsymbol{p_1 A_1 x_1 \widehat{E}_2 x_0}$. If $x_0 A_0 p_2$ then $p_1(E_1 \circ A_1)p_2$, so either $\boldsymbol{p_1 A_1 p_2}$ or $p_1 A_1 x_1 E_1 x'_1 A_1 p_2$; in the latter case $x_1, x'_1 \in \widehat{X}_1$ by (1) and (5), and $x_0 \sim x'_1$ by Lemma 99, so $\boldsymbol{p_1 A_1 \widehat{E}_1 x'_1 \sim x_0}$.

Case $x_0 E_0 x'_0$, $x_0 \notin \widehat{X}_0$. Then $x_0 A_0 p_2$ and $x'_0 A_0 p'_2$ with x_0, x'_0 closed in A_1, and also $p_2(E_1 \circ A_1)p'_2$; hence $p_2 A_1 p'_2$, so $\boldsymbol{x_0 A_0 p_2 A_1 p'_2 A_0 x'_0}$.

Case $x_0 E_0 x'_0$, $x_0 \in \widehat{X}_0$. Then $\boldsymbol{x_0 \widehat{E}_0 x'_0}$ by (5).

This completes the proof that $E \circ D_i = E_i$ $(i = 0, 1)$. It follows that E has no aliases, since an alias for E would be preserved by the composition as an alias for E_i, so it is any arrow in $\mathbf{Edg}_{\bar{a}}$. Furthermore, if any E' satisfies these equations then $E' = E$ since the D_i are epis. Thus $\vec{D}$ is a pushout for $\vec{A}$. ■

B.3 Further IPOs in $\mathbf{Edg}_{\bar{a}}$

We shall now characterise the IPOs for any pair $\vec{A}$ with common domain. In this subsection we use some of the notations introduced at the start of Section 6.

We begin with a simple lemma.

Lemma 103 *If $\vec{B}$ is an IPO for $\vec{A}$, then B_0 and B_1 have no inner ports in common.*

Proof Let $\vec{B}$ have co-domain Y. It is easy to construct a candidate $(\vec{C}, C)$ for $\vec{A}$ w.r.t. $\vec{B}$, in which $C : Z \to Y$ contains all inner ports common to B_0 and B_1. If any such ports exist, then no arrow $F : Y \to Z$ can exist with $C \circ F = \mathsf{id}_Y$, contradicting the assumption that $\vec{B}$ is an IPO. ■.

Our next proposition expresses any IPO for a pair having an idle name in terms of an IPO for the same pair with the name removed.

Proposition 104 (IPO extension for an idle name) *Let x not be a name of A_0. Then:*

(1) *Up to isomorphism any IPO $\vec{B}'$ for $x \otimes A_0, A_1$ takes the form either $\mathsf{id}_x \otimes B_0, x \otimes B_1$ or $(x \mapsto q)B_0, B_1$ where x is not a name or co-name of $\vec{B}$ and q is a port closed in B_0.*

(2) *In each case $\vec{B}$ is an IPO for $\vec{A}$.*

Note that q may be either a co-name or an inner port of B_0. The operation $(x \mapsto q)B_0$ is called an *elision* in Definition 78. This refers to the fact that the idle name x in A_0 is not exported in the IPO interface Y, but instead mapped into the body of B_i.

Proof (1) Suppose not; then, since there are no aliases, the IPO must take one of the forms $\underline{x} \otimes B_0, B_1$ or $((x \mapsto q)B_0, B_1$, with x not a name or co-name of B_0 and q a port open in B_0.

Now let B_i have co-domain Y $(i = 0, 1)$. In either case we construct a candidate $(\vec{C}, C)$ with $C : Z \to Y$, for which there is no mediating arrow $F : Y \to Z$ such that $F \circ B'_0 = C_0$, $F \circ B_1 = C_1$ and $C \circ F = \mathsf{id}_Y$, contradicting the assumption that B'_0, B_1 is an IPO.

In each case take $\vec{C} = \mathsf{id}_x \otimes B_0, x \otimes B_1$. In the first case take $C = \underline{x} \otimes \mathsf{id}_Y$; in the second case we have qB_0y for some $y \in Y$, so take $C = (x \mapsto \underline{y})\mathsf{id}_Y$. It is easy to check that each is a candidate. But in both cases $F \circ B'_0 \neq C_0$, whatever F. For in the first case we have $\underline{x}$ closed in $F \circ B'_0$ but open in C_0; in the second case, $\underline{x}(F \circ B'_0)q$ but not $\underline{x}C_0q$.

(2) Let $\vec{B}$ have co-domain Y. It is easy to prove it a bound for $\vec{A}$. Let $(\vec{C}, C)$ be candidate for $\vec{A}$ w.r.t. $\vec{B}$, whose intermediate domain Z we can assume w.l.o.g. does not contain x. We must find a unique mediating arrow $F : Y \to Z$, with $F \circ B_i = C_i$ and $C \circ F = \mathsf{id}_Y$.

Case $\vec{B}' = \mathsf{id}_x \otimes B_0, x \otimes B_1$.
It is easy to show that the following defines a candidate $(\vec{C}', C')$ for $\vec{A}'$ w.r.t. $\vec{B}'$, with $C' : x \uplus Z \to x \uplus Y$:

$$C'_0 \triangleq \mathsf{id}_x \otimes C_0,\ C'_1 \triangleq x \otimes C_1,\ C' \triangleq \mathsf{id}_x \otimes C\ .$$

Hence there exists unique $F' : x \uplus Y \to x \uplus Z$ satisfying $F' \circ B'_i = C'_i$ $(i = 0, 1)$ and $C' \circ F' = \mathsf{id}_{x \uplus Y}$. From $F' \circ B'_0 = C'_0$ we have $F' \circ (\mathsf{id}_x \otimes B_0) = (\mathsf{id}_x \otimes C_0)$, whence (since x is not a name or co-name for B_0, C_0) we deduce $F' = \mathsf{id}_x \otimes F$ where $F \circ B_0 = C_0$, yielding

$$\begin{array}{lll} & F \circ B_0 = C_0\ ; & \\ & F \circ B_1 = C_1\ , & \text{since } x \otimes (F \circ B_1) = (\mathsf{id}_x \otimes F) \circ (x \otimes B_1) = F' \circ B'_1 = x \otimes C_1\ ; \\ \text{and} & C \circ F = \mathsf{id}_Y\ , & \text{since } \mathsf{id}_x \otimes (C \circ F) = (\mathsf{id}_x \otimes C) \circ (\mathsf{id}_x \otimes F) = C' \circ F' = \mathsf{id}_{x \uplus Y} = \mathsf{id}_x \otimes \mathsf{id}_Y\ . \end{array}$$

Thus F is the required mediator; its unicity easily follows from that of F'.

Case $\vec{B}' = (x \mapsto q)B_0, B_1$, with q closed in B_0.
It is easy to show that the following defines a candidate $(\vec{C}', C')$ for $\vec{A}'$ w.r.t. $\vec{B}'$, with $C' : Z \to Y$:

$$C'_0 \triangleq (x \mapsto q)C_0,\ C'_1 \triangleq C_1,\ C' \triangleq C\ ;$$

in particular, Lemma 103 ensures that q is a port of C_0 (for if not, it would be an inner port of both B'_0 and B'_1). Hence there exists unique $F' : x \uplus Y \to x \uplus Z$ satisfying $F' \circ B'_i = C'_i$ $(i = 0, 1)$ and $C' \circ F' = \mathsf{id}_Y$. Taking $i = 0$ we get $F' \circ ((x \mapsto q)B_0) = (x \mapsto q)C_0$, whence $(x \mapsto q)(F' \circ B_0) = (x \mapsto q)C_0$; so we have

$$\begin{array}{ll} & F' \circ B_0 = C_0 \text{ immediately;} \\ & F' \circ B_1 = C_1\ ; \\ \text{and} & C \circ F' = \mathsf{id}_Y\ . \end{array}$$

Thus $F \triangleq F'$ itself is the required mediator, whose unicity is deduced easily from its unicity as mediator between $\vec{B}'$ and $\vec{C}'$.

This concludes the proof that $\vec{B}'$ is an IPO for $\vec{A}'$. ∎

This proposition shows that any IPO for $\vec{A}'$ can be derived by a sequence of steps (one for each idle name of $\vec{A}'$ from the IPO of a pair $\vec{A}$ of epis, provided the pair is consistent. It does not prove that any IPO for $\vec{A}'$ exists. We now complete the work of the section by showing that indeed this construction provides exactly all the IPOs for $\vec{A}'$. Let us next express the construction succinctly.

Construction 105 (IPOs in edge nets) Let $\vec{A}'$ be a pair of edge nets in $\mathbf{Edg}_{\bar{\mathsf{a}}}$. We construct a family of IPOs for $\vec{A}'$ as follows.

Choose any subset W_i of the idle names of A_i, and let its remaining idle names be X_i ($i = 0, 1$). Then $\vec{A}' = X_0 \otimes W_0 \otimes A_0, X \otimes W_1 \otimes A_1$ where $\vec{A} = A_0, A_1$ is a pair of epis. If the pair is consistent (otherwise $\vec{A}'$ is also inconsistent) then let $\vec{B}$ be an IPO for $\vec{A}$, unique up to isomorphism. We may assume that $\vec{B}$ has co-domain Y disjoint from X.

For $i = 0, 1$ choose an elision (Definition 78) η_i mapping W_i to the closed ports of B_i (which may be either inner ports or co-names). Now define

$$\vec{B}' \triangleq \mathsf{id}_{X_0} \otimes X_1 \otimes \eta_0 B_0,\ X_0 \otimes \mathsf{id}_{X_1} \otimes \eta_1 B_1\ ;$$

for each choice of W_i and η_i ($i = 0, 1$) this is an IPO for $\vec{A}'$. ■

We are now ready to validate this construction, and indeed show that it yields all and only the IPOs of $\vec{A}'$. First, two lemmas to deal with the idle names:

Lemma 106 *Let $\vec{B}$ be an IPO for $\vec{A}$, and x be not a name or co-name of either pair. Then $\vec{B}' = \mathsf{id}_x \otimes B_0, x \otimes B_1$ is an IPO for $x \otimes A_0, A_1$.*

Proof It is easy to show that $\vec{B}'$ is a bound for $x \otimes A_0, A_1$. Let $\vec{B}$ have co-domain Y, and let $(\vec{C}', C')$ be a candidate for $x \otimes A_0, A_1$ w.r.t. $\vec{B}'$, where $C' : Z' \to x \uplus Y$. We have to find a unique mediating arrow $F' : x \uplus Y \to Z'$ satisfying

$$F' \circ B_i' = C_i'\ (i = 0, 1) \text{ and } C' \circ F' = \mathsf{id}_{x \uplus Y}\ .$$

Now since $C' \circ C_0' = B_0' = \mathsf{id}_x \otimes B_0$, we find that Z' may be partitioned as $Z' = X \uplus Z$ and that there exist $C : Z \to Y$ and C_0 with co-domain Z such that

$$C_0' = I_x \otimes C_0,\ C' = J_x \otimes C \text{ and } C \circ C_0 = B_0$$

where $I_x : x \to X$ links x to a single member of X, while $J_x : X \to x$ links all members of X to x. Furthermore, since $C' \circ C_1' = B_1'$, i.e. $(J_x \otimes C) \circ C_1' = x \otimes B_1$, we deduce that for some C_1 with co-domain Z

$$C_1' = X \otimes C_1 \text{ and } C \circ C_1 = B_1\ ,$$

and we also find $X \otimes (C_0 \circ A_0) = C_0' \circ (x \otimes A_0) = C_1' \circ A_1 = X \otimes (C_1 \circ A_1)$, whence $C_0 \circ A_0 = C_1 \circ A_1$; so $(\vec{C}, C)$ is a candidate for $\vec{A}$ w.r.t. $\vec{B}$.

It follows that there exists a unique mediator $F : Y \to Z$ satisfying $F \circ B_1 = C_i$ ($i = 0, 1$) and $C \circ F = \mathsf{id}_Y$. Now take $F' \triangleq I_x \otimes F$. We readily deduce that $F' : x \uplus Y \to Z'$ is the mediator we seek between $\vec{B}'$ and the candidate $(\vec{C}', C')$. To see that it is unique, suppose that G' is another; then from $G' \circ B_0' = C_0'$ we find that $G' = I_x \otimes G$ where $G \circ B_0 = C_0$; from this we deduce that also $G \circ B_1 = C_1$ and $C \circ G = \mathsf{id}_Y$, so $G = F$ by the unicity of F, hence also $G' = F'$. ■

Lemma 107 *Let $\vec{B}$ be an IPO for $\vec{A}$, both epis. For $i = 0, 1$ let the set W_i ($i = 0, 1$) contain no name or co-name of either pair, and let η_i map W_i into the closed ports of B_i. Then $\vec{B}' = \eta_0 B_0, \eta_1 B_1$ is an IPO for $W_0 \otimes A_0, W_1 \otimes A_1$.*

By Theorem 102, $\vec{B}$ is the pushout (unique up to isomorphism) for $\vec{A}$ produced by Construction 98. Let $(\vec{C}', C')$ be a candidate for $W_0 \otimes A_0, W_1 \otimes A_1$ w.r.t $\vec{B}$. with Z the co-domain of C'; then we require a unique mediating arrow $F'' : Y \to Z$ satisfying

$$F' \circ B'_i = C'_i \ (i = 0, 1) \text{ and } C' \circ F' = \mathsf{id}_Y \ .$$

We shall first show that $\eta_0(w)$ is closed in C'_0 for each $w \in W_0$. (Likewise $\eta_1(w)$ is closed in C'_1 for each $w \in W_1$.) There are two cases. First, if $\eta_0(w) = p_1$ is an inner port of B_0, then the construction ensures that it is closed in A_1; hence by $C'_0 \circ A'_0 = C'_1 \circ A'_1$ it is also closed in C'_0. On the other hand if $\eta_0(w) = x_0$ is a co-name of B_i then $x \notin \widehat{X}_i$ (using the notation of Construction 98), so for some p_2 we have $x_0 A_0 p_2$ with p_2 closed in A_1; again $C'_0 \circ A'_0 = C'_1 \circ A'_1$ ensures that x_0 is closed in C'_0.

We deduce from this, and the known equation $C' \circ C'_i = B'_i$ $(i = 0, 1)$ that C'_i takes the form $C'_i = \eta_i C_i$ where $C' \circ C_i = B_i$. Also $C_0 \circ A_0 = C'_0 \circ A'_0 = C'_1 \circ A'_1$; so by taking $C \triangleq C'$ we have a candidate $(\vec{C}, C)$ for $\vec{A}$ w.r.t. $\vec{B}$. Since the latter is an IPO, we have that there is a unique mediator $F : Y \to Z$ satisfying

$$F \circ B_i = C_i \ (i = 0, 1) \text{ and } C \circ F = \mathsf{id}_Y \ .$$

From this, taking $F' \triangleq F$, we deduce quite simply the required mediating equations for F'; we use the fact that each $\eta_i(w)$ is closed in C'_i. Moreover, the unicity of F' can be deduced easily from the unicity of F, and the Lemma is proved. ■

We may now complete the proof of the theorem which characterises IPOs for edge nets:

Theorem 108 (characterizing IPOs for topographs) *The IPOs for any pair of edge nets in* $\mathbf{Edg}_{\bar{\mathsf{a}}}$ *are, up to isomorphism, exactly those produced by Construction 105.*

Proof Let $\vec{A}'$ be the pair. Then Proposition 104 ensures that an IPO for $\vec{A}'$ is generated by Construction 105; for by applying the Proposition inductively, removing the idle names of $\vec{A}$ one by one, we determine the pairs W_i, η_i $(i = 0, 1)$ for which the construction should be applied.

It remains to show that for each such application an IPO is created. Therefore let us assume that

$$\vec{A}' = X_0 \otimes W_0 \otimes A_0, X_1 \otimes W_1 \otimes A_1$$

with $\vec{A}$ both epis, and let $\vec{B}$ be an IPO for $\vec{A}$ with co-domain Y disjoint from the X_i. Choose elisions η_i (see Definition 78) mapping W_i into the closed ports of B_i $(i = 0, 1)$.

First, Lemma 107 shows that $\vec{B}'' = \eta_0 B_0, \eta_1 B_1$ is an IPO for $\vec{A}'' = W_0 \otimes A_0, W_1 \otimes A_1$.

It only remains to define $\vec{B}' \triangleq \mathsf{id}_{X_0} \otimes X_1 \otimes B''_0,\ X_0 \otimes \mathsf{id}_{X_1} \otimes B''_1$; for by inductive use of Lemma 106 we deduce that it is an IPO for $\vec{A}'$ w.r.t. $\vec{B}'$, and the theorem is proved. ■

C RPOs in edge nets

In Appendix B we have shown that IPOs exist for any consistent pair $\vec{A}$ in $\mathbf{Edg}_{\bar{\mathsf{a}}}$; furthermore we were able to give a construction to generate them all, which is necessary to characterise the labelled transition systems (LTSs) derived in this paper.

Here we shall focus our attention on RPOs in $\mathbf{Edg}_{\bar{\mathsf{a}}}$, which are needed to ensure that familiar behavioural equivalences for our derived LTSs are congruences. For this purpose we require an RPO to exist for $\vec{A}$ w.r.t. any bound $\vec{E}$ whenever A_1 is a redex. Elsewhere it has been argued that any reasonable redex should be an epi; and we are able to show that an RPO does indeed always exist if A_1 is epi, as required for Subsection 4.2.

RPOs do not *always* exist! Counter-examples arise with non-epis, i.e. with idle names. In the lower diagram of Figure 15 we show a simple case in which the pair $\vec{A}$, each with an idle name, has no RPO w.r.t. $\vec{E}$. The figure shows two candidates $(\vec{B}, B)$ and $(\vec{C}, C)$; their legs $\vec{B}$ and $\vec{C}$ are identical, and their heads B and C differ in the simplest possible way. It is easy to show that there is no mediating arrow between the two candidates, and then to show that no other candidate can dominate both.

We now proceed to the RPO construction. Section C.1 presents the construction; Section C.2 validates it.

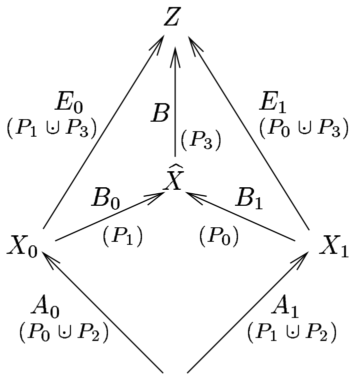
C.1 Construction

From now on we shall work in $\mathbf{Edg}_{\bar{\mathsf{a}}}$ with an arbitrary pair of edge nets $\vec{A} : W \to \vec{X}$ having a bound $\vec{E} : \vec{X} \to Z$. Under a simple assumption, we shall give and justify the construction of an RPO for $\vec{A}$ w.r.t. $\vec{E}$. In this section we summarise the construction and give examples; in Section C.2 the next section we repeat the construction piecemeal interleaved with a series of lemmas and propositions.

The internal ports of $E_i \circ A_i$ may be partitioned into four sets: P_i common to A_i and E_{1-i} $(i = 0, 1)$, P_2 common to A_0 and A_1, and P_3 common to E_0 and E_1, as indicated (in parentheses) in the diagram at the top of Figure 14. We shall consistently use $p_i, p_i' p_i'', \ldots$ $(i = 0, 1, 2, 3)$ to range over these sets, and x_i over X_i $(i = 0, 1)$. Furthermore, for ease of notation, from now on we shall now consider W to be empty; this loses no generality, since the construction of an RPO treats the set W just like P_2.

An RPO consists of a triple of edge nets $B_0 : X_0 \to \widehat{X}$, $B_1 : X_1 \to \widehat{X}$ and $B : \widehat{X} \to Z$, with the desired universal property. The first step in our construction is to take their internal ports sets to be P_1, P_0 and P_3 respectively. The next step is the RPO interface $\widehat{X}$. Each $x \in \widehat{X}$ will be an equivalence class of certain $x_i \in \widehat{X}_i$ $(i = 0, 1)$ which need to be exported. To help understand the construction, here is an example of an x_0 which should *not* be exported via $\widehat{X}$, i.e. for which x_0 must be closed in B_0: Suppose that $p_2 A_0 x_0$, but that p_2 is closed in A_1; then p_2 will be closed in $B_1 \circ A_1$, so to achieve $B_0 \circ A_0 = B_1 \circ A_1$ we must also have p_2 closed in $B_0 \circ A_0$, which precludes exporting x_0 via $\widehat{X}$.

The entire construction of $(\vec{B}, B)$ is illustrated and presented in Figure 14. An example of the construction, together with another candidate $(\vec{C}, C)$, is shown in the top diagram of Figure 15. Though very simple, involving a single port p_1 with $p_1 A_1 x_1$ and $x_0 E_0 p_1$, it exhibits some subtleties. Indeed our construction does not yield an RPO in some degenerate cases. The second diagram in the figure illustrates what happens if the single port p_1 is omitted from the first example; the RPO is lost! (The notes in the diagrams indicate why.) However, we shall find that RPOs always exist in $\mathbf{Edg}_{\bar{\mathsf{a}}}$ provided only that one of $\vec{A}$ is an epi, which in this precategory means that it has no idle names.



The RPO interface $\widehat{\boldsymbol{X}}$: The ports in X_i which will be exported via the interface are denoted by $\widehat{X}_i$, and those not exported are denoted by $\ddot{X}_i$, defined as follows ($i = 0, 1$):

$$
\begin{array}{rcl}
\widehat{X}_i & \triangleq & \{\, x_i \mid \;\; (\exists x_i'.x_i E_i x_i' A_i p_2) \Rightarrow p_2 \text{ open in } A_{\bar{\imath}}, \text{ and} \\
& & \qquad\;\; (\exists x_i'.x_i E_i p_{\bar{\imath}}) \Rightarrow p_{\bar{\imath}} \text{ open in } A_{\bar{\imath}} \,\} \\
\ddot{X}_i & \triangleq & X_i \backslash \widehat{X}_i \;.
\end{array}
$$

We next define an equivalence relation over $\widehat{X}_0$ and $\widehat{X}_1$, and take $\widehat{X}$ to be its quotient:

$$
\begin{array}{rcl}
\sim & \triangleq & \{(x_0, x_1) \mid x_0 A_0 p_2 A_1 x_1 \text{ for some } p_2\}^{\equiv} \\
\widehat{X} & \triangleq & (\widehat{X}_0 + \widehat{X}_1)/\!\sim \;.
\end{array}
$$

We extend $\sim$ to $\widehat{X}_0 + \widehat{X}_1 + \widehat{X}$ by declaring that $x_i \sim [x_i]_\sim$.

The RPO legs $\boldsymbol{B_0}, \boldsymbol{B_1}$: To define B_0, for example, we employ part of A_1 involving P_1 and $\widehat{X}_1$, and part of E_0 involving P_1 and $\ddot{X}_0$. To export both $\widehat{X}_0$ and $\widehat{X}_1$ in B_0, we incorporate $\sim$ and restrict away X_1:

$$
\begin{array}{rcl}
\widehat{A}_i & \triangleq & A_i \restriction (P_i \uplus \widehat{X}_i) \\
\ddot{E}_i & \triangleq & \{(x_i, q) \in E_i \mid x_i \in \ddot{X}_i,\, q \in \ddot{X}_i \uplus P_{\bar{\imath}}\}^{\equiv} \\
B_i & \triangleq & (\ddot{E}_i \sqcup \widehat{A}_{\bar{\imath}} \sqcup \sim) \backslash X_{\bar{\imath}} \;.
\end{array}
$$

The RPO head $\boldsymbol{B}$: We define a larger equivalence upon $\widehat{X}_0 + \widehat{X}_1$, and use it in defining B:

$$
\begin{array}{rcl}
\sim_E & \triangleq & \sim \sqcup \; \{(x_0, x_1) \mid \exists p_1.x_0 E_0 p_1 A_1 x_1 \text{ or } \exists p_0.x_0 A_0 p_0 E_1 x_1\}^{\equiv} \\
B & \triangleq & (\widehat{E}_0 \sqcup \widehat{E}_1 \sqcup \sim_E) \backslash X_0 \backslash X_1 \qquad \text{where } \widehat{E}_i \;\triangleq\; E_i \restriction (\widehat{X}_i \uplus P_3 \uplus Z) \; (i = 0, 1) \;.
\end{array}
$$

Figure 14: Construction of an RPO $(\vec{\boldsymbol{B}}, \boldsymbol{B})$ for $\vec{\boldsymbol{A}}$ w.r.t. $\vec{\boldsymbol{E}}$.

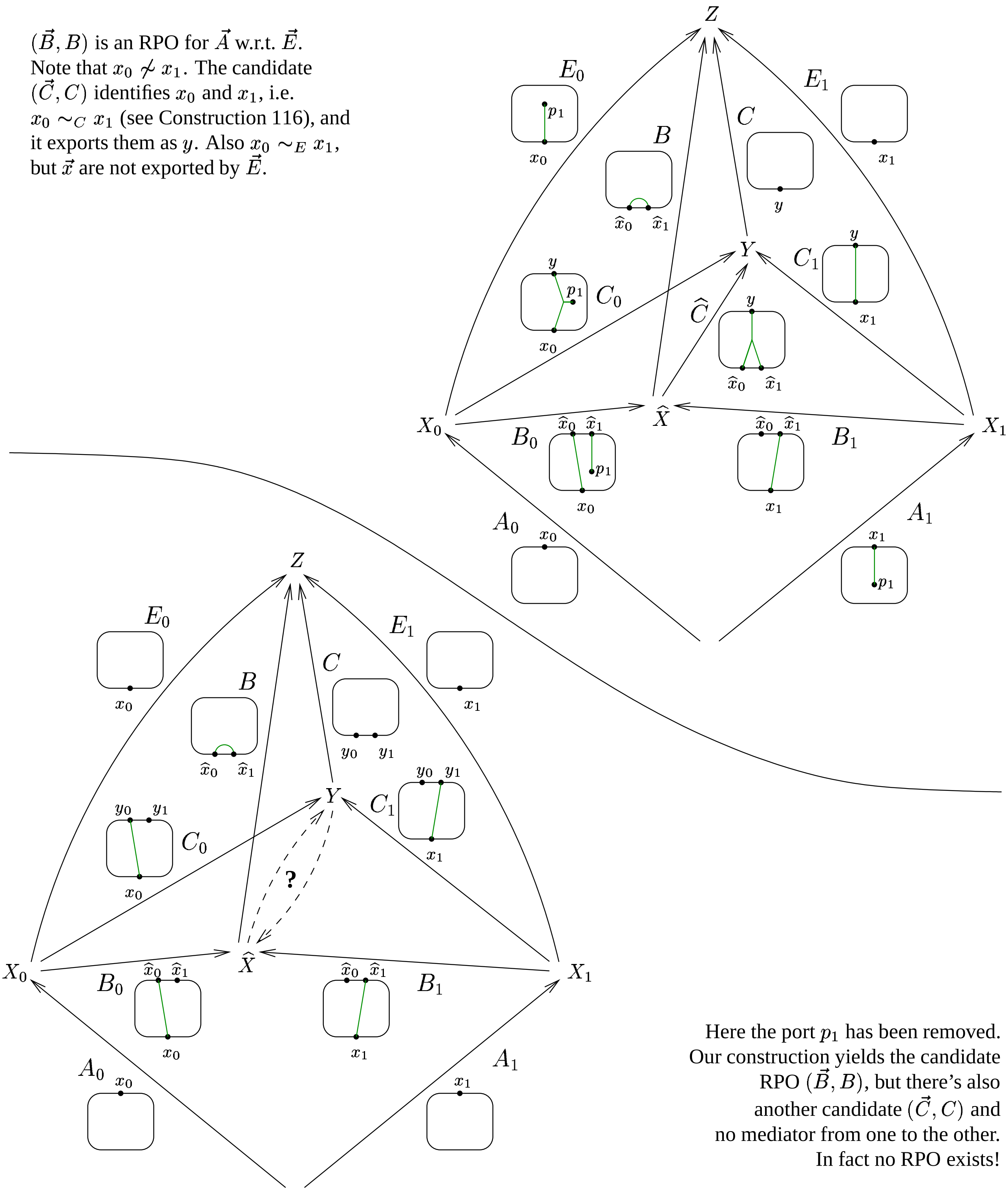


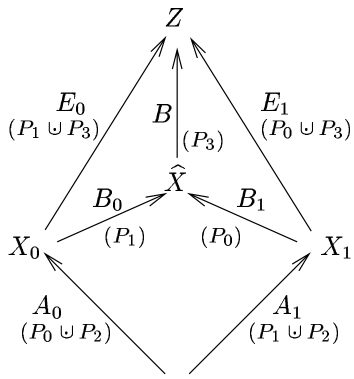
Figure 15: An example and a counter-example for RPOs

C.2 Validation

In this section we repeat the complete RPO construction at a slower pace, interleaved with lemmas and propositions and ending with a summary theorem. Some of them hold when subscripts 0 and 1 are systematically interchanged; in this case they are annotated (example: Lemma 110 $^{(0\leftrightarrow 1)}$).

Recall the notations and terminology of the previous section. In particular, A_i has internal ports $P_i \uplus P_2$ $(i = 0, 1)$. In our RPO construction we are taking their co-name set W to be empty; this loses no generality, since the full construction treats W just as it treats P_2.

An RPO of $\vec{A}$ w.r.t. $\vec{E}$ consists of a triple of edge nets $B_0 : X_0 \to \widehat{X}$, $B_1 : X_1 \to \widehat{X}$ and $B : \widehat{X} \to Z$, with the desired universal property. The first step in our construction is to take their internal ports sets to be P_1, P_0 and P_3 respectively.



The next step in the construction is the RPO interface $\widehat{X}$. Each $x \in \widehat{X}$ will be an equivalence class of certain $x_i \in \widehat{X}_i$ $(i = 0, 1)$ which need to be exported.

Construction 109 (RPO interface) The ports in X_i which are be exported via the RPO interface $\widehat{X}$ are denoted by $\widehat{X}_i$, and those not exported are denoted by $\ddot{X}_i$ $(i = 0, 1)$; they are defined as follows:

$$\begin{array}{rcl} \widehat{X}_i & \triangleq & \{ x_i \mid \ (\exists x_i'.x_i E_i x_i' A_i p_2) \Rightarrow p_2 \text{ open in } A_{\bar{\imath}}, \text{ and} \\ & & \qquad (\exists x_i'.x_i E_i p_{\bar{\imath}}) \Rightarrow p_{\bar{\imath}} \text{ open in } A_{\bar{\imath}} \} \\ \ddot{X}_i & \triangleq & X_i \backslash \widehat{X}_i \ . \end{array}$$

We next define an equivalence relation over $\widehat{X}_0$ and $\widehat{X}_1$, and take $\widehat{X}$ to be its quotient:

$$\begin{array}{rcl} \sim & \triangleq & \{(x_0, x_1) \mid x_0 A_0 p_2 A_1 x_1 \text{ for some } p_2\}^{\equiv} \\ \widehat{X} & \triangleq & (\widehat{X}_0 + \widehat{X}_1)/\sim \ . \end{array}$$

Finally, we extend $\sim$ to $\widehat{X}_0 + \widehat{X}_1 + \widehat{X}$ by declaring that $x_i \sim [x_i]_\sim$. ■

Lemma 110 $^{(0\leftrightarrow 1)}$ *If $x_0 A_0 p_2 A_1 x_1$ then $x_0 \in \widehat{X}_0$, $x_1 \in \widehat{X}_1$ and $x_0 \sim x_1$.*

Proof Suppose $x_0 \in \ddot{X}_0$; then there are two cases, both yielding a contradiction:

Case $x_0 E_0 x_0' A_0 p_2'$ and p_2' closed in A_1. Then $p_2 (E_0 \circ A_0) p_2'$, hence $p_2 (E_1 \circ A_1) p_2'$, so by Lemma 101(2) $p_2 A_1 p_2'$ and hence p_2 is closed in A_1, contra $p_2 A_0 x_0$.

Case $x_0 E_0 p_1$ and p_1 closed in A_1. So similarly $p_2(E_1 \circ A_1)p_1$, and again we find a contradiction.

So $x_0 \in \widehat{X}_0$, and similarly $x_1 \in \widehat{X}_1$. Thus by definition $x_0 \sim x_1$. ∎

Lemma 111 $^{(0\leftrightarrow 1)}$ *Each of the following conditions implies $x_0 \in \widehat{X}_0$:*

$$\begin{array}{llll} (1) & x_0 E_0 x'_0 A_0 p_0 & (2) & x_0 E_0 x'_0 A_0 p_2 \;\; (p_2 \textit{ open in } A_1) \\ (3) & x_0 E_0 p_3 & (4) & x_0 E_0 p_1 \;\; (p_1 \textit{ open in } A_1)\ . \end{array}$$

Proof The arguments from each condition are similar. For example, assuming condition (2), suppose $x_0 \in \ddot{X}_0$ and consider two cases:

Case $x_0 E_0 x'_0 A_0 p_2$, p_2 closed in A_1. Then we deduce $p_2(E_1 \circ A_1)p'_2$, whence by Lemma 101(2) $p_2 A_2 p'_2$, contra p_2 open in A_1.

Case $x_0 E_0 p_1, p_1$ closed in A_1. Again $p_2(E_1 \circ A_1)p_1$, whence a similar contradiction. ∎

We are now ready for the next stage of the RPO construction; we define its two legs $\vec{B}$.

Construction 112 (RPO legs) We define the legs B_0, B_1 of our intended RPO. First, in defining B_0 for example, we need the part of A_1 involving P_1 and the exported names:

$$\widehat{A}_i \;\triangleq\; A_i \restriction (P_i \uplus \widehat{X}_i)\ .$$

In building B_0 we also need a part of E_0 involving P_1 and the unexported names in X_0:

$$\ddot{E}_i \;\triangleq\; \{(x_i, q) \in E_i \mid x_i \in \ddot{X}_i,\ q \in \ddot{X}_i \uplus P_{\bar{\imath}}\}^{\equiv}\ .$$

Finally, in order to export both $\widehat{X}_0$ and $\widehat{X}_1$, we incorporate $\sim$ in B_0 and then restrict away X_1 to yield an equivalence on $X_0 \uplus P_1 \uplus \widehat{X}$:

$$B_i \;\triangleq\; (\ddot{E}_i \sqcup \widehat{A}_{\bar{\imath}} \sqcup \sim) \backslash X_{\bar{\imath}}\ .$$ ∎

A point about $\ddot{E}_0$ deserves comment. Why not simply $\ddot{E}_0 = E_0 \restriction (\ddot{X}_0 \uplus P_1)$? The latter would include any pair (p_1, p'_1) which is a member of E_0. But such a pair should not always be in B_0, because the corresponding leg C_0 of a candidate may not contain the pair – and an RPO leg should only link ports which will be linked by *every* candidate. The present definition ensures that if $p_1 E_0 x_0 E_0 p'_1$ for some x_0 then the pair will indeed be in B_0 – whether or not $x_0 \in \widehat{X}_0$.

We must check that B_0, B_1 are indeed in $\mathbf{Edg}_{\bar{\mathsf{a}}}$:

Proposition 113 (RPO legs are well-formed) *B_0 and B_1 have no aliases.*

Proof We use the notion of support. It is obvious that $\ddot{E}_0$ and $\sim$ have disjoint support, and by Lemma 111(4) it is easy to show that the supports of $\ddot{E}_0$ and $\widehat{A}_1$ are disjoint.

Now suppose $x B_0 x'$, where $x, x' \in \widehat{X}$; we shall prove $x = x'$. There must be a lub chain for $\{\ddot{E}_0, \widehat{A}_1, \sim\}$ whose end-members are x, x'. So the end-links of the chain are $\sim$ links, and from the above it follows that it has no $\ddot{E}_0$ links; so it consists of alternating $\sim$ and $\widehat{A}_1$ links. But any $\widehat{A}_1$ link can only take the form $x_1 A_1 x'_1$, implying $x_1 = x'_1$ by the no-alias condition; so there can be no $\widehat{A}_1$ links. Thus the chain consists of a single link $x \sim x'$. But x, x' are equivalence classes over $\widehat{X}_0 + \widehat{X}_1$, so $x = x'$, as required. ∎

We are now ready to prove that $\vec{B}$ make a commuting diagram with $\vec{A}$.

Proposition 114 (RPO legs yield a bound) *Let* $\widehat{A} = (A_0 \sqcup A_1 \sqcup \sim) \backslash X_0 \backslash X_1$. *Then* $B_0 \circ A_0 = \widehat{A} = B_1 \circ A_1$.

Proof It will be enough to prove $B_0 \circ A_0 = \widehat{A}$, which we do as two inequalities. Note first that $B_0 \circ A_0 = (A_0 \sqcup \ddot{E}_0 \sqcup \widehat{A}_1 \sqcup \sim) \backslash X_0 \backslash X_1$.

To prove $B_0 \circ A_0 \subseteq \widehat{A}$: For every instance of $B_0 \circ A_0$ there is a lub chain for $\{A_0, \ddot{E}_0, \widehat{A}_1, \sim\}$ whose extremal members are not in X_i $(i = 0, 1)$. So we need only show that each $\ddot{E}_0$ link in the chain, together perhaps with adjacent links, implies an instance of A_0 or A_1. We proceed by cases of $q\ddot{E}_0 q'$, where $q, q' \in \ddot{X}_0 \uplus P_1$. In each case, we emphasize the replacement instance in bold type.

Case $p_1 \ddot{E}_0 p'_1$. Then for some $x_0 \in \ddot{X}_0$ we have $p_1 E_0 x_0 E_0 p'_1$. Hence by Lemma 111(4) p_1, p'_1 are closed in A_1. But also $p_1 (E_1 \circ A_1) p'_1$, hence by Lemma 101(2) $\boldsymbol{p_1 A_1 p'_1}$.

Case $p_1 \ddot{E}_0 x_0$. Then $x_0 \in \ddot{X}_0$ and $p_1 E_0 x_0$. This link is not rightmost in the lub chain; so there is an adjacent instance, necessarily of A_0, of the form $x_0 A_0 q$ for $q \in P_0 \uplus P_2$. ($q \in X_0$ would imply $q = x_0$ since A_0 has no aliases, contra the definition of a lub chain.) But $q \in P_0$ contradicts Lemma 111(1), hence $q = p_2 \in P_2$ and we must replace $p_1 E_0 x_0 A_0 p_2$. Hence $p_1 (E_1 \circ A_1) p_2$; also by definition of $\ddot{X}_0$ we have p_2 closed in A_1; so by Lemma 101(2) we have $\boldsymbol{p_1 A_1 p_2}$.

Case $x_0 \ddot{E}_0 x'_0$. Then $x_0 E_0 x'_0$ with $x_0, x'_0 \in \ddot{X}_0$, and as above there must be adjacent instances $q A_0 x_0$ and $x'_0 A_0 q'$ with $q \in P_0 \uplus P_2$. Again, $q \in P_0$ or $q' \in P_0$ contradicts Lemma 111(1), so we have to replace $p_2 A_0 x_0 E_0 x'_0 A_0 p'_2$. But we deduce $p_2 (E_0 \circ A_0) p'_2$, whence as above $\boldsymbol{p_2 A_1 p'_2}$.

This concludes the elimination of $\ddot{E}_0$ from the chain, so we have proved $B_0 \circ A_0 \subseteq \widehat{A}$.

To prove $\widehat{A} \subseteq B_0 \circ A_0$: In a lub chain for $\{A_0, A_1, \sim\}$ with no member of X_0 or X_1 extremal, for each instance of A_1 we need find a replacement in terms of $\{A_0, \ddot{E}_0, \widehat{A}_1, \sim\}$.

Case $p_2 A_1 p'_2$. Then $p_2 (E_0 \circ A_0) p'_2$, so by Lemma 101(2) either $\boldsymbol{p_2 A_0 p'_2}$ or $p_2 A_0 x_0 E_0 x'_0 A_0 p'_2$. In the latter case, if p_2 (hence p'_2) is open in A_1 then $p_2 A_1 x_1$ and $p'_2 A_1 x_1$ for some x_1, whence by Lemma 110 we have $x_0, x'_0 \in \widehat{X}_0$ and $x_1 \in \widehat{X}_1$, with $x_0 \sim x_1 \sim x'_0$; hence $\boldsymbol{p_2 A_0 x_0 \sim x'_0 A_0 p'_2}$. On the other hand if p_2 and p'_2 are closed in A_1 then $x_0, x'_0 \in \ddot{X}_0$, hence $\boldsymbol{p_2 A_0 x_0 \ddot{E}_0 x'_0 A_0 p'_2}$.

Case $p_2 A_1 p_1$. Then $p_2 (E_0 \circ A_0) p_1$, so by Lemma 101(1) $p_2 A_0 x_0 E_0 p_1$. By similar reasoning to above, if p_2 is open in A_1 we find a replacement $\boldsymbol{p_2 A_0 x_0 \sim x_1 \widehat{A}_1 p_1}$; otherwise $\boldsymbol{p_2 A_0 x_0 \ddot{E}_0 p'_1}$.

Case $p_2 A_1 x_1$. Then $x_1 \in \widehat{X}_1$, since x_1 cannot be extremal in the chain and if $x_1 \in \ddot{X}_0$ then no right-adjacent link can exist. So p_2 is open in A_0, i.e. $p_2 A_0 x_0$ for some x_0; so by Lemma 110 we have $x_0 \in \widehat{X}_0$ and $x_0 \sim x_1$, yielding the replacement $\boldsymbol{p_2 A_0 x_0 \sim x_1}$.

Case $q A_1 q'$ $(q, q' \in P_1 \uplus X_1)$. Then q, q' cannot be in $\ddot{X}_0$, for the above reasons; so $\boldsymbol{q \widehat{A}_1 q'}$.

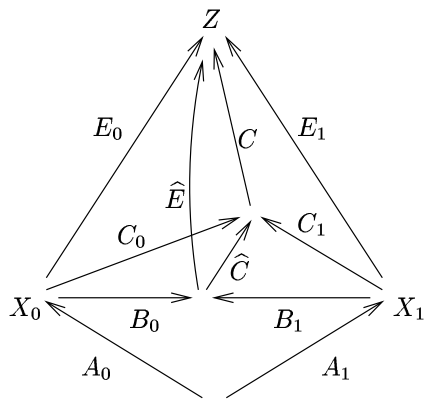
This concludes the proof of the second inequality, and hence of the proposition. ∎

For later use, we record an important property of $\vec{B}$.

Proposition 115 $^{(0 \leftrightarrow 1)}$ **(RPO legs keep an epi)** *If* A_1 *is epi then* B_0 *is epi.*

Proof Consider $x \in \widehat{X}$; there exists either $x_0 \in x$ with $x_0 B_0 x$, or $x_1 \in x$. In the latter case, as A_1 has no idle names, either some $p_1 A_1 x_1$, in which case by construction $p_1 B_0 x$, or some $p_2 A_1 x_1$. In the latter case, since $x_1 \in \widehat{X}_1$, p_2 is open in A_0 so some $x_0 A_0 p_2$, whence by Lemma 110 $x_0 \sim x_1$, so $x_0 \in x$, and again $x_0 B_0 x$. Thus x is not idle in B_0. ∎

We now proceed to justify our construction by finding $B : \widehat{X} \to Z$ which makes the triple $\vec{B}, B$ an RPO for $\vec{A}$ w.r.t. $\vec{E}$. Rather than define B directly, we prefer first to consider all RPO candidates, as defined in Definition 6. Observe that $(E_0, E_1, \mathsf{id}_Z)$ is trivially a candidate. Our work is lightened by treating candidates generally before taking account of this extremal one. The next step of our construction is, for any candidate $(\vec{C}, C)$, to define the equivalence upon $\widehat{X}_0 + \widehat{X}_1$ which it induces; then we can define the mediating arrow from the RPO legs to the candidate.



Construction 116 (RPO mediator) Let $(\vec{C}, C)$ be a candidate for $\vec{A}$ w.r.t. $\vec{E}$, where $C_i : X_i \to Y$ $(i = 0, 1)$ and $C : Y \to Z$. Define an equivalence upon $\widehat{X}_0 + \widehat{X}_1$ as follows:

$$\sim_C \quad \triangleq \quad \sim \sqcup \ \{(x_0, x_1) \mid x_0 C_0 p_1 A_1 x_1 \text{ or } x_0 A_0 p_0 C_1 x_1\}^{\equiv} \ .$$

Let the internal ports sets of C_0 and C_1 be respectively $P_1 \uplus P_4$ and $P_0 \uplus P_4$, where $P_3 = P_4 \uplus P_5$. Define the *mediator* $\widehat{C}$ as follows:

$$\begin{aligned} \widehat{C}_0 &\triangleq C_0 \restriction (\widehat{X}_0 \uplus P_4) \\ \widehat{C}_1 &\triangleq C_1 \restriction (\widehat{X}_1 \uplus P_4) \\ \widehat{C} &\triangleq (\widehat{C}_0 \sqcup \widehat{C}_1 \sqcup \sim_C) \backslash X_0 \backslash X_1 \ . \end{aligned}$$ ■

Our new equivalence deserved explanation. Intuitively, the larger – i.e. the closer to $(\vec{E}, \mathsf{id}_Z)$ – a candidate becomes, the more ports it will link. Thus, if C_0 (say) is to factor through the mediator $\widehat{C}$, the latter must link more ports. The equivalence $\sim_C$ is intended to represent this identification, at least upon the interface ports X_0 and X_1.

Before validating our mediator we need some lemmas. For the rest of this section we shall discuss an arbitrary candidate $(\vec{C}, C)$ with components denoted as in the above construction.

Lemma 117 $^{(0\leftrightarrow 1)}$

(1) *If* $x_0 \in \ddot{X}_0$ *then* x_0 *is closed in* C_0

(2) $\ddot{E}_0 \subseteq \ddot{C}_0$.

Proof (1) There are two cases:

Case $x_0 E_0 x_0' A_0 p_2$ (p_2 closed in A_1). Then from $C_0 \circ A_0 = C_1 \circ A_1$ we know p_2 is closed in $C_0 \circ A_0$, and hence x_0' closed in C_0. But $x_0 E_0 x_0'$ and $E_0 = C \circ C_0$, so $x_0 C_0 x_0'$ by Lemma 101(2), whence also x_0 closed in C_0.

Case $x_0E_0p_1$ (p_1 closed in A_1). Then again p_2 closed in $C_0 \circ A_0$, hence by Lemma 101(3) p_1 is closed in C_0, whence also x_0 closed in C_0.

(2) Recall that $\ddot{E}_0$ is generated by pairs (x_0, p_1) and (x_0, x_0') in E_0, where $x_0, x_0' \in \ddot{X}_0$; the rest follows from $E_0 = C \circ C_0$, by (1) and Lemma 101(2). ∎

Lemma 118 $^{(0\leftrightarrow 1)}$ *If* $x_0C_0p_1A_1x_1$ *or* $x_0A_0p_0C_1x_1$ *then* $x_0 \in \widehat{X}_0$, $x_1 \in \widehat{X}_1$ *and* $x_0 \sim_C x_1$.

Proof We first deduce contradictions from $x_0 \in \ddot{X}_0$, $x_1 \in \ddot{X}_1$ as was done in Lemma 110, but also using Lemma 117(2). Having thus established $x_0 \in \widehat{X}_0$ and $x_1 \in \widehat{X}_1$, the result is immediate by definition of $\sim_C$. ∎

The next lemma shows how $\sim_C$ relates C_0 and C_1:

Lemma 119 $^{(0\leftrightarrow 1)}$

(1) *If* $x_0 \sim_C x_1C_1q$ *then* x_0C_0q $(q \in P_4 \uplus Y)$
(2) *If* $x_0 \sim_C x_1C_1x_1'A_1p_1$ *then* $x_0C_0p_1$
(3) *If* $x_0 \sim_C x_1C_1x_1' \sim_C x_0'$ *then* $x_0C_0x_0'$.

Proof We shall prove (3); the others are similar. It is enough to treat the case in which both instances of $\sim_C$ take one of the three generating forms $x_0A_0p_2A_1x_1$, $x_0C_0p_1A_1x_1$ and $x_0A_0p_0C_1x_1$. In all (essentially six) combinations of these forms in the antecedent of (3), we deduce immediately one of the following, or a symmetric variant:

$$\begin{array}{lll} (i) & x_0A_0q(C_1 \circ A_1)p_1C_0x_0' & (q \in P_0 \uplus P_2) \\ (ii) & x_0A_0q(C_1 \circ A_1)q'A_0x_0' & (q, q' \in P_0 \uplus P_2) \\ (iii) & x_0C_0p_1(C_1 \circ A_1)p_1'C_0x_0' . & \end{array}$$

In each of these cases, we apply $C_1 \circ A_1 = C_0 \circ A_0$ and then the corresponding part of Lemma 101, also using the no-alias condition for A_0, to deduce that $x_0C_0x_0'$. ∎

We can now show that the mediator $\widehat{C}$ satisfies one of its required properties.

Proposition 120 (RPO mediator respects bounds) $\widehat{C} \circ B_i = C_i$ $(i = 0, 1)$.

Proof We shall treat the case $i = 0$. First note that, using the fact that $x_0 \sim [x_0]_\sim$, we can expand $\widehat{C} \circ B_0$ as follows:

$$\widehat{C} \circ B_0 = (\ddot{E}_0 \sqcup \widehat{C}_0 \sqcup \widehat{A}_1 \sqcup \widehat{C}_1 \sqcup \sim_C) \setminus X_1 \setminus \widehat{X} .$$

To prove $\widehat{C} \circ B_0 \subseteq C_0$: For each instance of $\widehat{C} \circ B_0$ there is a lub chain for $\{\ddot{E}_0, \widehat{C}_0, \widehat{A}_1, \widehat{C}_1, \sim_C\}$ in which members of X_1 and $\widehat{X}$ do not occur extremally. We have to eliminate instances of $\ddot{E}_0$, $\widehat{A}_1$, $\widehat{C}_1$ and $\sim_C$ from the chain. From Lemma 117 we have $\ddot{E}_0 \subseteq C_0$, so we need only consider the remaining three. We begin by eliminating $\widehat{C}_1$ links:

Case qC_1q' $(q, q' \in P_4 \uplus Y)$. Then $q(C_0 \circ A_0)q'$, hence by Lemma 101(3) $\boldsymbol{qC_0q'}$.

Case qC_1x_1 $(q \in P_4 \uplus Y,\ x_1 \in \widehat{X}_1)$. Then there is a right-adjacent link for x_1, in either in $\widehat{A}_1$ or $\sim_C$. If in $\widehat{A}_1$ it must be to some p_1; so by Lemma 101(3) we replace $qC_1x_1A_1p_1$ by $\boldsymbol{qC_0p_1}$. If the link is $x_1 \sim_C q'$

then $q' \neq x_1$, and if $q' \in \widehat{X}$ then no further adjacent link is possible, so $x_1 \sim_C x_0$ for some x_0; so by Lemma 119(1) we replace $qC_1x_1 \sim_C q'$ by $\boldsymbol{qC_0x_0 \sim_C q'}$.

Case $x_1C_1x_1'(x_1, x_1' \in \widehat{X}_1)$. Then there are left and right adjacent links in $\widehat{A}_1$ or $\sim_C$. There are essentially three cases for the three links. If they are $p_1A_1x_1C_1x_1'A_1p_1'$, then using $C_1 \circ A_1 = C_0 \circ A_0$ and by Lemma 101(3) we get $\boldsymbol{p_1C_0p_1'}$; if they are $p_1A_1x_1C_1x_1' \sim_C q$ with some $x_0 \sim_C x_1'$, then using $C_1 \circ A_1 = C_0 \circ A_0$ and by Lemma 101(1) we get $\boldsymbol{p_1C_0x_0 \sim_C q}$; finally, if the links are $q \sim_C x_1C_1x_1' \sim_C q'$ with some $x_0 \sim_C x_1$ and $x_0' \sim_C x_1'$, then by Lemma 119(3) we arrive at the replacement $\boldsymbol{q \sim_C x_0C_0x_0' \sim_C q'}$.

This completes the elimination of $\widehat{C}_1$. The elimination of $\widehat{A}_1$ is simpler:

Case $p_1A_1p_1'$. Then by $C_1 \circ A_1 = C_0 \circ A_0$ and Lemma 101(3) we get $\boldsymbol{p_1C_0p_1'}$.

Case $x_1A_1p_1$ $(x_1 \in \widehat{X}_1)$. There must be a left adjacent link $q \sim_C x_1$ (since it cannot be in $\widehat{A}_1$), and as argued above there must exist $x_0 \sim_C x_1$; by Lemma 119 we then replace $q \sim_C x_1A_1p_1$ by $\boldsymbol{q \sim_C x_0C_0p_1}$.

We are now left with a lub chain for C_0 and $\sim_C$, whose extremal members are not in X_1 or $\widehat{X}$; moreover no C_0 link involves these sets, so every remaining instance of $\sim_C$ can only be of the form $x_0 \sim_C x_0'$ which, by taking $x_1 = x_1'$ in Lemma 119(3), can be replaced by $\boldsymbol{x_0C_0x_0'}$. This completes the proof that $\widehat{C} \circ B_0 \subseteq C_0$.

To prove $C_0 \subseteq \widehat{C} \circ B_0$, we express each instance of C_0 in terms of $\{\ddot{E}_0, \widehat{C}_0, \widehat{A}_1, \widehat{C}_1, \sim_C\}$:

Case qC_0q' $(q, q' \in P_4 \uplus Y)$. Immediately $\boldsymbol{q\widehat{C}_0q'}$.

Case qC_0p_1 $(q \in P_4 \uplus Y)$. Using $C_0 \circ A_0 = C_1 \circ A_1$ and Lemma 101(1) we deduce $qC_1x_1A_1p_1$, with $x_1 \in \widehat{X}_1$ by Lemma 111(1); so $\boldsymbol{q\widehat{C}_1x_1\widehat{A}_1p_1}$.

Case qC_0x_0 $(q \in P_4 \uplus Y)$. Then qE_0x_0 since $E_0 = C \circ C_0$, and $x_0 \in \widehat{X}_0$ by Lemma 111(3); hence $\boldsymbol{q\widehat{C}_0x_0}$.

Case $p_1C_0p_1'$. Then using $C_0 \circ A_0 = C_1 \circ A_1$, by Lemma 101(2) either $p_1A_1p_1'$ whence $\boldsymbol{p_1A_1p_1'}$, or $p_1A_1x_1C_1x_1'A_1p_1'$, with $x_1, x_1' \in \widehat{X}_1$ by Lemma 111(1), so $\boldsymbol{p_1\widehat{A}_1x_1\widehat{C}_1x_1'\widehat{A}_1p_1}$.

Case $p_1C_0x_0$. Then $p_1E_0x_0$ since $E_0 = C \circ C_0$. If $x_0 \in \ddot{X}_0$ then $\boldsymbol{p_1\ddot{E}_0x_0}$. If $x_0 \in \widehat{X}_0$ then p_1 is open in A_1 by definition of $\widehat{X}_1$, so $x_1A_1p_1$. It follows from Lemma 118 that $x_1 \in \widehat{X}_1$ and $x_0 \sim_C x_1$, so we deduce $\boldsymbol{p_1\widehat{A}_1x_1 \sim_C x_0}$.

Case $x_0C_0x_0'$. Then $x_0E_0x_0'$ since $E_0 = C \circ C_0$. By definition of $\widehat{X}_0$, either $x_0, x_0' \in \widehat{X}_0$ or $x_0, x_0' \in \ddot{X}_0$; in the former case we deduce $\boldsymbol{x_0\widehat{C}_0x_0'}$, in the latter case $\boldsymbol{x_0\ddot{E}_0x_0'}$.

This concludes the proof that $C_0 \subseteq \widehat{C} \circ B_0$, and hence the proof of the proposition. ■

Now, since $(\vec{E}, \mathsf{id}_Z)$ is itself a candidate, we have also proved

Corollary 121 (RPO head respects bounds) $\widehat{E} \circ B_i = E_i$ $(i = 0, 1)$. *Hence* $(\vec{B}, \widehat{E})$ *is a candidate RPO for* $\vec{A}$ *w.r.t.* $\vec{E}$.

We have therefore completed our RPO construction:

Construction 122 (RPO head) We define the RPO of $\vec{A}$ w.r.t. $\vec{E}$ to be the triple $(\vec{B}, B)$, where $B \triangleq \widehat{E}$. ■

However, it remains to prove $\widehat{C}$ to be the unique edge net F such that $C \circ F = B$. We now address this. We begin with two crucial lemmas. The first relates two equivalences.

Lemma 123

(1) *If* $x_0 \sim_C x_1$ *then* $x_0 \sim_E x_1$
(2) *If* $x_0 \sim_E x_1$ *then either* $x_0 \sim_C x_1$ *or* $x_0\widehat{C}_0 yCy'\widehat{C}_1 x_1$ *for some* $y, y' \in Y$.

Proof In each case, we need only consider generating instances in the antecedent.

(1) For example, assume $x_0 C_0 p_1 A_1 x_1$. We have $C_0 \restriction (X_0 \uplus P_1) \subseteq E_0 \restriction (X_0 \uplus P_1)$ since $C \circ C_0 = E_0$; the rest follows.

(2) For example, assume $x_0 E_0 p_1 A_1 x_1$. From $x_0 E_0 p_1$, since $C \circ C_0 = E_0$ we deduce by Lemma 101(2) that either $x_0 C_0 p_1$, so $x_0 C_0 p_1 A_1 x_1$ and $x_0 \sim_C x_1$ as required, or $x_0 C_0 yCy' C_0 p_1$. In this case, from $y' C_0 p_1$ by $C_0 \circ A_0 = C_1 \circ A_1$ and Lemma 101(1), together with the no-alias condition for A_1, we deduce $y' C_1 x_1 A_1 p_1$; since $x_i \in \widehat{X}_1$ $(i = 0, 1)$, it immediately follows that $x_0\widehat{C}_0 yCy'\widehat{C}_1 x_1$. ∎

The second crucial lemma uses, for the first time, our assumption that one of $\vec{A}$, say A_1, has no idle names.

Lemma 124 *Assume that* A_1 *has no idle names. Let* $q_0 C_0 yCy' C_1 q_1$, *where* $q_i \in P_4 \uplus \widehat{X}_i$ $(i = 0, 1)$*. Then either* $q_0 = p_4 \in P_4$ *and* $q_0 \widehat{E}_1 q_1$, *or* $q_1 = p_4 \in P_4$ *and* $q_0 \widehat{E}_0 q_1$, *or* $q_i = x_i \in \widehat{X}_i$ $(i = 0, 1)$ *and* $x_0 \sim_C x_1$.

Proof If $q_0 = p_4 \in P_4$ then using $C_0 \circ A_0 = C_1 \circ A_1$ we find $p_4 C_1 yCy' C_1 q_1$, whence $q_0 \widehat{E}_1 q_1$. The case $q_1 \in P_4$ is similar.

We are left with the case $q_i = x_i$ $(i = 0, 1)$ and $x_0 C_0 yCy' C_1 x_1$. Since A_1 has no idle names, we have $x_1 A_1 q$ $(q \in P_1 \uplus P_2)$. If $x_1 A_1 p_1$ then by $C_1 \circ A_1 = C_0 \circ A_0$ and Lemma 101(3) we get $y' C_0 p_1$; hence $x_0 C_0 p_1$, so by Lemma 118 we deduce $x_0 \sim_C x_1$. The other possibility is $x_1 A_1 p_2$; then by $C_1 \circ A_1 = C_0 \circ A_0$ and Lemma 101(1) we get $y' C_0 x_0 A_0 p_2$, so by Lemma 110 we deduce $x_0 \sim x_1$, and again $x_0 \sim_C x_1$. ∎

We are now ready to complete the justification of our RPO construction.

Proposition 125 (unique RPO mediator) *Assume that* A_1 *has no idle names. Then* $\widehat{C}$ *is the unique edge net* $F : \widehat{X} \to Y$ *such that* $F \circ B_i = C_i$ $(i = 0, 1)$ *and* $C \circ F = B$.

Proof We have shown in Proposition 120 that $\widehat{C}$ satisfies the first equations; we now show that it satisfies the last. Recall that

$$\begin{aligned} B = \widehat{E} &= (\widehat{E}_0 \sqcup \widehat{E}_1 \sqcup \sim_E) \backslash X_0 \backslash X_1 \\ \widehat{C} &= (\widehat{C}_0 \sqcup \widehat{C}_1 \sqcup \sim_C) \backslash X_0 \backslash X_1 \\ &\text{and we deduce} \\ C \circ \widehat{C} &= (C \sqcup \widehat{C}_0 \sqcup \widehat{C}_1 \sqcup \sim_C) \backslash X_0 \backslash X_1 \backslash Y . \end{aligned}$$

To prove $C \circ \widehat{C} \subseteq B$, we first eliminate all C links from a lub chain for $\{C, \widehat{C}_0, \widehat{C}_1, \sim_C\}$ whose end-members are not in X_0, X_1 or Y.

Case qCq' $(q, q' \in P_5)$. Since $C \circ C_0 = E_0$, by Lemma 101(3) $\boldsymbol{q\widehat{E}_0 q'}$.

Case qCy $(q \in P_5)$. Then there is a right adjacent $\widehat{C}_i$ link, say $y\widehat{C}_0 q'$ with $q' \in P_4 \uplus \widehat{X}_0$ $(q' \notin Y$, else $q' = y)$. So as above we replace $qCy\widehat{C}_0 q'$ by $\boldsymbol{q\widehat{E}_0 q'}$.

Case yCy' $(y \neq y')$. Then there are left and right adjacent links $qC_i y$ and $y' C_j q'$, with (since C_0, C_1 have no aliases) $q \in P_4 \uplus \widehat{X}_i$ and $q' \in P_4 \uplus \widehat{X}_j$. We wish to replace $qC_i yCy' C_j q'$. If $i = j$ then $\boldsymbol{q\widehat{E}_0 q'}$. Now suppose, say, $i = 0$ and $j = 1$. Then by Lemma 124 we find the replacement $\boldsymbol{q_0 \widehat{E}_1 q_1}$ or $\boldsymbol{q_0 \widehat{E}_0 q_1}$, or else $q_i = x_i \in \widehat{X}_i$ $(i = 0, 1)$ and $\boldsymbol{x_0 \sim_C x_1}$.

Having eliminated C, we are left with an instance of $(\widehat{E}_0 \sqcup \widehat{E}_1 \sqcup \widehat{C}_0 \sqcup \widehat{C}_1 \sqcup \sim_C) \setminus X_0 \setminus X_1 \setminus Y$. Let us now eliminate $\widehat{C}_0$.

Case $q\widehat{C}_0 q'$ $(q, q' \in P_4 \uplus \widehat{X}_0)$. Then from $E_0 = C \circ C_0$ we get $\boldsymbol{q \widehat{E}_0 q'}$.

Case $q\widehat{C}_0 y$ $(q \in P_4 \uplus \widehat{X}_0)$. Then there is a right adjacent link $y\widehat{C}_i q'$ $(q' \in P_4 \uplus \widehat{X}_i)$, and we must replace $q\widehat{C}_0 y \widehat{C}_i q'$. If $i = 0$ then we have $q\widehat{C}_0 q'$ and get the replacement $\boldsymbol{q \widehat{E}_0 q'}$ as in the previous case. Otherwise, by Lemma 124 (with $y = y'$) we get the replacement $\boldsymbol{q_0 \widehat{E}_1 q_1}$ or $\boldsymbol{q_0 \widehat{E}_0 q_1}$, or else $q_i = x_i \in \widehat{X}_i$ $(i = 0, 1)$ and $\boldsymbol{x_0 \sim_C x_1}$.

This concludes the elimination of $\widehat{C}_0$ (recall that C_0 has no aliases, so the case $y\widehat{C}_0 y'$ would imply $y = y'$). $\widehat{C}_1$ is eliminated similarly. It only remains to note that $\sim_C \subseteq \sim_E$ by Lemma 123(1), and we have completed the proof that $C \circ \widehat{C} \subseteq B$.

To prove $B \subseteq C \circ \widehat{C}$, we first eliminate $\widehat{E}_0$ from a lub chain for $\{\widehat{E}_0, \widehat{E}_1, \sim_E\}$ with end-members not in X_0 or X_1. We use $E_0 = C \circ C_0$ frequently.

Case $p_5 E_0 p'_5$. Then by Lemma 101(3) $\boldsymbol{p_5 E_0 p'_5}$.

Case $p_5 E_0 q$ $(q \in P_4 \uplus \widehat{X}_0)$. Then by Lemma 101(1) $\boldsymbol{p_5 C y \widehat{C}_0 q}$.

Case $q E_0 q'$ $(q, q' \in P_4 \uplus \widehat{X}_0)$. Then by Lemma 101(2) either $p_4 C_0 p'_4$ so $\boldsymbol{p_4 \widehat{C}_0 p'_4}$, or $p_4 C_0 y C y' C_0 p'_4$ so $\boldsymbol{p_4 \widehat{C}_0 y C y' \widehat{C}_0 p'_4}$.

This concludes the elimination of $\widehat{C}_0$, and $\widehat{C}_1$ is treated the same. We now have a lub chain for $\{C, \widehat{C}_0, \widehat{C}_1, \sim_E\} \setminus X_0 \setminus X_1 \setminus Y$, and it only remains to replace any $\sim_E$ links. For this, it is enough to note that an instance of $\sim_E$ is representable by a chain of links of the form $x_i \sim_E [x_i]_\sim$, which is also an instance of $\sim_C$, or of the form $x_0 \sim_E x_1$, which by Lemma 123 can be represented by a combination of $C, \widehat{C}_0, \widehat{C}_1$ and $\sim_C$. This completes the proof that $C \circ \widehat{C} = B$.

To see that $\widehat{C}$ is the unique edge net F such that $F \circ B_i = C_i$ $(i = 0, 1)$ and $C \circ F = B$, we need merely observe that if F is another, then in particular $F \circ B_0 = C_0 = \widehat{C} \circ B_0$. But under our assumption that A_1 is epi (which we needed in Lemma 124) we also showed in Proposition 115 that B_0 is epi; hence from the foregoing equation we deduce $F = \widehat{C}$. ■

From Propositions 114 ,120 and 125 we deduce the main theorem:

Theorem 126 (monographs with one epi always have RPOs) *In the precategory* **Edg**$_{\bar{a}}$ *of edge nets without aliases, Constructions 109, 112 and 116 yield* $(\vec{B}, B)$ *as an RPO of* $\vec{A}$ *w.r.t.* $\vec{E}$*, provided only that one member of* $\vec{A}$ *is epi.*