Nordic Journal of Computing

Deriving Bisimulation Congruences using 2-categories

VLADIMIRO SASSONE
COGS
University of Sussex, United Kingdom

PAWEŁ SOBOCIŃSKI
BRICS
University of Aarhus, Denmark

Abstract. We introduce G-relative-pushouts (GRPO) which are a 2-categorical generalisation of relative-pushouts (RPO). They are suitable for deriving labelled transition systems (LTS) for process calculi where terms are viewed modulo structural congruence. We develop their basic properties and show that bisimulation on the LTS derived via GRPOs is a congruence, provided that sufficiently many GRPOs exist. The theory is applied to a simple subset of CCS and the resulting LTS is compared to one derived using a procedure proposed by Sewell.

CR Classification: F.3.2 Semantics of Programming Languages; F.3.1 Specifying and Verifying and Reasoning about Programs; D.1.3 Concurrent Programming .

Key words: bisimulation, congruence, observational and contextual equivalences, reduction and transition systems, relative pushout, 2-category

1. Introduction

Term rewriting is a cornerstone of sequential computation. The λ-calculus, for instance, is essentially a simple term rewriting system. Process calculi, which aim at modelling concurrent computation, may also be viewed as rewriting systems, where the rewrites, the so-called *reactions*, represent the systems' internal evolution. The setting is more complex, however, as terms are often quotiented by a non-trivial *structural congruence*, which is a relation expressing which different syntactic representations describe the same process.

Park's notion of *bisimulation* [Park 1981; Milner 1989] underpins a multitude of operational preorders and equivalences which allow reasoning about concurrent processes modelled in a particular process calculus. These rely on the presence of a *labelled transition system* (LTS) which may be seen as a description of how processes interact with their environment. A LTS is a description of what may be observed about processes, for this reason bisimulation is often called an *observational equivalence*. Such equivalences are most useful when they are congruences, as this allows equational reasoning and full substitutivity.

Supported by EU IHP Marie Curie Training Site '**DisCo**: Semantic Foundations of Distributed Computation', contract HPMT-CT-2001-00290, and by **BRICS** (`www.brics.dk`), Basic Research in Computer Science, funded by the Danish National Research Foundation.

Received March 14, 2003.

Reaction systems and LTS usually coexist. The former are more easily postulated, as they tend to describe systems behaviour directly by focusing on the interactions between different parts, and therefore correspond closely to the calculus designer's computational intuitions. Indeed, one can derive sensible process equivalences using reactions. One such approach is the *barbed congruence* of Milner and Sangiorgi [1992], which descends naturally from the sole choice of a specific notion of observable (a "barb"). A related approach is by Honda and Yoshida [1995] who, based on intuitions from the λ-calculus, build equational theories directly from rewrites requiring no a priori specification of observables.

On the contrary, LTS are often intensional: they aim at describing observable behaviours in a compositional way and, therefore, their labels may not be immediately justifiable in operational terms. In other words, it may not be obvious to identify a "natural" LTS for a given process calculus, even when its semantics is well understood. For example there are two alternative LTS for the π-calculus [Milner *et al.* 1992], the early and the late version, each giving a different bisimulation equivalence. Furthermore, once a LTS is given, it is usually non trivial to prove that bisimulation is a congruence.

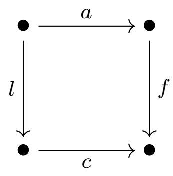
Due to the versatility of coinduction, LTS play a relevant role in applications. It is then important to be able to establish a correspondence between the two approaches. In particular, one may try to *synthesise* the LTS from the set of reactions. In a seminal work, Sewell [1998] proposed several ways of doing this for restricted classes of term rewriting systems. The common idea is that certain contexts which allow reaction are taken as labels. Consider for instance the rewrite system with signature $\Sigma = \{a, b, c\}$ where a is a unary function, b and c are constants and a single rewrite rule that transforms ab into c, written $\langle ab, c\rangle$. A possible LTS transition might then be

$$b \xrightarrow{a} c.$$

One needs however to be careful and selective about choosing the contexts and as a rule take only those which are the '*smallest*' allowing a particular reaction to occur, as this has a defining effect on the resulting bisimulation equivalence. Taking all possible contexts as labels gives larger LTS and results in equivalences which are too *coarse* even in the case of very simple process calculi. In principle, considering more labels in the bisimulation game delivers finer equivalences (i.e. equivalences which make less identifications); not so here, where the 'minimal' contexts characterise precisely the interaction between terms and effectively subsume the others. Intuitively, the observer who knows that contexts are essential and as small as possible for a particular reaction, knows more than the observer who can just take notice of reactions. Consider for instance $b \xrightarrow{aa} ac$. The redundant a in the label aa contributes nothing to the understanding of the term b. On the contrary, as will be clear in §3, it may garble the resulting bisimulation by introducing unjustified equivalences and make it much too coarse.

Sewell's method tackles that using dissection lemmas which provide a deep analysis of a term's structure, determining the missing triggers, if any. The proofs that bisimulation is a congruence on the resulting LTS is simple enough in the case of free syntax, but gets very complicated as soon as non trivial structural congruences are considered. Already in the case of the monoidal rules that govern parallel composition things become rather involved: the dissection method does not seem to scale to complex calculi.

A generalised approach was later developed by Leifer and Milner [2000], where the notion of smallest is formalised in categorical terms as a *relative-pushout* (RPOs). Informally, consider a category in which arrows are terms and composition is substitution. In such a framework a context f that allows a to react according to rule $\langle l, r\rangle$ can be given as a commuting square:

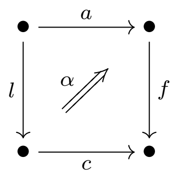


One derives a LTS by taking as labels those contexts f which make such squares "minimal." The proof of congruence takes form as a theorem in pure category theory, requiring only the existence of "enough" RPOs.

Again, applying the theory of RPOs to "categories of terms" fails for process calculi with even simple structural congruences. One problem is that considering a commuting square like above when arrows are terms quotiented by a structural congruence, we lose too much information for the RPO approach to give the expected results. In particular, as we shall discuss in detail in §3, in a simple calculus with a parallel operator which is associative and commutative we lose the information of *where* in the term the reaction occurs. The indication of exactly which occurrences of a term constructor belong to the redex is fundamental in order to derive sensible LTS. The same problem arises when we replace syntactic terms by algebraic structures such as action graphs [Milner 1996] and bigraphs [Milner 2001]. Indeed, essentially because of the problem of locating reactions, sufficient RPOs usually do not exist [Leifer 2001; Milner 2001].

For syntactic terms, Sewell [1998] proposes to deal with this issue using a notion of colouring, while Leifer [2002] suggests an abstract approach via *functorial reactive systems* and *precategories*. The technique presented here offers an alternative to the latter; we offer a brief, informal comparison with related work in §8. We believe that our approach is substantially simpler.

The approach proposed in this paper is to keep the information of "how the square commutes" by retaining the derivation of structural congruence. In this case, we think of the structural congruence as a set of rules governing how a term tree may be altered without changing the process (or in the algebraic case, as a set of suitable structure preserving isomorphisms). This information naturally gives a 2-categorical structure, where a 2-cell like the one below represents a "derivation" of the structural congruence of fa and cl.



Indeed, this geometric point of view is close to the spatial metaphor in the Chemical Abstract Machine of Berry and Boudol [1992] which served as an inspiration for the use of structural congruences in the operational semantics of process calculi. A similar approach, also based on 2-categories, has been identified and informally put forward by Sewell [2000] as a generalisation of colouring.

Observe that, since the structural alteration of term trees is always reversible in this setting, in our categories all 2-cells are isomorphisms. For such *locally groupoidal 2-categories*, G-categories for short, we propose a suitable generalisation of RPO, the GRPO, prove that it enjoys the fundamental properties of RPOs and, in particular, that bisimulation on a LTS derived using GRPOs is a congruence if sufficiently many GRPOs exist.

This paper completes [Sassone and Sobociński 2002] with the inclusion of all proofs and the exposition of the relationship between GRPOs and bicolimits, and provides the formal details of the extensions of the constructions of Leifer and Milner [2000] and their properties to a 2-categorical setting.

Structure of the paper. In §2 we review RPOs and Sewell's derivation of LTS for ground rewriting systems on free syntax, illustrating the relationship between the two approaches. In §3 we show in detail why the RPO approach fails when terms are viewed modulo structural congruence, and why a 2-categorical approach may be desirable. In §4 we recall the basic elements of the theory of 2-categories and their bicolimits, and we introduce the notion of bi-relative pushouts (biRPOs) in total generality. In §5 we specialise biRPOs to G-categories, so yielding the central notion of GRPOs, and we develop their theory; §6 illustrates the proof of congruence. In §7 we apply the theory to derive LTS for a simple fragment of CCS and compare the results to those of Sewell's approach. Finally, in section §8 we conclude by offering possible directions for future work.

2. Relative Pushouts

In this section we give a brief review of the theory of RPOs, a more complete presentation may be found in Leifer [2001]. We end with a comparison to the work of Sewell [1998] for ground rewriting systems on free syntax.

Consider a signature Σ. The (free) Lawvere theory for Σ [Lawvere 1963], denoted as $\mathsf{Th}(\Sigma)$, is a category with objects natural numbers and morphisms $t\colon m \to n$ being n-tuples of m-holed terms. Composition is substitution of terms. The category is cartesian, with 0 the terminal object and n being the product of 1 with itself n times. Identities $n \to n$ are $\langle -_1, -_2, \dots, -_n \rangle$. The theory is free in the sense that there are no equations between composite terms, apart from those imposed by the cartesian structure. A morphism $t\colon m \to n$ is linear if each of the m "holes" is used exactly once in t. Let $\mathbf{C}_\Sigma$ denote the subcategory of $\mathsf{Th}(\Sigma)$ consisting of the linear morphisms.

A term rewriting system can be given as a set $\mathcal{R}$ of pairs $\langle l, r\rangle$ where $l, r : n \to 1$ are arrows of $\mathbf{C}_\Sigma$.[1] The reaction relation $\longrightarrow\!\triangleright$ is derived from $\mathcal{R}$ by substitution under contexts, that is $a \longrightarrow\!\triangleright a'$ if $a = cl$, $a' = cr$ for some $c \in \mathbf{C}_\Sigma$. A term rewriting system is a ground term rewriting system when $\mathcal{R}$ consists only of pairs $\langle l, r\rangle$ with $l, r : 0 \to 1$.

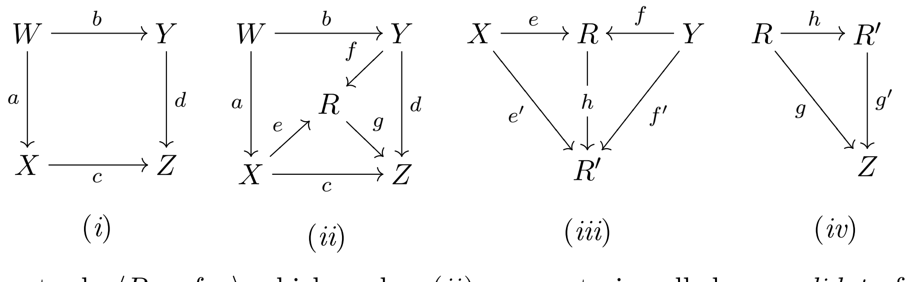
Generalising from ground term rewriting systems on $\mathbf{C}_\Sigma$, we give the definition of *reactive system* from Leifer and Milner [2000].

Definition 1. (Reactive System) A reactive system $\mathbf{C}$ consists of a category $\mathbb{C}$, a composition-reflecting subcategory $\mathbb{D}$ of *reactive contexts*, a distinguished object $I \in \mathbb{C}$ and a set or pairs $\mathcal{R} \subseteq \bigcup_{C \in \mathbb{C}} \mathbb{C}(I, C) \times \mathbb{C}(I, C)$.

By composition-reflecting we mean that $dd' \in \mathbb{D}$ implies d and $d' \in \mathbb{D}$. The reactive contexts are those contexts inside which evaluation may occur. The reaction relation $\longrightarrow\!\triangleright$ is derived from $\mathcal{R}$ by closing it under all reactive contexts. For simplicity, we shall henceforward assume that all contexts are reactive, that is, $\mathbb{D} = \mathbb{C}$. This will be the case for all the examples mentioned in this paper, while the proof of congruence needs only to be altered slightly to accommodate a smaller $\mathbb{D}$.

The notion of RPO formalises the idea of a context being the "smallest" that enables a reaction in a reactive system.

Definition 2. (RPO) Let $\mathbf{C}$ be a reactive system and (i) a commuting diagram in $\mathbf{C}$.



Any tuple $\langle R, e, f, g\rangle$ which makes (ii) commute is called a *candidate* for (i). A relative pushout is the "smallest" such candidate. More formally, it satisfies the universal property that given any other candidate $\langle R', e', f', g'\rangle$, there exists a *unique* mediating morphism $h\colon R \to R'$ such that (iii) and (iv) are commuting.

[1] For many applications it is reasonable to expect only l to be linear. Since we concentrate mainly on ground term rewriting, we do not elaborate here.

Another way of viewing RPOs is as ordinary pushouts in a slice-category. Indeed, the commuting square (i) above is simply a span

$$(X, c) \xleftarrow{a} (W, ca) \xrightarrow{b} (Y, d)$$

in the slice category $\mathbb{C}/Z$. It is straightforward to verify that to give a relative pushout of (i) above is to give a pushout of the span in $\mathbb{C}/Z$.

Definition 3. (IPO) A commuting square like (i) of Definition 2 is a idem-relative-pushout (IPO) if $\langle Z, c, d, \mathrm{id}_Z \rangle$ is its RPO.

For $\mathbf{C}$ a reactive system, a labelled transition system $\mathbf{TS}(\mathbf{C})$ can be derived using IPOs as follows:

- the states of $\mathbf{TS}(\mathbf{C})$ are arrows $a \colon I \to X$ of $\mathbf{C}$;
- there is a transition $a \xrightarrow{b} cr$ in $\mathbf{TS}(\mathbf{C})$ if and only if $\langle l, r \rangle \in \mathcal{R}$ and

$$\begin{array}{ccc} I & \xrightarrow{a} & X \\ {\scriptstyle l}\downarrow & & \downarrow{\scriptstyle b} \\ Y & \xrightarrow[c]{} & Z \end{array} \qquad \text{is an IPO.}$$

In other words, if insertion in context b makes a match l in context c (commutation of the diagram), where l is a redex, and b is the "smallest" such context (IPO condition), then a moves to cr with label b, where r is the reduct of l.

A reactive system $\mathbf{C}$ is said to *have redex RPOs* if every commuting square $cl = ba$ in $\mathbb{C}$, where $\langle l, r \rangle \in \mathcal{R}$, has an RPO. If this condition is satisfied, then $\sim$, the largest bisimulation on $\mathbf{TS}(\mathbf{C})$, is a congruence [Leifer and Milner 2000]. This result is generalised in this paper to a 2-categorical notion of RPOs (cf. Theorem 1).

Often it is desirable to consider only terms a of a fixed arity $I \to T$ and labels of type $T \to T$. Let $\mathbf{TS}(\mathbb{C})_T$ denote the labelled transition system so obtained and let $\sim_T$ be corresponding bisimulation. If $\mathbb{C}$ has redex RPOs, then it follows from the general proof that $\sim_T$ is also a congruence. Clearly, $\sim \subseteq \sim_T$, and the converse does not hold in general.

For ground term rewriting on $\mathbf{C}_\Sigma$, Sewell derives a LTS $\mathbf{Sew}(\mathbf{C}_\Sigma)$ with states being terms $a : 0 \to 1$ and labels $f : 1 \to 1$ as follows:

- $s \xrightarrow{-} t$ iff $s \longrightarrow\!\!\rhd\, t$
- $s \xrightarrow{f} t$ iff there is $\langle l, r \rangle \in \mathcal{R}$ such that $fs = l$ and $r = t$ (for $f \neq -$).

The two definitions are related. Indeed, using dissections [Sewell 1998], one can prove that redex RPOs exist in $\mathbf{C}_\Sigma$. The following lemma is due to Sewell [2000].

LEMMA 1. $\mathbf{TS}(\mathbf{C}_\Sigma)_1 = \mathbf{Sew}(\mathbf{C}_\Sigma)$.
PROOF. It suffices to show that

$$\begin{array}{ccc} 0 & \xrightarrow{a} & 1 \\ {\scriptstyle l}\downarrow & & \downarrow{\scriptstyle c} \\ 1 & \xrightarrow[d]{} & 1 \end{array}$$

is an IPO iff either $c = -$ or $d = -$. Indeed, suppose that $d \neq -$ and $c \neq -$. Then d and c, viewed as term trees, contain a topmost node labelled by $\sigma : n \to 1$ where $\sigma \in \Sigma$. This σ can be used to construct a non-trivial candidate for the diagram above, contradicting the assumption that the square is an IPO. □

3. Structural Congruence

In this section we discuss the motivation for a notion of relative-pushout in a 2-categorical setting.

EXAMPLE 1. Consider the following simple subset of CCS:

$$P \ ::= \ \mathbf{0} \ \mid \ a \ \mid \ \overline{a} \ \mid \ P \,\|\, P' \qquad\qquad \text{where } a \in N.$$

The signature consists of constants for channel names and for the null process and a binary operator, that is $\Sigma = \{\mathbf{0}, a, \overline{a}, -_1 \,\|\, -_2\}$. The reaction relation is the closure of the relation $\{\,(a \,\|\, \overline{a}, \mathbf{0}) \ \mid \ a \in N\,\}$ under all contexts. The standard operational semantics can be summarised by the following rules,

$$a \xrightarrow{a} \mathbf{0} \qquad\qquad \overline{a} \xrightarrow{\overline{a}} \mathbf{0} \qquad\qquad a \,\|\, \overline{a} \xrightarrow{\tau} \mathbf{0}$$

$$\frac{P \xrightarrow{x} P'}{Q \,\|\, P \xrightarrow{x} Q \,\|\, P'} \qquad\qquad \frac{P \equiv P' \quad P \xrightarrow{x} Q \quad Q' \equiv Q}{P' \xrightarrow{x} Q'}$$

where $\equiv$ is the smallest congruence on the set of terms over Σ which makes $\|$ an action of a commutative monoid with $\mathbf{0}$ as identity.

Let $\mathbf{D}_\Sigma$ be a category with the same objects as $\mathbf{C}_\Sigma$ but whose arrows are terms quotiented by $\equiv$. One may ask what happens if we use the RPO approach to generate an LTS. Consider the term $a \,\|\, \overline{a}$. Using the standard operational semantics we should expect three transitions,

$$a \,\|\, \overline{a} \xrightarrow{a} \overline{a}, \qquad a \,\|\, \overline{a} \xrightarrow{\overline{a}} a \qquad \text{and} \qquad a \,\|\, \overline{a} \xrightarrow{\tau} \mathbf{0}.$$

Consider the three squares in $\mathbf{D}_\Sigma$ below, where we use subscripts to distinguish different occurrences of the term a (that may float around in larger terms because of $\equiv$). Observe that such distinction is for the sake of exposition only: arrows in $\mathbf{D}_\Sigma$ up to structural congruence, and therefore individual occurrences of terms are not discernible.

$$
\begin{array}{ccc}
0 & \xrightarrow{a_1|\overline{a}_2} & 1 \\
{\scriptstyle a_1|\overline{a}_2}\downarrow & & \downarrow{\scriptstyle -} \\
1 & \xrightarrow[-]{} & 1
\end{array}
\qquad
\begin{array}{ccc}
0 & \xrightarrow{a_1|\overline{a}_2} & 1 \\
{\scriptstyle \overline{a}_2|a_3}\downarrow & & \downarrow{\scriptstyle -|a_3} \\
1 & \xrightarrow[a_1|-]{} & 1
\end{array}
\qquad
\begin{array}{ccc}
0 & \xrightarrow{a_1|\overline{a}_2} & 1 \\
{\scriptstyle \overline{a}_3|a_1}\downarrow & & \downarrow{\scriptstyle \overline{a}_3|-} \\
1 & \xrightarrow[-|\overline{a}_2]{} & 1.
\end{array}
$$

Only the left one could possibly be an IPO, and it is easy to see that it is a *candidate* for the middle and the right squares. However, also the upper bounds given in the middle and the right square are in some sense *minimal.* Indeed, if we keep track of the place in the term where the reaction occurs, then the middle square is the smallest upper bound whose redex (viz. $\overline{a}_2 \mid a_3$) *only* uses $\overline{a}$ (as opposed to both a and $\overline{a}$) from the term. Similarly, in the right square the redex created by insertion into a context (viz. $\overline{a}_3 \mid -$) only uses a. It is precisely the fact that terms in $\mathbf{D}_\Sigma$ are quotiented by $\equiv$ that makes it impossible to place reaction within a term.

For many applications it makes sense to have the extra power of placing reaction. Indeed, the reader may verify that the LTS generated using RPOs on $\mathbf{D}_\Sigma$ generates the same set of labels for the terms $a \mid \overline{a}$ and $b \mid \overline{b}$ and thus no operational equivalence can distinguish between these two terms. That is, against the intuition, $a \mid \overline{a}$ and $b \mid \overline{b}$ would be bisimilar. However, when reaction can be placed, bisimulation equivalence coincides with the standard one (viz. Examples 2, 3 and 4).

At this point it is important to focus on what exactly is a commuting square in $\mathbf{D}_\Sigma$. To verify that a diagram like (i) below is commuting one has to exhibit a proof of structural congruence constructed from the basic rules closed under all contexts.

$$
\begin{array}{ccc}
k & \xrightarrow{p} & l \\
{\scriptstyle q}\downarrow & & \downarrow{\scriptstyle r} \\
m & \xrightarrow[s]{} & n
\end{array}
\qquad\qquad
\begin{array}{ccc}
k & \xrightarrow{p} & l \\
{\scriptstyle q}\downarrow & \overset{\rho}{\nearrow\!\!\!\nearrow} & \downarrow{\scriptstyle r} \\
m & \xrightarrow[s]{} & n
\end{array}
$$

(i) $\qquad$ (ii)

Different proofs can be chosen to exhibit commutativity. Indeed, as we exemplify later in §5, with a bit of massaging, such "proofs" can be represented as 2-cells and used to give a 2-categorical structure on $\mathbf{C}_\Sigma$.

Along with a suitable 2-categorical notion of a relative-pushout we get a natural notion of "smallest" which remembers the location of the redex and that for the example above works much the same way as Sewell's colouring of terms [Sewell 1998]. However, since our definition is abstract, it provides the framework to approach process calculi with structural congruences different from $(|, \mathbf{0})$.

In the following sections we give the details of our generalisation of RPOs to 2-categories and we show that they enjoy the congruence properties of their 1-dimensional cousins.

4. 2-categories, Bicolimits, and biRPOs

We start by recalling the definition of 2-categories and bicolimits. For a thorough introduction we refer the reader to Kelly and Street [1974].

A ***2-category*** $\mathbb{C}$ is a category enriched over Cat [Kelly 1982], the category of (small) categories and functors. That is, $\mathbb{C}$ is a category whose homsets are categories and, correspondingly, whose composition maps are functors. In explicit terms, a 2-category $\mathbb{C}$ consists of what follows.

- A class of *objects* $X, Y, Z, \ldots$.
- For any $X, Y \in \mathbb{C}$, a category $\mathbb{C}(X, Y)$. The objects $\mathbb{C}(X, Y)$ are called *1-cells*, or simply arrows, and denoted by $f\colon Y \to X$. Its morphisms are called *2-cells*, are written $\alpha\colon f \Rightarrow g$ and drawn as

  $$X \underset{g}{\overset{f}{\Downarrow\alpha}} Y.$$

  Composition in $\mathbb{C}(X, Y)$ is denoted by $\bullet$ and referred to as '*vertical*' composition. Identity 2-cells are denoted by $\mathbf{1}_f\colon f \Rightarrow f$.
- For each X, Y, Z there is a functor $.\colon \mathbb{C}(X, Y) \times \mathbb{C}(Y, Z) \to \mathbb{C}(X, Z)$, the so-called '*horizontal*' composition, which we shall usually denote by juxtaposition. Horizontal composition is associative and admits $\mathbf{1}_{\mathrm{id}_X}$ as identities.
  Note that the functoriality of . can be spelled out as follows: for all $\alpha\colon f \Rightarrow g$, $\beta\colon g \Rightarrow h$, and $\gamma\colon u \Rightarrow v$, $\delta\colon v \Rightarrow w$, where $f, g, h\colon X \to Y$ and $u, v, w\colon Y \to Z$, we have $(\delta \bullet \gamma)(\beta \bullet \alpha) = \delta\beta \bullet \gamma\alpha$ and $\mathbf{1}_g.\mathbf{1}_f = \mathbf{1}_{gf}$. The former equation is known as the *middle-four interchange law*. As a notation, we write αf and $g\alpha$ for, respectively, $\alpha\mathbf{1}_f$ and $\mathbf{1}_g\alpha$.

Syntactically, we follow the convention that horizontal composition binds tighter than vertical composition.

The canonical example of a 2-category is Cat, the 2-category of categories, functors and natural transformations. Given a 2-category $\mathbb{C}$ we shall sometimes refer to the "underlying category" $\mathbb{C}_1$, which is the category obtained from $\mathbb{C}$ by forgetting its 2-cells.

The remainder of this section is a brief introduction to the notion of ***bicolimits*** in 2-categories. The reader not interested in the abstract categorical notions underlying the theory of GRPOs can safely skip ahead to §5.

Bicolimits were introduced by Street [1980] in the context of *bicategories* [Bénabou 1967], which consist of the same basic data as 2-categories where associativity and identity laws for 1-cells hold *only up to* coherent isomorphisms. Any 2-category may be thought of as a special kind of bicategory (where the coherent isomorphisms are all identities), and indeed the notion of bicolimit applies to 2-categories as well. Bicolimits are also briefly discussed by Kelly [1989] along with other notions of 2-categorical (co)limits. We remark that our exposition here is simplified, as the more general notion of *indexed-bicolimits* is not needed in this paper.

A ***2-functor*** $F : \mathbb{C} \to \mathbb{D}$ maps objects to objects, arrows to arrows and 2-cells to 2-cells in a way which respects all identities and composition.

A ***pseudo-natural transformation*** α between 2-functors $F, G : \mathbb{C} \to \mathbb{D}$ consists of an arrow $a_C : FC \to GC$ for every object $C \in \mathbb{C}$, and invertible 2-cells a_f for every arrow f in $\mathbb{C}$ with domain and codomain as below,

$$\begin{array}{ccc} FC & \xrightarrow{a_C} & GC \\ {\scriptstyle Ff}\downarrow & \overset{a_f}{\swarrow\!\!\!\swarrow} & \downarrow{\scriptstyle Gf} \\ FC' & \xrightarrow[a_{C'}]{} & GC', \end{array}$$

such that for any 2-cell $\alpha : f \to g$ in $\mathbb{C}$, we have $a_{C'}F\alpha \bullet a_f = a_g \bullet G\alpha.a_C$. Additionally, we require that $a_{\mathrm{id}} = \mathbf{1}$ and that a_{gf} is the pasting composite of a_f and a_g.

A ***modification*** ξ between two pseudo-natural transformations a and b is a family of 2-cells $\xi_C : a_C \Rightarrow b_C$ suitably compatible with a_f and b_f, i.e. $b_f \bullet Gf.\xi_C = \xi_{C'}Ff \bullet a_f$.

2-functors from $\mathbb{C}$ to $\mathbb{D}$, pseudo-natural transformations and modifications form a *2-category* that we will indicate by $\mathbf{Psd}[\mathbb{C}, \mathbb{D}]$. For any 2-category $\mathbb{C}$, we shall denote by $\mathbf{!}$ the unique 2-functor to a terminal 2-category, that is a 2-category with one object, one identity arrow and one identity 2-cell.

Two objects C, D of a 2-category $\mathbb{C}$ are *equivalent* when there are arrows $f : C \to D$, $g : D \to C$ and 2-cells $\alpha : \mathrm{id}_C \Rightarrow gf$, $\beta : fg \Rightarrow \mathrm{id}_D$. We refer to f and g as equivalences. Given any equivalence f in the 2-category Cat, one may always choose the 2-cells α and β so that they form the unit and the counit of an *adjunction*, which we shall refer to as an ***equivalence of categories***.

Definition 4. (Bicolimit) Let $G : \mathbb{J} \to \mathbb{C}$ be a 2-functor. A ***bicolimit*** object of G is an object $\mathrm{Bic}\, G$ of $\mathbb{C}$ which satisfies

$$\mathbb{C}(\mathrm{Bic}\, G, A) \simeq \mathbf{Psd}[\mathbb{J}^{\mathrm{op}}, \mathsf{Cat}](\mathbf{!}, \mathbb{C}(G-, A)) \qquad (1)$$

where $\simeq$ denotes an equivalence of categories natural in $A \in \mathbb{C}$.[2] We shall

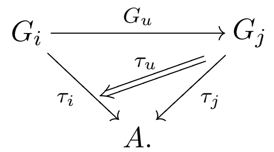
[2] This means that the functor $H = \lambda A.\, \mathbf{Psd}[\mathbb{J}^{\mathrm{op}}, \mathsf{Cat}](\mathbf{!}, \mathbb{C}(G-, A)) : \mathbb{C} \to \mathsf{Cat}$ admits a *birepresentation*, that is an object $\mathrm{Bic}\, G \in \mathbb{C}$ such that $\lambda A.\mathbb{C}(\mathrm{Bic}\, G, A)$ and H are equivalent as objects of $[\mathbb{C}, \mathsf{Cat}]$, the 2-category of functors $\mathbb{C} \to \mathsf{Cat}$, 2-natural transformations and modifications.

usually refer to the bicolimit of G as the pair $\langle \mathrm{Bic}\, G, \eta \rangle$ where η is the unit of (the image of the object $\mathrm{id}_{\mathrm{Bic}\, G}$ under the equivalence) (1).

Bicolimits are defined *up to equivalence*. Certainly, an equivalence $f : C \to D$ induces an equivalence of categories $\mathbb{C}(D, A) \simeq \mathbb{C}(C, A)$ natural in A. This, together with the fact that equivalences of categories compose, proves that C is a bicolimit of G if and only if D is. It also holds that if C and D are two bicolimits of G, they must be equivalent as objects of $\mathbb{C}$.

We can spell out Definition 4 in elementary terms. We use the fact that a functor is an equivalence of categories iff it is fully-faithful and essentially surjective on objects. It is then not difficult to verify that Definition 4 and Definition 5 are equivalent.

Definition 5. (Bicolimit) Let $G : \mathbb{J} \to \mathbb{C}$ be a 2-functor. A *pseudo-cocone* τ from G to $A \in \mathbb{C}$ is a family of arrows $\tau_i : G_i \to A$ for $i \in \mathbb{J}$ and invertible 2-cells $\tau_u : \tau_j G_u \Rightarrow \tau_i$ for $u : i \to j$ in $\mathbb{J}$ as illustrated below:



Additionally, $\tau_{\mathrm{id}} = \mathbf{1}$ and τ_{vu} must be the pasting composite of τ_u and τ_v. A bicolimit is a tuple $\langle \mathrm{Bic}\, G, \eta \rangle$ where $\mathrm{Bic}\, G \in \mathbb{C}$ and η is a pseudo-cocone from G to $\mathrm{Bic}\, G$. Two properties are required to ensure equivalence (1):

(a) For any pseudo-cocone α from G to A there exists an arrow $h : \mathrm{Bic}\, G \to A$ and a family of invertible 2-cells $\varphi_i : h\eta_i \Rightarrow \alpha_i$ which make the pseudo-cocones η and α compatible.

(b) Given another arrow $h' : \mathrm{Bic}\, G \to A$ and a family of 2-cells $\psi_i : h\eta_i \Rightarrow h'\eta_i$ which makes the two pseudo-cocones compatible, there exists a unique 2-cell $\xi : h \Rightarrow h'$ such that $\psi_i = \xi\eta_i$.

We experiment with a particular choice of G to examine the 'bi'-version of pushouts, as this will be useful in the definition of biRPOs.

Proposition 1. (Bipushout) Suppose that $\mathbb{J} = \cdot \leftarrow \cdot \rightarrow \cdot$. Then to give a 2-functor $G : \mathbb{J} \to \mathbb{C}$ is to give a cospan in $\mathbb{C}$.

A ***bipushout*** of arrows $X \xleftarrow{a} W \xrightarrow{b} Y$ is a quadruple $\langle Z, c, d, \rho \rangle$ where $c \colon X \to Z$, $d \colon Y \to Z$ and $\rho \colon ca \Rightarrow db$ is an isomorphism such that, for any other such quadruple $\langle Z', c', d', \rho' \rangle$:

(a) There exists an arrow $u \colon Z \to Z'$ and isomorphisms $\varphi \colon c' \Rightarrow uc$, $\psi \colon ud \Rightarrow d'$ satisfying the obvious compatibility condition, namely $\psi b \bullet u\rho \bullet \varphi a = \rho'$.

(b) For any other arrow $u' \colon Z \to Z'$ and 2-cells $\eta \colon u'c \Rightarrow uc$, $\mu \colon u'd \Rightarrow ud$ satisfying $u\rho \bullet \eta a = \mu b \bullet u'\rho$, there exists a unique $\xi \colon u' \Rightarrow u$ such that $\eta = \xi c$ and $\mu = \xi d$.

Analogously to RPOs being pushouts in a slice-category, we define a biRPOs to be bipushouts in a pseudo-slice category as below. In the next section we specialise biRPOs to a special kind of 2-categories, the G-categories, and provide a definition in elementary terms for the notion of GRPOs that so arises.

Given a 2-category $\mathbb{C}$ and an object Z, a ***pseudo-slice*** category $\mathbb{C}/Z$ is a 2-category with

- objects $C \xrightarrow{f} Z$;
- arrows $(C, f) \xrightarrow{(h,\epsilon)} (D, g)$ where $h \colon C \to D$ and $\epsilon \colon f \Rightarrow gh$ is an isomorphism; and
- 2-cells $\xi \colon (h, \epsilon) \Rightarrow (h', \epsilon')$ being 2-cells $\xi \colon h \Rightarrow h'$ satisfying the obvious compatibility requirement, namely $g\xi \bullet \epsilon = \epsilon'$.

Definition 6. (Bi-Relative Pushout) Let $\rho : ca \Rightarrow db : W \to Z$ be a 2-cell in a 2-category $\mathbb{C}$ (cf. diagram (i) in Definition 8). A ***bi-relative pushout*** (biRPO) for ρ is a bipushout of the pair or arrows $(a, \mathbf{1}) : ca \to c$ and $(b, \rho) : ca \to d$ of $\mathbb{C}/Z$.

5. GRPOs uncovered

Since the role of 2-cells in our approach is to represent (proofs of) structural congruences, we shall usually consider 2-categories whose 2-cells are all *isomorphisms*. As the categories all of whose morphisms are iso are commonly known as *groupoids*, our 2-categories are precisely the *groupoid-enriched* categories.

Definition 7. (G-Category) A G-***category*** is a category enriched over G, the category of groupoids.

Example 2. Consider the subset of CCS introduced in Example 1. The usual structural congruence rules assert that

$$P \mid (Q \mid R) \equiv (P \mid Q) \mid R,\ \ P \mid \mathbf{0} \equiv P \text{ and } P \mid Q \equiv Q \mid P.$$

Let $\mathbf{M}_\Sigma$ be the G-category with:

- a single object I;
- arrows strings $a_1 \mid a_2 \mid \ldots \mid a_n$, $a_i \in N$ with composition by juxtaposition (eg. $(a_3 \mid a_4)(a_1 \mid a_2) = a_1 \mid a_2 \mid a_3 \mid a_4$) and the empty string denoted by $\mathbf{0}$ serving as the identity;

- 2-cells permutations; namely, each arrow $a_1 \mid a_2 \mid \dots \mid a_n$ is the source of $n!$ 2-cells determined by the permutations $\varphi\colon [n] \to [n]$, where $[n] = \{1, 2, \dots, n\}$. Each such φ determines a 2-cell

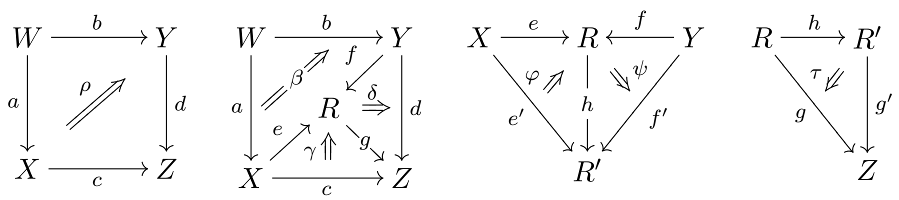
$$\varphi_{a_1,a_2,\dots,a_n}\colon a_1 \mid a_2 \mid \cdots \mid a_n \Rightarrow a_{\varphi^{-1}(1)} \mid a_{\varphi^{-1}(2)} \mid \cdots \mid a_{\varphi^{-1}(n)}.$$

  For clarity we will usually leave out the subscripts. So, for example, there are two 2-cells $a \mid a \Rightarrow a \mid a$: the identity, and the permutation that "swaps" the two as. Vertical composition is via composition of permutations, horizontal composition is via "juxtaposition," i.e. for $\varphi\colon [m] \to [m]$ and $\psi\colon [n] \to [n]$, we define $\psi\varphi\colon [m+n] \to [m+n]$ by $(\psi\varphi)(i) = \varphi(i)$ for $i \leq m$ and $(\psi\varphi)(i) = m + \psi(i-m)$ for $i > m$.

It should be clear to the reader that $p \equiv q$ iff there exists a 2-cell $\rho\colon p \Rightarrow q$.

We now present a generalisation of the notion of RPO to G-categories.

Definition 8. (GRPO) Let $\mathbb{C}$ be a G-category. A candidate (GRPO) for square *(i)* below



(i) (ii) (iii) (iv)

is a tuple $\langle R, e, f, g, \beta, \gamma, \delta\rangle$ such that $\delta b \bullet g\beta \bullet \gamma a = \rho$ (cf. diagram (ii)).

A G-relative-pushout (GRPO) for *(i)* is a candidate which satisfies a universal property, namely, for any other candidate $\langle R', e', f', g', \beta', \gamma', \delta'\rangle$ there exists a quadruple $\langle h, \varphi, \psi, \tau\rangle$ where $h\colon R \to R'$, $\varphi\colon e' \Rightarrow he$ and $\psi\colon hf \Rightarrow f'$ (cf. diagram (iii)) and $\tau\colon g'h' \Rightarrow g$ (diagram (iv)) which makes the two candidates compatible in the obvious way.

Spelling this out, the equations that need to be satisfied are:

(1) $\tau e \bullet g'\varphi \bullet \gamma' = \gamma$;

(2) $\delta' \bullet g'\psi \bullet \tau^{-1} f = \delta$;

(3) $\psi b \bullet h\beta \bullet \varphi a = \beta'$.

We shall refer to such a quadruple as a *mediating morphism*. Such a morphism must be *essentially unique*, namely, for any other mediating morphism $\langle h', \varphi', \psi', \tau'\rangle$ there must exist a unique two cell $\xi\colon h \Rightarrow h'$ which makes the two mediating morphisms compatible, i.e.:

(1) $\xi e \bullet \varphi = \varphi'$;

(2) $\psi \bullet \xi^{-1} f = \psi'$;

(3) $\tau' \bullet g'\xi = \tau$.

In order to verify that GRPOs correspond to biRPOs it is enough to observe that when $\mathbb{C}$ is a G-category clause (b) in Definition 5 takes the form:

(b′) Given another arrow $h' : \mathrm{Bic}\, G \to A$ and a family of invertible 2-cells $\varphi'_i : h'\eta_i \Rightarrow \alpha_i$ which makes the two pseudo-cocones compatible, there exists a unique invertible 2-cell $\xi : h \Rightarrow h'$ such that $\varphi'_i \bullet \xi\eta_i = \varphi_i$.

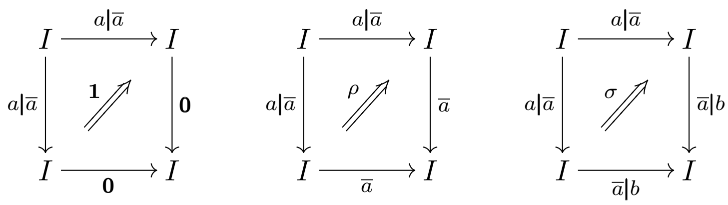
In turn this means that clause (b) of Proposition 1 can be simplified as:

(b′) For any other triple $\langle u', \varphi', \psi' \rangle$ satisfying the equations of item (a) there exists a unique $\xi\colon u' \Rightarrow u$ such that $\xi c \bullet \varphi' = \varphi$ and $\psi' \bullet \xi^{-1} d = \psi$.

Observe that whereas RPOs are defined up to isomorphism, GRPOs (due to their nature of bicolimits) are defined up to equivalence. GRPOs (and a fortiori biRPOs) clearly generalise RPOs: if one considers a category as a discrete 2-category (the only 2-cells are identities) then a GRPO is simply a RPO.

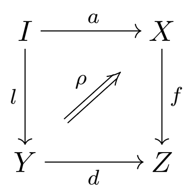
Definition 9. (GIPO) Diagram (i) of Definition 8 is said to be a G-idem-pushout (GIPO) if $\langle Z, c, d, \mathrm{id}_Z, \rho, 1_c, 1_d \rangle$ is its GRPO.

Example 3. Consider the category $\mathbf{M}_\Sigma$ from Example 2.



Consider the three squares above, where $\rho(2) = \sigma(2) = 3$ and $\rho(3) = \sigma(3) = 2$. Informally, the two copies of $\overline{a}$ are swapped in the middle and the right squares. The first two squares are GIPOs, but the third square is not. We leave the proofs to the reader.

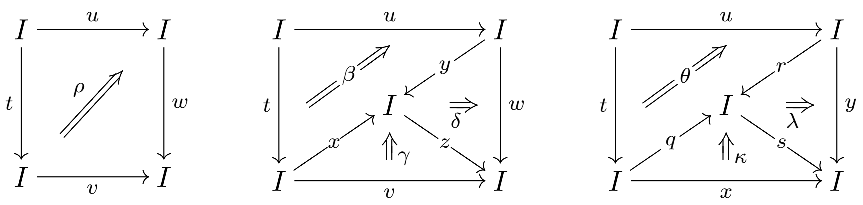
Definition 10. (Redex GRPOs) A reactive system $\mathbf{C} = \langle \mathbb{C}, \mathbb{D}, \mathcal{R}, I \rangle$, for $\mathbb{C}$ a G-category, is said to *have redex-GRPOs* if every square



(2)

where l is the left hand side of a reaction rule $\langle l, r \rangle \in \mathcal{R}$ has a GRPO.

EXAMPLE 4. The category $\mathbf{M}_\Sigma$ from Example 2 has redex GIPOs for any choice of $\mathcal{R}$. Indeed, consider any square *(i)* as below



(i) *(ii)* *(iii)*

where $t = t_1 \mid \cdots \mid t_{|t|}$, $u = u_1 \mid \cdots \mid u_{|u|}$, $v = v_1 \mid \cdots \mid v_{|v|}$ and $w = w_1 \mid \cdots \mid w_{|w|}$. We use $|-|$ to count the number of parallel components in terms. Then take $z = v_{\ell_1} \mid \cdots \mid v_{\ell_k}$ and $x = v_{\ell_{k+1}} \mid \cdots \mid v_{\ell_{|v|}}$, where $\{\ell_1, \ldots, \ell_k\} + \{\ell_{k+1}, \ldots, \ell_{|v|}\}$ is a partition of $[|v|]$ such that $\rho(\ell_j + |t|) \leq |l|$ iff $j > k$ (recall that $[n] = \{1, 2, \ldots, n\}$). Similarly, let $y = w_{m_1} \mid \cdots \mid w_{m_h}$, for $\{m_1, \ldots, m_h\}$ the subset of $[|w|]$ such that $\rho^{-1}(m_j) < |t|$. Then γ, δ and β are uniquely determined so that $\delta u \bullet z\beta \bullet \gamma t = \rho$, as illustrated in *(ii)*. Note that

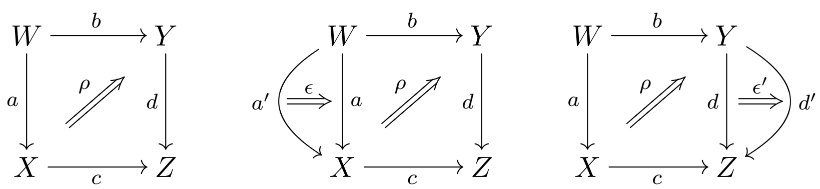
$$\beta(i + |t|) \leq |u| \text{ for } i \geq 1. \tag{3}$$

We claim that the resulting square *(iii)* is a GIPO. Indeed, suppose that $\langle q, r, s, \theta, \kappa, \lambda\rangle$ is a candidate. By (3), $s = \mathbf{0}$ and $\langle \mathbf{0}, \kappa^{-1}, \lambda^{-1}, 1_{\mathbf{0}}\rangle$ is the unique mediating morphism.

Notice that in fact *(i)* is a GIPO if and only if (3) holds.

The next lemma proves one of the fundamental properties of GRPOs.

LEMMA 2. Suppose that diagram *(i)* below is a GIPO and than $\epsilon \colon a' \Rightarrow a$, $\epsilon' \colon d \to d'$ are isomorphisms.



(i) *(ii)* *(iii)*

Then the regions obtained by pasting the 2-cells in *(ii)* and *(iii)* are GIPOs.

PROOF. Suppose that $\mathbf{R} = \langle R, e, f, g, \beta, \gamma, \delta\rangle$ is a candidate for *(ii)*. Then then tuple $\mathbf{R}' = \langle R, e, f, g, \beta \bullet e\epsilon^{-1}, \gamma, \delta\rangle$ is a candidate for *(i)* and we obtain the mediating morphism $\langle u, \phi, \psi, \tau\rangle$ between $\langle Z, c, d, \mathrm{id}_Z, \rho, 1_c, 1_d\rangle$ and $\mathbf{R}'$. It is straightforward to check that this is also a mediating morphism between $\langle Z, c, d, \mathrm{id}_Z, \rho \bullet c\epsilon, 1_c, 1_d\rangle$ and $\mathbf{R}$ and that the universal property follows from the universal property of *(i)*.

If $\langle R, e, f, g, \beta, \gamma, \delta\rangle$ is a candidate for *(iii)*, then $\langle R, e, f, g, \beta, \gamma, \epsilon'^{-1} \bullet \delta\rangle$ is a candidate for *(i)*. Hence there is a mediating morphism $\langle u, \phi, \psi, \tau\rangle$ and it is easy to check that $\langle u, \phi, \psi \bullet u\epsilon'^{-1}, \tau\rangle$ is a mediating morphism for the region. It is clear that the universal property also follows. □

Definition 11. (LTS) For $\mathbf{C}$ a reactive system whose underlying category $\mathbb{C}$ is a G-category, define $\mathbf{GTS}(\mathbf{C})$ as follows:

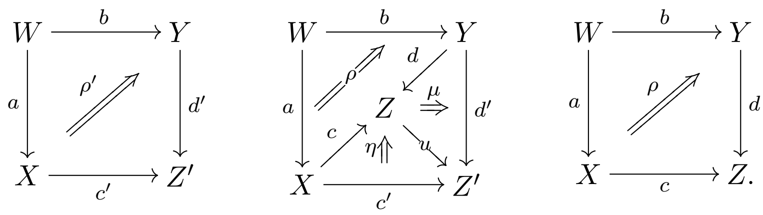
- the states $\mathbf{GTS}(\mathbf{C})$ are iso-classes of arrows $[a]\colon I \to X$ in $\mathbf{C}$
- there is a transition $[a] \xrightarrow{[f]} [a']$ if there exists a 2-cell ρ, $\langle l, r\rangle \in \mathcal{R}$ and d in $\mathbb{C}$ such that Diagram 2 is a GIPO and $a' = dr$.

Observe that the LTS is well defined by Lemma 2.

6. Congruence Theorem

In this section we formulate and prove our central Theorem 1 which states that bisimulation on the LTS derived as in Definition 11 is a congruence provided that sufficient GRPOs exist. Following a proof strategy similar to that of Leifer [2001], our proof unfolds in the Lemmas 3, 4 and 5.

Lemma 3. (GIPOs from GRPOs) *If $\langle Z, c, d, u, \rho, \eta, \mu\rangle$ is a* G*RPO for (i) below, as illustrated in (ii), then (iii) is a* G*IPO.*



(i) *(ii)* *(iii)*

Proof. Suppose that $\langle R, e, f, g, \beta, \gamma, \delta\rangle$ is a candidate for *(iii)*. Then it is easy to verify that $\langle R, e, f, ug, \beta, u\gamma \bullet \eta, \mu \bullet u\delta\rangle$ is a candidate for *(i)*.

Thus there exists $h\colon Z \to R$ and isomorphisms $\varphi\colon e \Rightarrow hc$, $\psi\colon hd \Rightarrow f$ and $\tau\colon ugh \Rightarrow u$ satisfying $\tau c \bullet ug\varphi \bullet u\gamma \bullet \eta = \eta$, $\mu \bullet u\delta \bullet ug\psi \bullet \tau^{-1}d = \mu$ and $\psi b \bullet h\rho \bullet \varphi a = \beta$ (†).

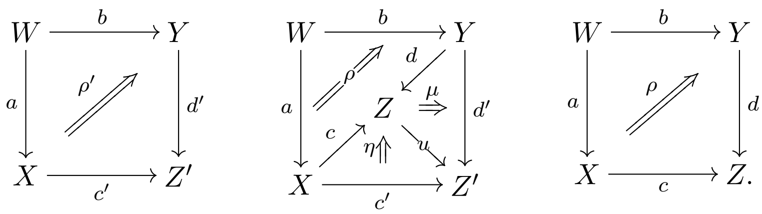
It follows that $\langle id_Z, 1_c, 1_d, 1_u\rangle$ and $\langle gh, g\varphi \bullet \gamma, \delta \bullet g\psi, \tau\rangle$ are both mediating morphisms from $\langle Z, c, d, u, \rho, \eta, \mu\rangle$ to $\langle Z, c, d, u, \rho, \eta, \mu\rangle$. Therefore there exists a unique 2-cell $\xi\colon gh \Rightarrow \mathrm{id}_Z$ such that $\xi c \bullet g\varphi \bullet \delta = 1_c$ (‡), $\delta \bullet g\psi \bullet \xi^{-1}d = 1_d$ (♠) and $u\xi = \tau$.

Equations (†), (‡) and (♠) ensure that $\langle h, \varphi, \psi, \xi\rangle$ is a mediating morphism from $\langle Z, c, d, \mathrm{id}, \rho, 1_c, 1_d\rangle$ to $\langle R, e, f, g, \beta, \delta, \gamma\rangle$ as candidates for *(iii)*.

Let $\langle h', \varphi', \psi', \xi'\rangle$ be another such mediating morphism. Then it is easy to verify that $\langle h', \varphi', \psi', u\xi'\rangle$ constitutes another mediating morphism from $\langle Z, c, d, u, \rho, \eta, \mu\rangle$ to $\langle R, e, f, ug, \beta, u\gamma \bullet \eta, \mu \bullet u\delta\rangle$. Thus there exists an unique

$\lambda\colon h \Rightarrow h'$ which satisfies $\lambda c \bullet \varphi = \varphi'$, $\psi \bullet \lambda^{-1} d = \psi'$ and $u\xi' \bullet ug\lambda = \tau (= u\xi)$. It remains to check that $\xi' \bullet g\lambda = \xi$, and this follows from the uniqueness of ξ. □

Lemma 4. (GRPOs from GIPOs) *If square (iii) below is a* G*IPO, (i) has a* G*RPO, and* $\langle Z, c, d, u, \rho, \eta, \mu\rangle$ *is a candidate for it as shown in (ii), then* $\langle Z, c, d, u, \rho, \eta, \mu\rangle$ *is a* G*RPO for (i).*



(i) (ii) (iii)

Proof. $\langle R, e, f, g, \beta, \gamma, \delta\rangle$ be an RPO for *(i)*. Using its defining property, there exists a morphism $v\colon R \to Z$ and isomorphisms $\varphi\colon c \Rightarrow ve$, $\psi\colon vf \Rightarrow d$ and $\tau\colon uv \Rightarrow g$ which satisfy $\tau e \bullet u\varphi \bullet \eta = \gamma$ $(\star)$, $\mu \bullet u\psi \bullet \tau^{-1} f = \delta$ and $\psi b \bullet v\beta \bullet \varphi a = \rho$ $(\dagger)$. The last equation asserts that $\langle R, e, f, v, \beta, \varphi, \psi\rangle$ is a candidate for the square on the right, thus there exists an arrow $w\colon Z \to R$ and isomorphisms $\varphi'\colon e \Rightarrow wc$, $\psi'\colon wd \Rightarrow f$ and $\tau'\colon vw \Rightarrow \mathrm{id}_Z$ such that $\tau' c \bullet v\varphi' \bullet \varphi = 1_c$ $(\star\star)$, $\psi \bullet v\psi' \bullet {\tau'}^{-1} d = 1_d$ and $\psi' b \bullet w\rho \bullet \varphi' a = \beta$ $(\ddagger)$.

We claim that $\left\langle wv, w\varphi \bullet \varphi', \psi' \bullet w\psi, \tau \bullet u\tau' v \bullet \tau^{-1} wv\right\rangle$ and $\langle \mathrm{id}_R, 1_e, 1_f, 1_g\rangle$ are mediating morphisms from $\langle R, e, f, g, \beta, \gamma, \delta\rangle$ to $\langle R, e, f, g, \beta, \gamma, \delta\rangle$. First, $(\dagger)$ and $(\ddagger)$ together imply that $(\psi' \bullet w\psi) b \bullet wv\beta \bullet (w\varphi \bullet \varphi') a = \beta$. It remains to show that

$$(\tau \bullet u\tau' v \bullet \tau^{-1} wv) e \bullet g(w\varphi \bullet \varphi') \bullet \gamma = \gamma$$

and

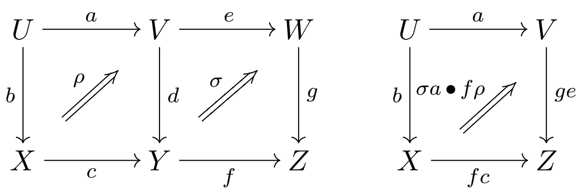
$$\delta \bullet g(\psi' \bullet w\psi) \bullet (\tau \bullet u\tau' v \bullet \tau^{-1} wv)^{-1} f = \delta.$$

We show that the first holds, the second is similar. Indeed,

$$\begin{aligned} (\tau \bullet u\tau' v \bullet \tau^{-1} wv) e \bullet g(w\varphi \bullet \varphi') \bullet \gamma &= && \text{(pasting)} \\ \tau e \bullet u\varphi \bullet u\tau' c \bullet uv\varphi' \bullet \tau^{-1} e \bullet \gamma &= && (\star) \\ \tau e \bullet u\varphi \bullet u\tau' c \bullet uv\varphi' \bullet u\varphi \bullet \eta &= && \\ \tau e \bullet u\varphi \bullet u(\tau' c \bullet v\varphi') \bullet u\varphi \bullet \eta &= && (\star\star) \\ \tau e \bullet u\varphi \bullet \eta &= \gamma. && (\star) \end{aligned}$$

Therefore, there exists a unique $\xi\colon wv \Rightarrow \mathrm{id}_R$ which makes the two mediating morphisms compatible. Since GRPOs are defined up to an equivalence, this completes the proof. □

LEMMA 5. Suppose that diagram *(ii)* below has a GRPO.



(i) (ii)

(1) If both squares in *(i)* are GIPOs then the rectangle of *(i)* is a GIPO
(2) If the left square and the rectangle of *(i)* are GIPOs then so is the right square.

PROOF. **(1).** By Lemma 4, $\langle Y, c, d, f, \rho, 1_{fc}, \sigma\rangle$ is a GRPO for *(ii)*. Suppose that $\langle R, u, v, w, \beta, \gamma, \delta\rangle$ is a candidate for the rectangle of *(i)*, that is $\delta ea \bullet w\beta \bullet \gamma b = \sigma a \bullet f\rho$.

Thus $\langle R, u, ve, \beta, \gamma, \delta e\rangle$ is a candidate for *(ii)* and so there exists an arrow $m\colon Y \to R$ and two-cells $\varphi\colon u \Rightarrow mc$, $\psi\colon md \Rightarrow ve$ and $\tau\colon wm \Rightarrow f$ satisfying the usual compatibility requirements.

In particular, $\delta e \bullet w\psi \bullet \tau^{-1}d = \sigma$, and so $\left\langle R, m, v, w, \psi, \tau^{-1}, \delta\right\rangle$ is a candidate for the right square of *(i)*. Thus there exists an arrow $n\colon Z \to R$ and two-cells $\varphi'\colon m \Rightarrow nf$, $\psi'\colon ng \Rightarrow v$ and $\tau'\colon wn \Rightarrow \mathrm{id}_Z$. The reader should verify that $\langle n, \varphi' c \bullet \varphi, \psi', \tau'\rangle$ constitutes a mediating morphism from $\langle Z, fc, g, \mathrm{id}_Z, \sigma a \bullet f\rho, 1_{fc}, 1_g\rangle$ to $\langle R, u, v, w, \beta, \gamma, \delta\rangle$.

Let $\langle n', \varphi'', \psi'', \tau''\rangle$ be another such mediating morphism. Then it follows that $\langle n'f, \varphi'', \psi'' e \bullet n'\sigma, \tau'' f\rangle$ is a mediating morphism from candidate $\langle Y, c, d, f, \rho, 1_{fc}, \sigma\rangle$ to $\langle R, u, ve, \beta, \gamma, \delta e\rangle$. Thus there exists a unique $\xi\colon m \Rightarrow n'f$ which makes it compatible with $\langle m, \varphi, \psi, \tau\rangle$, in particular, $\xi c \bullet \varphi = \varphi''$. Now $\langle n', \xi, \psi'', \tau''\rangle$ is a mediating morphism between $\langle Z, f, g, \mathrm{id}_Z, \sigma, 1_f, 1_g\rangle$ and $\left\langle R, m, v, w, \psi, \tau^{-1}, \delta\right\rangle$. Hence there exists a unique $\xi'\colon n \Rightarrow n'$ which makes the mediating morphism compatible with $\langle n, \varphi', \psi', \tau'\rangle$. It is easy to check that ξ' makes $\langle n, \varphi' c \bullet \varphi, \psi', \tau'\rangle$ compatible with $\langle n', \varphi'', \psi'', \tau''\rangle$ also. If there is another such ξ'' then by uniqueness of ξ it also makes $\langle n, \varphi', \psi', \tau'\rangle$ and $\langle n', \xi, \psi'', \tau''\rangle$ compatible, hence ξ'' must equal ξ'.

(2). Suppose that $\langle R, u, v, w, \beta, \delta, \gamma\rangle$ is a candidate for the right square of *(i)*. Then

$$\langle R,\ uc,\ v,\ w,\ \beta a \bullet u\rho,\ \gamma c,\ \delta\rangle$$

is a candidate for the rectangle and so there exists an arrow $m\colon Z \to R$ and two-cells $\varphi\colon uc \Rightarrow mfc$, $\psi\colon mg \Rightarrow v$ and $\tau\colon wm \Rightarrow \mathrm{id}_Z$ satisfying the three compatibility equations.

Recall that by Lemma 4, candidate $\langle Y, c, d, f, \rho, 1_{fc}, \sigma\rangle$ is a GRPO for *(ii)*. Now $\left\langle u, 1_{uc}, \beta, \gamma^{-1}\right\rangle$ and $\langle mf, \varphi, \psi e \bullet m\sigma, \tau f\rangle$ are mediating morphisms between

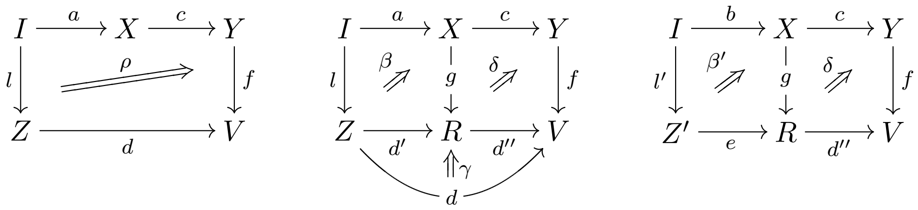
$$\langle Y, c, d, f, \rho, 1_{fc}, \sigma\rangle \quad \text{and} \quad \langle R, uc, ve, w, \beta a \bullet u\rho, \gamma, \delta\rangle\,.$$

Thus there exists a unique two-cell $\xi \colon u \Rightarrow mf$ making the two mediating morphisms compatible. In particular , $\xi c = \varphi$ which implies that $\langle m, \xi, \psi, \tau \rangle$ is a mediating morphism from $\langle Z, f, g, \mathrm{id}_Z, \sigma, 1_f, 1_g \rangle$ to $\langle R, u, v, \beta, \gamma, \delta \rangle$ in the right square.

If $\langle m', \varphi', \psi', \tau' \rangle$ is another such mediating morphism then $\langle m', \varphi' c, \psi', \tau' \rangle$ is a mediating morphism for the rectangle. Hence there is a unique $\xi' \colon m \Rightarrow m'$ which makes this mediating morphism compatible with $\langle m, \varphi, \psi, \tau \rangle$. The universal property of the left square implies that ξ' also makes $\langle m, \xi, \psi, \tau \rangle$ compatible with $\langle m', \varphi', \psi', \tau' \rangle$. Uniqueness follows from the universal property of the rectangle. □

THEOREM 1. *Let* $\mathbf{C}$ *be a reactive system whose underlying* G*-category* $\mathbb{C}$ *has redex* G*RPOs. The largest bisimulation* $\sim$ *on* $\mathbf{GTS}(\mathbf{C})$ *is a congruence.*

PROOF. It suffices to show that $S = \{ ([ca], [cb]) \mid [a] \sim [b] \}$ is a bisimulation. Suppose that $[a] \sim [b]$ and $[ca] \xrightarrow{[f]}\!\!\triangleright [a']$. Then



(*i*) (*ii*) (*iii*)

there exists $\langle l, r \rangle \in \mathcal{R}$, $d \colon Z \to V$ and $\rho \colon dl \Rightarrow fca$ such that (i) is a GIPO and $[a'] = [dr]$. Since $\mathbb{C}$ has redex-GRPOs, there exists $\langle R, d', g, d'', \beta, \gamma, \delta \rangle$ as shown in (ii) which is a GRPO. By Lemma 3, the left square in (ii) is a GIPO. Thus $[a] \xrightarrow{[g]}\!\!\triangleright [d'r]$ and so $[b] \xrightarrow{[g]}\!\!\triangleright [b']$ where $[b'] \sim [d'r]$. By definition, there is a pair $\langle l', r' \rangle \in \mathcal{R}$, an arrow $e \colon Z' \to R$ and a two-cell $\beta' \colon el' \Rightarrow gb$ so that the left square of (iii) is a GIPO and $[b'] = [er']$.

Now Lemma 2 implies that the composite of the two squares in (ii) is a GIPO and therefore, by part 2 of Lemma 5 the right square is a GIPO. Since we have deduced that both the squares in (iii) are GIPOs, part 1 of Lemma 5 ensures that the entire region is a GIPO and that $[cb] \xrightarrow{[f]}\!\!\triangleright [d''er']$. Since $[d'r] \sim [er']$, we conclude that $([d''d'r], [d''er']) \in S$. □

7. Comparison with Colouring

Sewell proposed an elegant derivation of LTS for ground term rewriting systems on syntax containing $\{\mathsf{|}, \mathbf{0}\}$ where terms are viewed modulo the standard structural congruence rules. The derivation procedure uses the notion of *colouring* [Sewell 1998]. We shall briefly recall the details and compare the LTS derived with the one derived using the theory of GRPO for the simple calculus of Example 2 (see Example 3 for sample labels). The reader should note that Sewell considers arbitrary signatures Σ with $\{|, \mathbf{0}\}$ and the relevant structural congruence; here we only consider signatures Σ with $\{|, \mathbf{0}\}$ and constants. The GIPO approach can be extended to arbitrary signatures by adopting a suitable 2-categorical extension of linear Lawvere theories (cf. Section 2). Such structures are called Lawvere 2-theories and have been used, e.g., by Meseguer [1990] to provide presentation-independent realisations of rewrite theories. We plan to pursue this direction in future work.

Let $\{|, \mathbf{0}\} \subseteq \Sigma$. Let $C = \{red, blue\}$ be a set of colours and let Σ^C denote the coloured signature, it consists of $\{|, \mathbf{0}\}$ and coloured symbols σ^c, $c \in \mathbb{C}$, $\sigma \notin \{|, \mathbf{0}\}$.

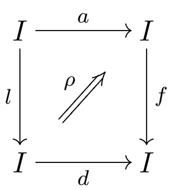
Let $\mathbf{M}_{\Sigma^C}$ and $\mathbf{M}_\Sigma$ denote the categories constructed as in Example 2. There is an obvious "underlying symbol" 2-functor $|-| \colon \mathbf{M}_{\Sigma^C} \to \mathbf{M}_\Sigma$. There are also 2-functors $(-)^{red} \colon \mathbf{M}_\Sigma \to \mathbf{M}_{\Sigma^C}$ and $(-)^{blue} \colon \mathbf{M}_\Sigma \to \mathbf{M}_{\Sigma^C}$ which colour non $\{|, \mathbf{0}\}$ symbols red and blue respectively.

DEFINITION 12. Define a labelled transition system $\mathbf{Sew}^c(\mathbf{M}_\Sigma)$ as follows

- the states are elements $a \in \mathbf{M}_\Sigma$;
- $s \xrightarrow{f} t$ iff there exists $\langle l, r\rangle \in \mathcal{R}$, $f^{red}\mathbf{s} \equiv d^{blue} l^{red}$, $|\mathbf{s}| = s$ and $t \equiv dr$

Intuitively, f contains only information necessary for the reaction.

THEOREM 2. $\mathbf{GTS}(\mathbf{M}_\Sigma) = \mathbf{Sew}^c(\mathbf{M}_\Sigma)$. PROOF. It suffices to show that there exists a 2-cell ρ such that the diagram below is a GIPO if and only if there exists a colouring $\mathbf{a}$ of a such that $f^{red}\mathbf{a} = d^{blue} l^{red}$.



Recall from Example 4 that such a square is a GIPO if and only if $\rho(i + |l|) \leq |a|$ for $0 < i \leq |d|$.

Suppose that the square is a GIPO. Assume that a, l, f and d are coloured red. Certainly $f^{red} a^{red} \equiv d^{red} l^{red}$, as exhibited by ρ. We show that d can be coloured blue while not changing the colour of f. Indeed supposing that $d = d_1 \mid \ldots \mid d_k \mid \ldots \mid d_{|d|}$, we have $\rho(k + |l|) \leq |a|$ and we can colour d_k and its image under ρ blue as the image lies in a.

Now assume that $f^{red}\mathbf{a} \equiv d^{blue} l^{red}$ and let ρ be a two-cell which exhibits this equivalence. Suppose that $\langle q, r, s, \theta, \kappa, \lambda\rangle$ is a candidate. Then s consists of elements which are both in f and in d. Since f and d are monochrome and differ in colour, $s = \mathbf{0}$. This implies that $\rho(i + |l|) \leq |a|$ for $0 < i \leq |d|$ and therefore the square is a GIPO. $\square$

8. Conclusion and Future Work

We have presented the theory of G-relative-pushouts, a generalisation of Leifer and Milner's relative-pushouts to 2-categories and, in a particular, to locally groupoidal 2-categories (G-categories). The theory allows derivation of labelled transition systems which are automatically congruences under certain general conditions. The novelty of the approach is that, by keeping track of the application of structural congruence rules (as 2-cells), we are able to derive more informative labelled transition systems in comparisons to approaches which forget the placement of reaction. We have demonstrated an application to a simple calculus with an associative and commutative parallel operator. Work is underway to apply the theory in the presence of complex structural congruences, in particular replication. We hope that eventually this research will lead to a uniform treatment of an interesting class of process calculi. We envisage that such an approach may be based on suitable Lawvere 2-theories of calculi, as mentioned briefly in §7.

Moving away from syntax based reactive systems, our 2-categorical approach could prove useful when syntactic terms are replaced by algebraic objects, such as graphs, action graphs [Milner 1996] or bigraphs [Milner 2001]. In such cases the 2-cells would be suitable structure preserving isomorphisms.

A simple example is the category of bunches, as considered by Leifer and Milner [2000]. By taking the 2-cells as permutations of the leaves, one can specify bunches elegantly, leaving out the so-called "trailing" data. We proved in [Sassone and Sobociński 2003] that GRPOs give the same LTS on such simpler bunches as RPOs do on the original definition.

The synthesis of LTS for action graphs and bigraphs relies on the *functorial reactive systems*, introduced by Leifer [2001]. They feature a category "above" related to the category of interest via a functor and decorated with trailing information so as to guarantees enough RPOs. Labels are derived accordingly and the LTS enjoys the expected congruence properties, under suitable conditions on the functor. Categories "above" and the corresponding functors can usually be generated automatically from so-called *precategories*. We recently found, however, that GRPOs subsume the use of precategories. Indeed, in [Sassone and Sobociński 2003] we present a general procedure which constructs a 2-category from any given precategory so that the LTS derived using the theory of GRPOs coincides with the one generated using the theory of functorial reactive systems.

The notion of GRPO seems also natural in the context of graph transformation systems (GTS) [Corradini *et al.* 1997] realised as graph cospans, similarly to Gadducci and Heckel [1997]. Applying the theory of GRPO in this setting will provide, we hope, interesting new LTS-based semantic theories for GTS.

Acknowledgement. The authors would like to thank Peter Sewell for generously providing access to his unpublished working notes and Steve

Lack for pointing out useful references and for many helpful comments on categorical aspects of this work. Thanks also goes to James Leifer and to the referees for their helpful suggestions.

References

BÉNABOU, J. 1967. Introduction to Bicategories. In *Midwest Category Seminar*, Volume 42 of *Lecture Notes in Mathematics*. Springer-Verlag, 1–77.

BERRY, G. AND BOUDOL, G. 1992. The chemical abstract machine. *Theoretical Computer Science 96*, 217–248.

CORRADINI, A., EHRIG, H., HECKEL, R., KORFF, M., LOWE, M., RIBEIRO, L., AND WAGNER, A. 1997. Algebraic Approaches to Graph Transformation - Part I: Single Pushout Approach and Comparison with Double Pushout Approach. In *Handbook of Graph Grammars and Computing by Graph Transformation. Vol. I: Foundations*, Rozenberg, G., Editor. World Scientific, 247–312.

GADDUCCI, F. AND HECKEL, R. 1997. An inductive view of graph transformation. In *Workshop on Algebraic Development Techniques*, 223–237.

HONDA, K. AND YOSHIDA, N. 1995. On reduction-based process semantics. *Theoretical Computer Science 151*, 2, 437–486.

KELLY, G. M. 1982. *Basic Concepts of Enriched Category Theory*. Volume 64 of *London Mathematical Society Lecture Note Series*. Cambridge University Press.

KELLY, G. M. 1989. Elementary Observations on 2-categorical Limits. *Bull. Austral. Math. Soc. 39*, 301–317.

KELLY, G. M. AND STREET, R. H. 1974. Review of the Elements of 2-categories. *Lecture Notes in Mathematics 420*, 75–103.

LAWVERE, F. W. 1963. Functorial semantics of algebraic theories. *Proceedings, National Academy of Sciences 50*, 869–873.

LEIFER, J. 2001. *Operational congruences for reactive systems*. Phd thesis, University of Cambridge.

LEIFER, J. 2002. Synthesising labelled transitions and operational congruences in reactive systems, part 2. Tech. Report RR-4395, INRIA Rocquencourt.

LEIFER, J. AND MILNER, R. 2000. Deriving Bisimulation Congruences for Reactive Systems. In *International Conference on Concurrency Theory*, 243–258.

MESEGUER, J. 1990. Conditional Rewriting Logic: Deduction, Models and Concurrency. Tech. Report CSL-93-02-R, SRI International.

MILNER, R. 1989. *Communication and Concurrency*. Prentice Hall.

MILNER, R. 1996. Calculi for Interaction. *Acta Informatica 33*, 8, 707–737.

MILNER, R. 2001. Bigraphical Reactive Systems: Basic Theory. Tech. Report 523, Computer Laboratory, University of Cambridge.

MILNER, R., PARROW, J., AND WALKER, D. 1992. A Calculus of Mobile Processes, (Parts I and II). *Information and Computation 100*, 1–77.

MILNER, R. AND SANGIORGI, D. 1992. Barbed Bisimulation. In *9th Colloquium on Automata, Languages and Programming, ICALP92*, Volume 623 of *Lecture Notes in Computer Science*. Springer Verlag, 685–695.

PARK, D.M. 1981. Concurrency on Automata and Infinite Sequences. In *Conf. on Theoretical Computer Science*, Volume 104 of *Lecture Notes in Computer Science*. Springer Verlag.

SASSONE, V. AND SOBOCIŃSKI, P. 2002. Deriving Bisimulation Congruences: A 2-categorical Approach. In *Proceedings of Express 2002*, Number 68(2) in ENTCS. Elsevier.

SASSONE, V. AND SOBOCIŃSKI, P. 2003. Deriving Bisimulation Congruences: 2-categories vs precategories. In *To appear in the proceedings of FOSSACS 2003*, LNCS.

SEWELL, P. 1998. From Rewrite Rules to Bisimulation Congruences. *Lecture Notes in Computer Science 1466*, 269–284.

SEWELL, P. 2000. Working note PS12.

Street, R. H. 1980. Fibrations in Bicategories. *Cahiers de topologie et géométrie différentielle XXI-2*, 111–159.