Reversible structures

Luca Cardelli
Microsoft Research, Cambridge
luca@microsoft.com

Cosimo Laneve
Università di Bologna
laneve@cs.unibo.it

ABSTRACT

Reversible structures are computational units that may progress forward and backward. We study weak coherent structures that are primarily inspired by DNA circuits and may be compiled in these systems and demonstrate a standardization theorem. When units have unique id, the standardization theorem may be strengthened in a form that bears a quadratic algorithm for reachability, a problem that is EXPSPACE-complete for generic structures. We then define a compilation of a concurrent calculus – the asynchronous RCCS – to DNA *via* reversible structures, thus yielding a fine-grain implementation of memories of the past into chemistry.

1. INTRODUCTION

In abstract computation systems, such as automata, lambda calculus, process calculi, etc., we usually model the forward progress of computations through a sequence of irreversible steps. But physical implementations of these steps are usually reversible: in physics and chemistry operations are reversible, and only an appropriate injection of energy and entropy can move the computational system in a desired direction. It is therefore relevant to discuss the implementation of a simple computational calculus into a chemical system, reflecting the reversibility of the chemical system into the calculus instead of abstracting it.

In general, since process calculi are not confluent and processes are non-deterministic, reversing a (forward) computation history means undoing the history not in a deterministic way but in a causally consistent fashion, where states that are reached during a backward computation are states that could have been reached during the computation history by just performing independent actions in a different order. In RCCS [7], Danos and Krivine achieve this with CCS without recursion by attaching a memory m to each process P, in the monitored process construct $m : P$. Memories in RCCS are stacks of information needed for processes to backtrack.

Chemical systems, however, are naturally reversible without retaining any backtracking memory. Reversibility there means reversibility of configurations, while time of course keeps marching forward. The only way to make such a system exactly reversible is to remember the position and momentum of each molecule, which is precisely contrary to the well-mixing assumption of chemical soups, namely that the probability of collision between two molecules is independent of their position [9]. In order to comply with the chemical well-mixing assumption, notions of causality and independence of events need to be adapted to reflect the fundamental fact that different molecules of the same chemical species are indistinguishable. Their interactions can cause effects, but not to the point of being able to identify the precise molecule that caused an effect.

In this paper we study the formal interplay between causal dependency and a computational system where terms bear multiplicities, which are a way of expressing the presence of different molecules of the same species – the *reversible structures*. Following Lévy [11], we define an equivalence on computations that abstracts away from the order of causally independent reductions – the *permutation equivalence*. Because of multiplicities this abstraction does not always exchange independent reductions. For example, two reductions that use a same signal cannot be exchanged because one cannot grasp whether the two reductions are competing on a same signal or are using two different occurrences of a same signal. Notwithstanding this inadequacy, permutation equivalence in reversible structures yields a standardization theorem that allows one to remove converse reductions from computations. To our knowledge, the study of causality in a language with multiplicities is original (similar studies have been carried out in models such as Petri nets [8]).

We then provide a scheme for the implementation of significant computational primitives – the *weak coherent* reversible structures – in DNA chemical systems. As discussed in [4], this latter systems can be precisely and programmably orchestrated in order to model CCS-style interaction and (massive) concurrency and to naturally model well-mixed chemical solutions by structural congruence [2]. It turns out that DNA systems may achieve irreversible computations, but they cannot avoid using reversible steps to do it (for example, for binary operators), and hence they are a natural implementation target for reversible calculi.

We finally study *coherent* reversible structures where multiplicities are dropped (terms have multiplicity one). Coherence in this strong sense is not realizable in well-mixed chemical solutions, but may become realizable in the future if we learn how to control individual molecules. We demon-

strate that the reachability problem in these structures has a computational complexity that is quadratic with respect to the size of the structures, a problem that is EXPSPACE-complete in weak coherent structures. We also measure the expressive power of coherent reversible structures by drawing a precise comparison with a sub-calculus of RCCS [7], its *asynchronous* fragment.

A discussion of the integration of irreversible operators in our model completes the work.

Related work. The studies about reversibility in calculi date back at least to the seventies when Bennett theorized reversible Turing machines that compute by dissipating less energy than irreversible ones [1]. Already Bennett's machines use histories for backtracking computations that are deterministic in that case.

More recently, areas such as bio-systems and quantum computing have stimulated foundational studies of reversible and distributed computations. For this reason, several reversible process calculi have been developed. In [7], Danos and Krivine define a reversible concurrent calculus – RCCS – and undertake a thorough algebraic study of reversibility. In RCCS the histories are recorded in memories that need a complex ad-hoc management. In particular, the congruence rule of distribution of memories in parallel contexts requires a global synchronization in the backward direction. Using a similar technique, [10] studies reversibility in the context of higher order concurrent languages and demonstrate that reversibility does not augment the expressive power of the language.

A general technique for reversing process calculi without using memories is proposed in [15]. As in our structures, in this technique, the structure of processes is not destroyed and the progress is noted by underlining the actions that have been performed (while we use the symbol $\hat{}$). Unlike our structures, the technique, in order to tag the communicating processes, generates ids on-the-fly during the communications. In reversible structures the ids are stored in outputs and we statically enforce their unicity (coherence). As for RCCS, when the computation must be reverted in a distributed setting, this technique requires a global synchronization between parallel processes that have been spawned at the same time.

The authors of the above papers have all noticed that reversing a computation history means undoing the history not in a deterministic way but in a way that is consistent with causal dependency. This is discussed in some detail in [14].

Structure of the paper. In Section 2 we define reversible structures and study the encoding in DNA circuits. In Section 3 we study weak coherent reversible structures and the theory of permutation equivalence. In Section 4 we analyze coherent reversible structures and Section 5 is devoted their relationship with asynchronous RCCS. In Section 6 we discuss the extension of our model with irreversible operators. We conclude in Section 7 by outlining some future work.

2. THE ALGEBRA OF REVERSIBLE STRUCTURES

The syntax of reversible structures uses five disjoint infinite sets: *names* $\mathcal{N}$, ranged over by $a, b, c, \cdots$, *co-names* $\overline{\mathcal{N}}$, ranged over by $\overline{a}, \overline{b}, \overline{c}, \cdots$, and a countable set of *id*s, ranged over u, v, w, $\cdots$. Names and co-names are ranged over by $\alpha, \alpha', \cdots$ and $\overline{\overline{\alpha}} = \alpha$. Names and ids are ranged over x, x', $\cdots$. The following notations for *sequences of actions* will be taken:

- sequences of $\mathcal{N}$ are ranged over by $\mathtt{A}$, $\mathtt{B}$, $\cdots$;
- sequences of elements $u{:}\overline{a}$ are ranged over by $\overline{\mathtt{A}}$, $\overline{\mathtt{B}}$, $\cdots$;
- sequences of elements $u{:}a$ are ranged over by $\mathtt{A}^{\perp}$, $\mathtt{B}^{\perp}$, $\cdots$;

Sequences of ids are ranged over by $\widetilde{u}$, $\widetilde{v}$, $\cdots$. The dots in sequences of ids are always omitted, that is $u.v.w$ is shortened into uvw, and the empty sequence is represented by ε. The length of a sequence is given by the function $\mathit{length}(\cdot)$.

The syntax of *reversible structures* includes *gates* g and *structures* $\mathtt{S}$ and consists of the rules:

$$
\begin{array}{rcll}
g & ::= & \mathtt{A}^{\perp}\,\mathtt{.}\,\hat{}\,\mathtt{B.}\overline{\mathtt{C}} & (\mathit{length}(\mathtt{A}^{\perp}\,\mathtt{.}\,\mathtt{B}) > 0) \\
 & | & \mathtt{A}^{\perp}\,\mathtt{.}\,\overline{\mathtt{B}}\,\mathtt{.}\,\hat{}\,\overline{\mathtt{C}} & (\mathit{length}(\mathtt{A}^{\perp}) > 0)
\end{array}
$$

$$
\begin{array}{rcll}
\mathtt{S} & ::= & & \\
 & & \mathbf{0} & \text{(null)} \\
 & | & u{:}\overline{a} & \text{(signal)} \\
 & | & g & \text{(gate)} \\
 & | & \mathtt{S} \mid \mathtt{S} & \text{(parallel)} \\
 & | & (\mathtt{new}\ x)\ \mathtt{S} & \text{(new)}
\end{array}
$$

A gate is a term that accepts input signals $u{:}\overline{a}$ and emits output signals, reversibly. The form $\mathtt{A}^{\perp}\,\mathtt{.}\,\hat{}\,\mathtt{B.}\overline{\mathtt{C}}$ represents input-accepting gates, at least when not considering reverse reactions. $\mathtt{A}^{\perp}$ are the inputs that have been processed, $\mathtt{B}$ are the inputs still to be processed, and $\overline{\mathtt{C}}$ are the outputs to be emitted. The other form $\mathtt{A}^{\perp}\,\mathtt{.}\,\overline{\mathtt{B}}\,\mathtt{.}\,\hat{}\,\overline{\mathtt{C}}$ represents an output-producing gate (when not considering reverse reactions). The $\mathtt{A}^{\perp}$ is as before, $\overline{\mathtt{B}}$ are the outputs that have been emitted, and $\overline{\mathtt{C}}$ are the outputs still to be emitted. Since all the inputs in a gate have to be processed before the outputs are produced, we do not need to consider other forms. In both forms, the symbol $\hat{}$ indicates the next operations (one forward and one backward) that the gate can perform. A structure may be either a void structure $\mathbf{0}$, or a signal $u{:}\overline{a}$ denoting an elementary message a with an id u, or a gate g, or a parallel composition " | " that collects gates and signals and allow them to interact. A structure may also be $(\mathtt{new}\ x)\ \mathtt{S}$ that limits the scope of a name or id x to $\mathtt{S}$; x is said to be *bound* in $(\mathtt{new}\ x)\ \mathtt{S}$. This is the only binding operator in reversible structures.

For example, a *transducer gate* transforming a signal from a name a to b is defined by $\hat{}a\,\mathtt{.}\,u{:}\overline{b}$. This gate may evolve into $v{:}a\,\mathtt{.}\,\hat{}u{:}\overline{b}$ by inputting a signal $v{:}\overline{a}$. At this stage it may emit the signal $u{:}\overline{b}$, thus becoming $v{:}a\,\mathtt{.}\,u{:}\overline{b}\hat{}$ or may backtrack to $\hat{}a\,\mathtt{.}\,u{:}\overline{b}$ by releasing the signal $v{:}\overline{a}$ (see the following semantics). Another example is a *sink gate*, such as $\hat{}a\,\mathtt{.}\,b$, that collects signals (and, in a stochastic model, may hold them for a while). This gate may evolve into $u{:}a\,\mathtt{.}\,\hat{}b$, and then may become $u{:}a\,\mathtt{.}\,v{:}b\hat{}$.

We often abbreviate the parallel of $\mathtt{S}_i$ for $i \in I$, where I is a finite set, with $\prod_{i\in I} \mathtt{S}_i$. We write $(\mathtt{new}\ x_1, \cdots, x_n)\ \mathtt{S}$ for $(\mathtt{new}\ x_1)\ \cdots (\mathtt{new}\ x_n)\ \mathtt{S}$, $n \geq 0$, and sometimes we shorten $x_1, \cdots, x_n$ into $\widetilde{x}$. The *free names and ids* in $\mathtt{S}$, denoted $\mathsf{fn}(\mathtt{S})$, are the names and ids in $\mathtt{S}$ with a non-bound occurrence.

Structures we will never want to distinguish for any semantic reason are identified by a congruence. Let $\equiv$, called

structural congruence, be the least congruence between structures containing alpha equivalence and satisfying the abelian monoid laws for parallel (associativity, commutativity and $\mathbf{0}$ as identity), and the scope laws

$$(\texttt{new}\ x)\ \mathbf{0} \equiv \mathbf{0} \qquad (\texttt{new}\ x)\ (\texttt{new}\ x')\ \texttt{S} \equiv (\texttt{new}\ x')\ (\texttt{new}\ x)\ \texttt{S},$$

$$\texttt{S} \mid (\texttt{new}\ x)\ \texttt{S}' \equiv (\texttt{new}\ x)\ (\texttt{S} \mid \texttt{S}'), \quad \textit{if } x \notin \mathsf{fn}(\texttt{S})$$

It is easy to demonstrate the following property.

PROPOSITION 2.1. *For every* $\texttt{S}$, $\texttt{S} \equiv (\texttt{new}\ \widetilde{x})\ (\prod_{i\in I} g_i \mid \prod_{j\in J} u_j{:}\overline{a_j})$. *The structure* $(\texttt{new}\ \widetilde{x})\ (\prod_{i\in I} g_i \mid \prod_{j\in J} u_j{:}\overline{a_j})$, *which is unique up-to the order of names and ids in the sequence* $\widetilde{x}$ *and the order of gates and signals, is called the* normal form *of* $\texttt{S}$.

The semantics of reversible structures is defined operationally by means of a reduction relation.

DEFINITION 2.2. *The reduction relation of reversible structures is the least relation* $\longrightarrow$ *satisfying the axioms*

(*input capture*) $u{:}\overline{a} \mid \texttt{A}^{\perp}\texttt{.}\hat{}a\texttt{.B.}\overline{\texttt{C}} \longrightarrow \texttt{A}^{\perp}\texttt{.}u{:}a\texttt{.}\hat{}\texttt{B.}\overline{\texttt{C}}$,

(*input release*) $\texttt{A}^{\perp}\texttt{.}u{:}a\texttt{.}\hat{}\texttt{B.}\overline{\texttt{C}} \longrightarrow u{:}\overline{a} \mid \texttt{A}^{\perp}\texttt{.}\hat{}a\texttt{.B.}\overline{\texttt{C}}$,

(*output release*) $\texttt{A}^{\perp}\texttt{.}\overline{\texttt{B}}\texttt{.}\hat{}u{:}\overline{a}\texttt{.}\overline{\texttt{C}} \longrightarrow u{:}\overline{a} \mid \texttt{A}^{\perp}\texttt{.}\overline{\texttt{B}}\texttt{.}u{:}\overline{a}\texttt{.}\hat{}\overline{\texttt{C}}$,

(*output capture*) $u{:}\overline{a} \mid \texttt{A}^{\perp}\texttt{.}\overline{\texttt{B}}\texttt{.}u{:}\overline{a}\texttt{.}\hat{}\overline{\texttt{C}} \longrightarrow \texttt{A}^{\perp}\texttt{.}\overline{\texttt{B}}\texttt{.}\hat{}u{:}\overline{a}\texttt{.}\overline{\texttt{C}}$,

and closed under the rules

$$\frac{\texttt{S} \longrightarrow \texttt{S}'}{(\texttt{new}\ a)\ \texttt{S} \longrightarrow (\texttt{new}\ a)\ \texttt{S}'} \qquad \frac{\texttt{S} \longrightarrow \texttt{S}'}{\texttt{S} \mid \texttt{S}'' \longrightarrow \texttt{S}' \mid \texttt{S}''}$$

$$\frac{\texttt{S}_1 \equiv \texttt{S}_1' \quad \texttt{S}_1' \longrightarrow \texttt{S}_2' \quad \texttt{S}_2' \equiv \texttt{S}_2}{\texttt{S}_1 \longrightarrow \texttt{S}_2}$$

Sequences of reductions, called computations*, are noted* $\longrightarrow^*$.

The reductions (*input capture*) and (*output release*) are called *forward reductions*, the reductions (*input release*) and (*output capture*) are called *backward reductions*.

We explain the axioms of reversible structures semantics by discussing the reductions of the transducer $\hat{}a\texttt{.}u{:}\overline{b}$ when exposed to signals $v{:}\overline{a}$ and $w{:}\overline{a}$. The transducer may behave either as $v{:}\overline{a} \mid w{:}\overline{a} \mid \hat{}a\texttt{.}u{:}\overline{b} \longrightarrow w{:}\overline{a} \mid v{:}a\texttt{.}\hat{}u{:}\overline{b}$ or as $v{:}\overline{a} \mid w{:}\overline{a} \mid \hat{}a\texttt{.}u{:}\overline{b} \longrightarrow v{:}\overline{a} \mid w{:}a\texttt{.}\hat{}u{:}\overline{b}$ according to whether the axiom (*input capture*) is instantiated either with the signal $v{:}\overline{a}$ or with $w{:}\overline{a}$ – in these cases $\texttt{A}^{\perp}$ is empty. In turn, $w{:}\overline{a} \mid v{:}a\texttt{.}\hat{}u{:}\overline{b}$ may reduce with (*output release*) as $w{:}\overline{a} \mid v{:}a\texttt{.}\hat{}u{:}\overline{b} \longrightarrow w{:}\overline{a} \mid v{:}a\texttt{.}u{:}\overline{b}\hat{} \mid u{:}\overline{b}$ or may backtrack with (*input release*) as follows $w{:}\overline{a} \mid v{:}a\texttt{.}\hat{}u{:}\overline{b} \longrightarrow v{:}\overline{a} \mid w{:}\overline{a} \mid \hat{}a\texttt{.}u{:}\overline{b}$. This backtracking is always possible in our algebra. In fact, it is a direct consequence of the property that, for every axiom $\texttt{S} \longrightarrow \texttt{S}'$ of Definition 2.2, there is a "converse one" $\texttt{S}' \longrightarrow \texttt{S}$.

PROPOSITION 2.3. *For any reduction* $\texttt{S} \longrightarrow \texttt{S}'$ *there exists a converse one* $\texttt{S}' \longrightarrow \texttt{S}$.

We notice that, $\hat{}a\texttt{.}u{:}\overline{b} \mid v{:}\overline{a} \mid \hat{}a\texttt{.}u{:}\overline{b} \equiv v{:}\overline{a} \mid \hat{}a\texttt{.}u{:}\overline{b} \mid \hat{}a\texttt{.}u{:}\overline{b}$ (and similarly for every permutation of gates and signals). In these structures, the two occurrences of $\hat{}a\texttt{.}u{:}\overline{b}$ are indistinguishable, that is it is not possible to identify the precise gate $\hat{}a\texttt{.}u{:}\overline{b}$ that performs the reduction $\hat{}a\texttt{.}u{:}\overline{b} \mid v{:}\overline{a} \mid \hat{}a\texttt{.}u{:}\overline{b} \longrightarrow v{:}a\texttt{.}\hat{}u{:}\overline{b} \mid \hat{}a\texttt{.}u{:}\overline{b}$. This feature formalizes the well-mixing assumption of chemical solutions, namely that the probability of collision between two molecules is independent of their position. This is also the main difference between our model and reversible process calculi models as [7, 14], where every element has a unique tag. (We will study reversible structures where elements have unique tags in Section 4.) We finally notice that, as a consequence of the above identities, the notions of causality and independence of reductions need to be adapted to reflect the fundamental fact that different molecules of the same chemical species are indistinguishable.

By Proposition 2.1 and the definition of the reduction relation, it is possible to restrict the arguments about the dynamics of reversible structures to structures in normal forms. In turns, the following statement allows one to limit the analysis to the subclass of structures without $\texttt{new}$s when the interest is in computations of "closed" structures, namely structures that do not interact with the external environment. This will simplify the following notions (such as weak coherence and labels).

PROPOSITION 2.4. $(\texttt{new}\ \widetilde{x})\ (\prod_{i\in I} g_i \mid \prod_{j\in J} u_j{:}\overline{a_j}) \longrightarrow (\texttt{new}\ \widetilde{x})\ (\prod_{i\in I} g_i' \mid \prod_{j\in J'} u_j'{:}\overline{a_j'})$ *if and only if* $\prod_{i\in I} g_i \mid \prod_{j\in J} u_j{:}\overline{a_j} \longrightarrow \prod_{i\in I} g_i' \mid \prod_{j\in J'} u_j'{:}\overline{a_j'}$.

In the following, if not otherwise specified, the structures will be considered without $\texttt{new}$s.

DEFINITION 2.5. *A structure* $\texttt{S}$ *is* weak coherent *whenever, ids are uniquely associated to names and co-names. That is, if* $u{:}\alpha$ *and* $u{:}\alpha'$ *occur in* $\texttt{S}$ *then either* $\alpha = \alpha'$ *or* $\alpha = \overline{\alpha'}$.

For example, the structure $u{:}a\texttt{.}v{:}\overline{b}\hat{} \mid v{:}\overline{c}$ is not weak coherent because v is associated to two different co-names, while $u{:}a\texttt{.}v{:}\overline{b}\hat{} \mid v{:}\overline{b}$ is weak coherent. Weak coherence is an invariance of the reduction relation.

PROPOSITION 2.6. *If* $\texttt{S}$ *is weak coherent and* $\texttt{S} \longrightarrow \texttt{S}'$ *then* $\texttt{S}'$ *is weak coherent.*

The compilation into DNA circuits. Weak coherent reversible structures may be implemented into the DSD language, a formalism for defining DNA strands and gates and study the biological mechanisms for binding and unbinding of strands [13], as a variation of the irreversible structures of [4]. We conclude this section by defining the encoding of the reversible structures into the DSD terms. This part may be safely skipped by uninterested readers.

The syntax of DSD uses *domains* $\texttt{a}$, $\texttt{u}$, $\texttt{t}$, $\cdots$, and *toehold domains* $\texttt{a~}$, $\texttt{u~}$, $\texttt{t~}$, $\cdots$. DSD terms $\texttt{D}$ are similar to structures, except that gates and structures are replaced by strands and DNA gates. The strands and gates we use in the DSD encoding are in particular:

- *strands* are `<u t~ b>` or `<a t~ >` or `<t~ u>`;
- DNA *gates* are $\texttt{G}_1\texttt{:}\texttt{G}_2\texttt{:}\cdots\texttt{:}\texttt{G}_n$ where $\texttt{G}_i$ may be either `t~` or `[a t~ ]` or `[t~ a]` or `<b>[a t~ ]` or `[a t~ ]<b>`;

Given a structural congruence definition similar to the one of reversible structures, the semantics of DSD is the least relation $\longrightarrow$ containing (structural congruence and reduction will be denoted as in reversible strand algebra):

– $\mathtt{G_1}{:}\cdots{:}\mathtt{t\tilde{}}\ {:}[\mathtt{a\ t\tilde{}}\ ]{:}\cdots{:}\ \mathtt{G}_n \mid \mathtt{<u\ t\tilde{}\ a>} \longleftrightarrow \mathtt{G_1}{:}\cdots{:}\mathtt{<u>[t\tilde{}\ a]}{:}\mathtt{t\tilde{}}\ {:}\cdots{:}\ \mathtt{G}_n \mid \mathtt{<a\ t\tilde{}\ >}$

– $\mathtt{G_1}{:}\cdots{:}\mathtt{t\tilde{}}\ {:}[\mathtt{a\ t\tilde{}}\ ]\mathtt{<b>}{:}\cdots{:}\ \mathtt{G}_n \mid \mathtt{<t\tilde{}\ a>} \longleftrightarrow \mathtt{G_1}{:}\cdots{:}[\mathtt{t\tilde{}\ a}]{:}\mathtt{t\tilde{}}\ {:}\cdots{:}\ \mathtt{G}_n \mid \mathtt{<a\ t\tilde{}\ b>}$

(axioms are bidirectional, hence the symbol $\longleftrightarrow$) and closed under the same rules of reversible structures. Figure 1 illustrates strands, DNA gates and reductions of the DSD language. The encoding $(\!|\ \cdot\ |\!)$ of reversible structures to DSD terms is homomorphic with respect to parallel and new and it is defined on signals and gates as follows (for gates we only illustrate the encodings of configurations of $a_1 \,.\, a_2 \,.\, v_1{:}\overline{b_1} \,.\, v_2{:}\overline{b_2}$):

– $(\!|\ u{:}\overline{a}\ |\!) \quad = \quad \mathtt{<u\ t\tilde{}\ a>}$

– $(\!|\ \hat{}a_1 \,.\, a_2 \,.\, v_1{:}\overline{b_1} \,.\, v_2{:}\overline{b_2}\ |\!) \quad =$
$\mathtt{t\tilde{}}\ {:}[\mathtt{a}_1\ \mathtt{t\tilde{}}\ ]{:}[\mathtt{a}_2\ \mathtt{t\tilde{}}\ ]{:}[\mathtt{v}_1\ \mathtt{t\tilde{}}\ ]\mathtt{<b_1>}{:}[\mathtt{v}_2\ \mathtt{t\tilde{}}\ ]\mathtt{<b_2>}$
$\mid \mathtt{<t\tilde{}\ v_1>} \mid \mathtt{<t\tilde{}\ v_2>}$

– $(\!|\ u_1{:}a_1 \,.\, \hat{}a_2 \,.\, v_1{:}\overline{b_1} \,.\, v_2{:}\overline{b_2}\ |\!) \quad =$
$\mathtt{<u_1>[t\tilde{}\ a_1]}{:}\mathtt{t\tilde{}}\ {:}[\mathtt{a}_2\ \mathtt{t\tilde{}}\ ]{:}[\mathtt{v}_1\ \mathtt{t\tilde{}}\ ]\mathtt{<b_1>}{:}[\mathtt{v}_2\ \mathtt{t\tilde{}}\ ]\mathtt{<b_2>}$
$\mid \mathtt{<a_1\ t\tilde{}\ >} \mid \mathtt{<t\tilde{}\ v_1>} \mid \mathtt{<t\tilde{}\ v_2>}$

– $(\!|\ u_1{:}a_1 \,.\, u_2{:}a_2 \,.\, v_1{:}\overline{b_1} \,.\, \hat{}v_2{:}\overline{b_2}\ |\!) \quad =$
$\mathtt{<u_1>[t\tilde{}\ a_1]}{:}\mathtt{<u_2>[t\tilde{}\ a_2]}{:}[\mathtt{t\tilde{}\ v_1}]{:}\mathtt{t\tilde{}}\ {:}[\mathtt{v}_2\ \mathtt{t\tilde{}}\ ]\mathtt{<b_2>}$
$\mid \mathtt{<a_1\ t\tilde{}\ >} \mid \mathtt{<a_2\ t\tilde{}\ >} \mid \mathtt{<t\tilde{}\ v_2>}$

Figure 1 illustrates a sample encoding of 1 input and 1 output gate and its reactions. The strict correspondance between reversible structures and the DSD language is fixed by the following statement.

PROPOSITION 2.7. $\mathtt{S} \longrightarrow \mathtt{S}'$ *implies* $(\!|\ \mathtt{S}\ |\!) \longrightarrow (\!|\ \mathtt{S}'\ |\!)$. *Additionally, if* $\mathtt{S}$ *is weak coherent then* $(\!|\ \mathtt{S}\ |\!) \longrightarrow \mathtt{S}'$ *implies there is* $\mathtt{S}''$ *such that* $\mathtt{S}' \equiv (\!|\ \mathtt{S}''\ |\!)$ *and* $\mathtt{S} \longrightarrow \mathtt{S}''$.

The second part of Proposition 2.7 is restricted to weak coherent structures. In fact, $(\!|\ \mathtt{S}\ |\!) \longrightarrow (\!|\ \mathtt{S}'\ |\!)$ implies $\mathtt{S} \longrightarrow \mathtt{S}'$ is false in the unrestricted case. Consider the encoding of a gate $u{:}a \,.\, v{:}\overline{b}\hat{}$, namely $\mathtt{<u>[t\tilde{}\ a]}{:}[\mathtt{t\tilde{}\ v}]{:}\mathtt{t\tilde{}}\ \mid \mathtt{<a\ t\tilde{}\ >}$, and observe that, in this DNA gate, the co-name $\overline{b}$ never appears. If the (not weak coherent) structure also contained the signal $v{:}\overline{c}$, which is compiled into $\mathtt{<v\ t\tilde{}\ c>}$, then the DNA structure might reduce to $\mathtt{<u>[t\tilde{}\ a]}{:}\mathtt{t\tilde{}}\ {:}[\mathtt{v\ t\tilde{}}\ ]\mathtt{<c>}$ $\mid \mathtt{<a\ t\tilde{}\ >}$. This last DNA structure encodes $u{:}a \,.\, \hat{}v{:}\overline{c}$ and cannot be obtained from the structure $u{:}a \,.\, v{:}\overline{b}\hat{}\ \mid\ v{:}\overline{c}$.

The correspondence between the a subset of the DSD language and reversible structures has been crucial in the design of the latter ones. However, at this point a reader may wonder whether ids are really needed in these two formalisms: is it possible to define an id-free reversible structure and an encoding in the DSD language? The answer is positive. In this case signals are encoded in two-domains strands and gates have no overhangs [5] (instead of the three-domain strands above). However, in two-domains DSD structures it is not possible to define coherence (see Section 4) and to encode (in a causally consistent way) process calculi such as asynchronous RCCS. Said in a more effective way: the three-domains DSD (sub)language has the shortest domains that correctly implement reversibility of reversible process calculi.

3. WEAK COHERENCE AND CAUSALITY

Computations of reversible structures may have a lot of forward and backward reductions that continuously do and undo stuff. For example, in the transducer of Section 2, the computation

$$v{:}\overline{a} \mid w{:}\overline{a} \mid \hat{}a \,.\, u{:}\overline{b} \longrightarrow w{:}\overline{a} \mid v{:}a \,.\, \hat{}u{:}\overline{b} \longrightarrow v{:}\overline{a} \mid w{:}\overline{a} \mid \hat{}a \,.\, u{:}\overline{b}$$

is actually equivalent to the empty one – the computation performing no reduction at all. Clearly the above two reductions may be repeated at will, still being equivalent to the empty computation. Therefore, it is meaningful to analyze whether a computation may be simplified, *i.e. shortened*, without altering its computational meaning. In general, these simplifications may require swapping of independent reductions. For example, in

$$\begin{aligned} v{:}\overline{a} \mid w{:}\overline{a} \mid \hat{}a \,.\, u{:}\overline{b} \mid \hat{}a \,.\, z{:}\overline{c} & \\ \longrightarrow\ & w{:}\overline{a} \mid v{:}a \,.\, \hat{}u{:}\overline{b} \mid \hat{}a \,.\, z{:}\overline{c} \qquad (1)\\ \longrightarrow\ & v{:}a \,.\, \hat{}u{:}\overline{b} \mid w{:}a \,.\, \hat{}z{:}\overline{c} \qquad (2)\\ \longrightarrow\ & v{:}\overline{a} \mid \hat{}a \,.\, u{:}\overline{b} \mid w{:}a \,.\, \hat{}z{:}\overline{c} \qquad (3) \end{aligned}$$

the reductions (1) and (3) may be simplified because one is the reverse of the other. In order to achieve this simplification one may observe that reductions (1) and (2) involve disjoint structures – are independent, *there is no causal dependency between them* (similarly for (2) and (3)). After the swapping of (1) and (2), the reduction (1) occurs immediately before (3) and they may be removed, thus obtaining

$$v{:}\overline{a} \mid w{:}\overline{a} \mid \hat{}a \,.\, u{:}\overline{b} \mid \hat{}a \,.\, z{:}\overline{c} \longrightarrow\ v{:}\overline{a} \mid \hat{}a \,.\, u{:}\overline{b} \mid w{:}a \,.\, \hat{}z{:}\overline{c}$$

The standard equivalence in literature that identifies the above computations is *permutation equivalence* [11, 3]. We follow Lévy that uses *labels* for defining permutation equivalence.

DEFINITION 3.1. *Let* labels*, noted* μ*,* ν*,* $\cdots$*, be* input capture labels $u|\widetilde{v}\hat{}\mathtt{A}\circ\widetilde{w}$*,* input release labels $\widetilde{v}u\hat{}\mathtt{A}\circ\widetilde{w}$*,* output release labels $\widetilde{v}\circ\widetilde{w}\hat{}u\widetilde{z}$*, and* output capture labels $u|\widetilde{v}\circ\widetilde{w}u\hat{}\widetilde{z}$*. The sub-labels* $\widetilde{v}\hat{}\mathtt{A}\circ\widetilde{w}$ *and* $\widetilde{v}\circ\widetilde{w}\hat{}\widetilde{z}$*, noted* ℓ*,* ℓ'*,* $\cdots$*, are called* labels of gates.

The symbol "$\circ$" in labels separates the ids that refer to the input part of a gate from the ids that refer to the output part. Labels will be used for marking reductions (see Definition 3.2). In weak coherent structures where ids are uniquely associated to names and co-names, labels carry the minimal informations for identifying gates and signals that are reduced (up-to multiplicities). For example, if a is the name associated to the id u, then the label $u_\circ u\hat{}u$ refers to the gate $u{:}a \,.\, u{:}\overline{a} \,.\, \hat{}u{:}\overline{a}$. In fact, this gate may be reduced with an (*output release*) axiom. The label $u \mid u_\circ u\hat{}u$ addresses a signal $u{:}\overline{a}$ and a gate $u{:}a \,.\, u{:}\overline{a} \,.\, \hat{}u{:}\overline{a}$ (the same as before). In fact, $u{:}\overline{a} \mid u{:}a \,.\, u{:}\overline{a} \,.\, \hat{}u{:}\overline{a}$ may be reduced with an (*output capture*) axiom. We notice that two different labels identify the same gate $u{:}a \,.\, u{:}\overline{a} \,.\, \hat{}u{:}\overline{a}$. This is not surprising because labels mark reductions and the above gate may be actually involved in two different reductions. Finally, the two labels $u_\circ u\hat{}u$ and $uu_\circ\hat{}u$ mark te output-release reductions of the gates $u{:}a \,.\, u{:}\overline{a} \,.\, \hat{}u{:}\overline{a}$ and $u{:}a \,.\, u{:}a \,.\, \hat{}u{:}\overline{a}$, respectively. Without the symbol "$\circ$" these two reductions should have been confused.

It is worth to observe that the main difference between our labelling technique and those in [11, 3] is that labels are already available in the structures (and in the DNA, by Proposition 2.7) as ids of signals and gates.

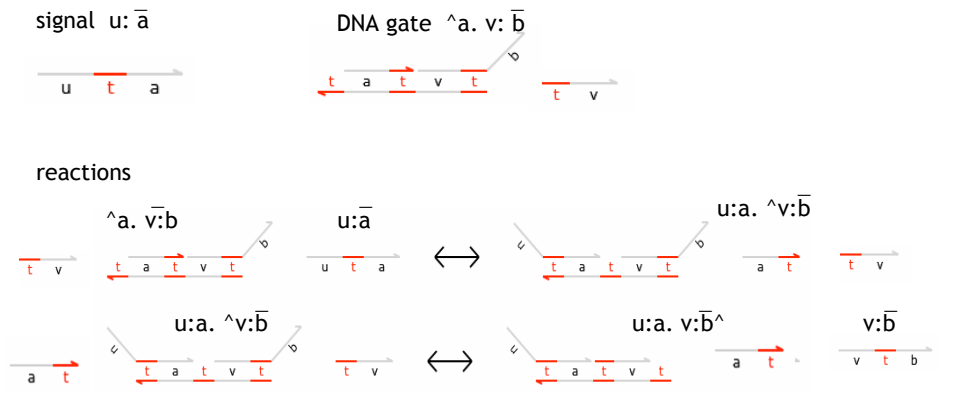


Figure 1: The dsd encoding of 1 input and 1 output gates and their reactions

DEFINITION 3.2. *Let* $id(\mathtt{A}^\perp) = \widetilde{v}$, $id(\overline{\mathtt{B}}) = \widetilde{w}$ *and* $id(\overline{\mathtt{C}}) = \widetilde{z}$; *we write* $\mu : \mathtt{S} \longrightarrow \mathtt{S}'$, *when the axiom used in the proof tree is*

- *(*input capture*)* $u{:}\overline{a} \mid \mathtt{A}^\perp \,.\, \widehat{\ }a \,.\, \mathtt{A}' \,.\, \overline{\mathtt{C}} \longrightarrow \mathtt{A}^\perp \,.\, u{:}a \,.\, \widehat{\ }\mathtt{A}' \,.\, \overline{\mathtt{C}}$ *and* $\mu = u \mid \widetilde{v}\widehat{\ }\mathtt{A}'\circ\widetilde{z}$,
- *(*input release*)* $\mathtt{A}^\perp \,.\, u{:}a \,.\, \widehat{\ }\mathtt{A}' \,.\, \overline{\mathtt{C}} \longrightarrow u{:}\overline{a} \mid \mathtt{A}^\perp \,.\, \widehat{\ }a \,.\, \mathtt{A}' \,.\, \overline{\mathtt{C}}$ *and* $\mu = \widetilde{v}u\widehat{\ }\mathtt{A}'\circ\widetilde{z}$,
- *(*output release*)* $\mathtt{A}^\perp \,.\, \overline{\mathtt{B}} \,.\, \widehat{\ }u{:}\overline{a} \,.\, \overline{\mathtt{C}} \longrightarrow u{:}\overline{a} \mid \mathtt{A}^\perp \,.\, \overline{\mathtt{B}} \,.\, u{:}\overline{a} \,.\, \widehat{\ }\overline{\mathtt{C}}$ *and* $\mu = \widetilde{v}\circ\widetilde{w}\widehat{\ }u\widetilde{z}$,
- *(*output capture*)* $u{:}\overline{a} \mid \mathtt{A}^\perp \,.\, \overline{\mathtt{B}} \,.\, u{:}\overline{a} \,.\, \widehat{\ }\overline{\mathtt{C}} \longrightarrow \mathtt{A}^\perp \,.\, \overline{\mathtt{B}} \,.\, \widehat{\ }u{:}\overline{a} \,.\, \overline{\mathtt{C}}$ *and* $\mu = u \mid \widetilde{v}\circ\widetilde{w}u\widehat{\ }\widetilde{z}$.

Letting a and b be the names associated to the ids u and v, respectively, in a weak coherent structure, both the labels $u\circ\widehat{\ }v$ and $u\widehat{\ }\circ v$ identify the same gate $u{:}a \,.\, \widehat{\ }u{:}\overline{b}$. In fact, the former label marks an (*output release*) reduction, the latter one marks an (*input release*) reduction. It is worth to notice that, if structures are not weak coherent, then labels may fail to address gates/signals that are reduced. For example, in $u{:}\overline{a} \mid u{:}\overline{b} \mid \widehat{\ }a \,.\, v{:}\overline{c} \mid \widehat{\ }a \,.\, v{:}\overline{d}$, the two reductions are labelled $u \mid \widehat{\ }a\circ v$ even if they address two different pairs of signal and gate.

Let $\mu \cap \nu \neq \emptyset$ if and only if one of the following holds (we recall that ℓ, ℓ' range over terms $\widetilde{v}\widehat{\ }\mathtt{A}\circ\widetilde{w}$ and $\widetilde{v}\circ\widetilde{w}\widehat{\ }\widetilde{z}$):

1. $\mu = u \mid \ell$ and $\nu = u \mid \ell'$;
2. $\mu = u \mid \ell$ and $\nu = v \mid \ell$;
3. $\mu = \ell$ and $\nu = \ell$;
4. $\mu = \widetilde{u}\widehat{\ }\circ\widetilde{v}$ and $\nu = \widetilde{u}\circ\widehat{\ }\widetilde{v}$;
5. $\mu = \widetilde{u}\circ\widehat{\ }\widetilde{v}$ and $\nu = \widetilde{u}\widehat{\ }\circ\widetilde{v}$.

We notice that, when $\mu \cap \nu \neq \emptyset$, the gates/signals that are reduced by μ and ν are not disjoint. We write $\mu \cap \nu = \emptyset$ when $\mu \cap \nu \neq \emptyset$ does not hold.

LEMMA 3.3. *Let* $\mu : \mathtt{S} \longrightarrow \mathtt{S}'$ *and* $\nu : \mathtt{S} \longrightarrow \mathtt{S}''$ *be such that* $\mu \cap \nu = \emptyset$. *Then there exists* $\mathtt{S}'''$ *such that* $\nu : \mathtt{S}' \longrightarrow \mathtt{S}'''$ *and* $\mu : \mathtt{S}'' \longrightarrow \mathtt{S}'''$. *The reductions* μ *and* ν *are said* causally independent.

Lemma 3.3 is known in the literature as "diamond lemma" because the two computations $\mu;\nu$ and $\nu;\mu$ have same initial and final structures – they are *coinitial* and *cofinal*. The condition $\mu \cap \nu = \emptyset$ means that gates/signals that are reduced by the two reductions are disjoint, therefore reductions are not causally related and may be swapped. Contrary to other formalisms [11, 3, 7], in (weak coherent) reversible structures, the condition $\mu \cap \nu = \emptyset$ does not completely catch reductions that may be performed concurrently. For example, in $u{:}\overline{a} \mid u{:}\overline{a} \mid \widehat{\ }a \,.\, u{:}\overline{a} \mid w{:}c \,.\, u{:}\overline{a} \,.\, \widehat{\ }v{:}\overline{b}$ we have the possibility of one input capture and one output capture of the same signal and Lemma 3.3 does not apply (even if there are two copies of the signal). The problem follows from the fact that labels do not convey details about multiplicities of signals and gates.

DEFINITION 3.4. *Let* $[\mu]^+$, *read the* converse label of μ, *be the following labels (let* a *be the name associated to* u*):*

$$\begin{array}{rcl} [u \mid \widetilde{v}\widehat{\ }a \,.\, \mathtt{A}\circ\widetilde{w}]^+ & \stackrel{\text{def}}{=} & \widetilde{v}u\widehat{\ }\mathtt{A}\circ\widetilde{w} \\ [\widetilde{v}u\widehat{\ }\mathtt{A}\circ\widetilde{w}]^+ & \stackrel{\text{def}}{=} & u \mid \widetilde{v}\widehat{\ }a \,.\, \mathtt{A}\circ\widetilde{w} \\ [\widetilde{v}\circ\widetilde{w}\widehat{\ }u\widetilde{z}]^+ & \stackrel{\text{def}}{=} & u \mid \widetilde{v}\circ\widetilde{w}u\widehat{\ }\widetilde{z} \\ [u \mid \widetilde{v}\circ\widetilde{w}u\widehat{\ }\widetilde{z}]^+ & \stackrel{\text{def}}{=} & \widetilde{v}\circ\widetilde{w}\widehat{\ }u\widetilde{z} \end{array}$$

Let $\mu_1 : \mathtt{S}_1 \longrightarrow \mathtt{S}_2, \cdots, \mu_n : \mathtt{S}_n \longrightarrow \mathtt{S}_{n+1}$. The computation $\mathtt{S}_1 \longrightarrow^* \mathtt{S}_{n+1}$ performing the reductions $\mu_1, \cdots, \mu_n$ will be denoted with $\mu_1;\cdots;\mu_n$. For example, the computation $u{:}\overline{a} \mid \widehat{\ }a \,.\, v{:}\overline{b} \longrightarrow^2 v{:}\overline{b} \mid u{:}a \,.\, v{:}\overline{b}\widehat{\ }$ is noted $u|\widehat{\ }a\circ v \;;\; u\circ\widehat{\ }v$. We observe that $\mu\,;[\mu]^+$ and $[\mu]^+\,;\mu$ do not change the initial structure (see Definition 3.5). Therefore the name given to $[\mu]^+$.

DEFINITION 3.5. Permutation equivalence, *written* $\sim$, *is the least equivalence relation between computations closed under composition and such that:*

$$\begin{array}{rcll} \mu;[\mu]^+ & \sim & \varepsilon & \\ \mu;\nu & \sim & \nu;\mu & \text{if } \mu \text{ and } \nu \text{ are coinitial and } \mu \cap \nu = \emptyset \end{array}$$

For example, the computation

$$\begin{array}{rcl} u : \overline{a} \mid \widehat{\ }a \,.\, v{:}\overline{b} \mid u{:}a \,.\, \widehat{\ }v{:}\overline{b} & \longrightarrow & u{:}a \,.\, \widehat{\ }v{:}\overline{b} \mid u{:}a \,.\, \widehat{\ }v{:}\overline{b} \\ & \longrightarrow & u{:}a \,.\, \widehat{\ }v{:}\overline{b} \mid u{:}a \,.\, v{:}\overline{b}\widehat{\ } \mid v{:}\overline{b} \\ & \longrightarrow & u{:}\overline{a} \mid \widehat{\ }a \,.\, v{:}\overline{b} \mid u{:}a \,.\, v{:}\overline{b}\widehat{\ } \mid v{:}\overline{b} \end{array}$$

that is represented by the sequence of labels $u|\widehat{\ }a\circ v \;;\; u\circ\widehat{\ }v \;;\; u\widehat{\ }\circ v$ is permutation equivalent to $u\circ\widehat{\ }v$.

Permutation equivalence as defined in Definition 3.5 is more discriminant than usual. As already discussed, the computations $u \mid \hat{}a \circ u \;;\; u \mid w \circ u\hat{}v$ and $u \mid w \circ u\hat{}v \;;\; u \mid \hat{}a \circ u$ of the structure $u{:}\overline{a} \mid u{:}\overline{a} \mid \hat{}a \,.\, u{:}\overline{a} \mid w{:}c \,.\, u{:}\overline{a} \,.\, \hat{}v{:}\overline{b}$ are not equal even if the two reductions concern different terms. The reason for this discriminating power is due to multiplicities of gates and signals and the fact that labels do not distinguish different occurrences of a same term. Of course we might have defined more informative labels recording the proof-tree of a reductions, in the style of [3], but this would have been a twist of well-mixed chemical solutions in the theory of reversible structures. In fact, in these solutions, molecules have concentrations and two occurrences of a same molecule cannot be separated. Anyhow, reversible structures without multiplicities (where labels uniquely identify the terms) and their properties are studied in the next section.

Weak coherence guarantees the soundness of Definition 3.5. The (not weak coherent) structure $u{:}\overline{a} \mid \hat{}a \,.\, v{:}\overline{b} \mid u{:}a \,.\, \hat{}v{:}\overline{c}$ has a computation

$$\begin{array}{rcl} u{:}\overline{a} \mid \hat{}a \,.\, v{:}\overline{b} \mid u{:}a \,.\, \hat{}v{:}\overline{c} & \longrightarrow & u{:}a \,.\, \hat{}v{:}\overline{b} \mid u{:}a \,.\, \hat{}v{:}\overline{c} \\ & \longrightarrow & u{:}a \,.\, \hat{}v{:}\overline{b} \mid \hat{}a \,.\, \overline{c} \mid u{:}\overline{a} \end{array}$$

whose labels are $u \mid \hat{}a \circ v \;;\; u\hat{} \circ v$, with $[u \mid \hat{}a \circ v]^+ = u\hat{} \circ v$. However, the two labels specify different gates and the above computation is not equivalent to ε.

Let the *gate of a label* be the gate, up-to structural congruence, involved in the reduction. Let $\mu : \mathtt{S} \longrightarrow \mathtt{S}'$ and let g be a gate in $\mathtt{S}$. We define g/μ, the *residual* of g after μ, the following gate in $\mathtt{S}'$:

$$g/\mu \stackrel{\text{def}}{=} \begin{cases} g & \text{if } g \text{ is not the gate of } \mu \\ g' & \text{if } g \text{ is the gate of } \mu \text{ and } g' \text{ is the gate of } [\mu]^+ \end{cases}$$

Proposition 3.6. *Let* $\mathtt{S}$ *be weak coherent and* $\mu : \mathtt{S} \longrightarrow \mathtt{S}'$. *(1) For every gate* g *in* $\mathtt{S}$ *there is a gate* g/μ *in* $\mathtt{S}'$*; (2) for every gate* g' *in* $\mathtt{S}'$ *there is a gate* g *in* $\mathtt{S}$ *such that* $g' = g/\mu$.

Theorem 3.7 (Standardization theorem). *Let* $\mathtt{S}$ *be weak coherent and* $\mu_1 \;;\; \cdots \;;\; \mu_n$ *be a computation of* $\mathtt{S}$ *such that* μ_n *is the converse of* μ_1. *Then there is a shorter computation that is permutation equivalent to* $\mu_1 \;;\; \cdots \;;\; \mu_n$.

Proof. By induction on n. The cases $n \leq 2$ are either vacuous or obvious. Let $n \geq 3$. The argument for the inductive case analyzes the sequence $\mu_1 \;;\; \cdots \;;\; \mu_n$. Since μ_1 and μ_n are one the converse of the other, let μ_1 be the reduction whose label specifies the set $\{g_1\}$, where g_1 is a gate (it does not specify any signal). The argument is by cases on μ_2. Let $\mu_1 : \mathtt{S} \longrightarrow \mathtt{S}'$ and g_2 be the gate of μ_2 in $\mathtt{S}'$. By Proposition 3.6, let g_2' in $\mathtt{S}$ be such that $g_2 = g_2'/\mu$. If $g_1 \neq g_2'$ then μ_1 and μ_2 are coinitial and $\mu_1 \cap \mu_2 = \emptyset$. Therefore $\mu_1 \;;\; \mu_2 \;;\; \cdots \;;\; \mu_n \sim \mu_2 \;;\; \mu_1 \;;\; \cdots \;;\; \mu_n$ and we may apply the inductive hypothesis to $\mu_1 \;;\; \mu_3 \;;\; \cdots \;;\; \mu_n$. If $g_1 = g_2'$ and $\mu_2 = [\mu_1]^+$ then $\mu_1 \;;\; \mu_2 \;;\; \cdots \;;\; \mu_n \sim \mu_3 \;;\; \cdots \;;\; \mu_n$ and we are done. It remains the case $g_1 = g_2'$ and $\mu_1 = \mu_2$ then μ_n is the converse of μ_2 as well. We therefore use the inductive hypothesis on $\mu_2 \;;\; \cdots \;;\; \mu_n$. □

The definitions of permutation equivalence and weak coherence imply that two permutation equivalent computations are cofinal. The converse direction is false, as witnessed by the above two computations $u \mid \hat{}a \circ u \;;\; u \mid w \circ u\hat{}v$ and $u \mid w \circ u\hat{}v \;;\; u \mid \hat{}a \circ u$. This problem, that we will amend in the next section by refining weak coherence, is well-known in the theory of Petri nets [8].

We conclude this section with a comment about the computational complexity of the reachability problem in reversible structures – the existence of a computation from one structure to another –, which is a relevant practical issue when structures represent DNA solutions. It is straightforward to encode reversible structures into symmetric (*a.k.a.* reversible) and bounded place-transition Petri nets. However, for these nets, the reachability marking problem is EXPSPACE complete [12, 6]. We are not aware of any better algorithm that improve, in our case, this limit. It is worth to remark that, even restricting our analysis to weak coherent structures, the conditions do not change very much because the problem reduces (by Theorem 3.7) to find the shortest computation in symmetric and bounded place-transition Petri net, which is the one returned by the algorithms in [12, 6]. In the next section we will study a refinement of coherence that retains better reachability algorithms.

4. COHERENT STRUCTURES

The mismatch between cofinality and permutation equivalence (of coinitial computations) may be eliminated by strengthening the notion of weak coherence. Following the remarks in Section 3, the refinement may be achieved by removing multiplicities from initial structures. We recall from the introduction that the constraint of molecules without multiplicities (with unique identity) is not realizable in well-mixed chemical solutions. The aim of this section and the next one is to study the expressive power of reversible structures and compare them with a concurrent calculus.

Let an occurrence of an id u be *positive* in a structure $\mathtt{S}$ if u occurs in a signal or in a gate $\mathtt{A}^\perp \,.\, \overline{\mathtt{B}} \,.\, \hat{}\overline{\mathtt{C}}$ or $\mathtt{A}^\perp \,.\, \hat{}\mathtt{B} \,.\, \overline{\mathtt{C}}$ in the $\mathtt{A}^\perp$ sequence or in the $\overline{\mathtt{C}}$ sequence. The occurrence of u is *negative* if it is in the $\overline{\mathtt{B}}$ sequence of a gate $\mathtt{A}^\perp \,.\, \overline{\mathtt{B}} \,.\, \hat{}\overline{\mathtt{C}}$. Let the *type of* g, written $\mathit{type}(g)$, be the sequence of ids of co-names in g. For example $\mathit{type}(v{:}a \,.\, \hat{}a \,.\, u{:}\overline{a} \,.\, w{:}\overline{c}) = uw$ (as usual, dots are omitted in sequences of ids). Let the *type of a label* be the type of the gate involved in the reduction.

Definition 4.1. *A weak coherent structure* $\mathtt{S}$ *is* coherent *whenever*

- *different gates in* $\mathtt{S}$ *have types with no id in common;*
- *ids occur at most twice: one occurrence is positive and the other is negative.*

Proposition 4.2. *If* $\mathtt{S}$ *is coherent and* $\mathtt{S} \longrightarrow \mathtt{S}'$ *then* $\mathtt{S}'$ *is coherent.*

We notice that, replacing the second part of Definition 4.1 with the simpler constraint that ids occur linearly in structures, then Proposition 4.2 should have been definitely threatened. For example, the structure $u{:}\overline{a} \mid \hat{}a \,.\, v{:}\overline{b}$ reduces to $u{:}a \,.\, v{:}\overline{b}\hat{} \mid v{:}\overline{b}$ where v occurs twice – one occurrence is positive, the other is negative: the reader may verify that this last structure matches the constraints of Definition 4.1.

Back to the mismatch between cofinality and permutation equivalence, if we weaken Definition 4.1 by admitting unconstrained occurrences of ids (the second constraint of coherence), then the problem remains. For example, the (not coherent) structure $u{:}\overline{a} \mid u{:}\overline{a} \mid \hat{}a \,.\, v{:}\overline{b} \mid \hat{}a \,.\, w{:}\overline{c}$ has two cofinal computations $u \mid \hat{}a \circ v \;;\; u \mid \hat{}a \circ w$ and $u \mid \hat{}a \circ w \;;\; u \mid \hat{}a \circ v$ that are not permutation equivalent.

THEOREM 4.3. *Let $\mu_1;\mu_2;\cdots;\mu_m$ and $\nu_1;\nu_2;\cdots;\nu_n$ be two coinitial computations of a coherent structure. Then $\mu_1;\mu_2;\cdots;\mu_m \sim \nu_1;\nu_2;\cdots;\nu_n$ if and only if they terminate in the same structure, up-to structural congruence (they are cofinal).*

PROOF. The only-if direction is immediate by definition of $\sim$ and (strong) coherence. The if-direction is demonstrated by induction on $m+n$. The base case ($m+n=0$) is immediate. Assume the theorem holds when $m+n=h$, we prove the case $m+n=h+1$.

The interesting case is when Theorem 3.7 cannot be applied to the (sub)com-putations $\mu_1;\mu_2;\cdots;\mu_m$ and $\nu_1;\nu_2;\cdots;\nu_n$. Otherwise we use the inductive hypothesis.

So assume that $\mu_1;\mu_2;\cdots;\mu_m$ and $\nu_1;\nu_2;\cdots;\nu_n$ have no pair of converse labels. By coherence, if the types of two labels are equal then the corresponding reductions have the same direction (they are both forward or backward). By cofinality, the two computations must address the same gates and every gate appears the same number of times in labels. (Therefore $m=n$.) Since the computations do not contain converse labels then, projecting out labels of a same type, we obtain either sequences of input captures followed by output releases or sequences of output captures followed by input releases – therefore subsequences are *monotone*. Additionally, the projections of $\mu_1;\mu_2;\cdots;\mu_m$ and of $\nu_1;\nu_2;\cdots;\nu_n$ corresponding to a same type must be pairwise equal.

There are two cases: (a) one of the sequences has empty subsequence of captures, (b) every sequence has nonempty subsequence of captures. In case (a), the first label of the empty subsequence of captures may be permuted till the beginning of the sequences $\mu_1;\cdots;\mu_m$ and of $\nu_1;\cdots;\nu_n$, therefore we are reduced to shorter computations (because the first two labels are equal) and we may apply inductive hypothesis. In case (b), assume μ_1 is an input capture of a signal $u{:}\overline{a}$ and let ν_h be the first occurrence in $\nu_1;\cdots;\nu_n$ of the gate in μ_1. Because of cofinality, $\mu_1=\nu_h$. We demonstrate that ν_h may be permuted with $\nu_1;\cdots;\nu_{h-1}$ and the argument is as in case (a). If no label ν_1, $\cdots$, ν_{h-1} is an input capture of $u{:}\overline{a}$ then the permutation is possible. Otherwise, take the reduction ν_i, $1\le i\le h-1$, with largest i that performs an input capture. Then ν_i is either equal to $u\,|\,\widetilde{v}\,\widehat{}\,a\,.\,\mathtt{A}_\circ\widetilde{w}$ or equal to $u\,|\,\widetilde{v}_\circ\widetilde{v'}u\,\widehat{}\,\widetilde{w}$. In both cases, after the reduction ν_i the occurrence of u is positive. Since $u{:}\overline{a}$ cannot be released by the gate of ν_i (because this would contradict the monotony) and by any other gate in ν_{i+1} ; $\cdots$; ν_{h-1} (because this would contradict coherence) then it is impossible that ν_h performs an input capture.

The case when μ_1 is an output capture is similar. □

We conclude this section by proving that the reachability problem for coherent structures has a computational complexity that is quadratic with respect to the number of gates in the structures. We use the notion of *distance* between two gates with the same type as the least number of reductions to convert one gate into the other (it is infinite, otherwise). Formally, the *distance* between two gates g and g' *of the same type*, written $|g-g'|$, is the commutative operation defined as follows:

– if $g=\mathtt{A}^\perp\,.\,\mathtt{A}_1{}^\perp\,.\,\widehat{}\,\mathtt{A}\,.\,\overline{\mathtt{B}}$ and $g'=\mathtt{A}^\perp\,.\,\mathtt{A}_2{}^\perp\,.\,\widehat{}\,\mathtt{A}\,.\,\overline{\mathtt{B}}$, where the first id of $\mathtt{A}_1{}^\perp$ is different from the first id of $\mathtt{A}_2{}^\perp$, then

$$|g-g'| \stackrel{\text{def}}{=} \mathit{length}(\mathtt{A}_1{}^\perp)+\mathit{length}(\mathtt{A}_2{}^\perp)$$

– if $g=\mathtt{A}^\perp\,.\,\mathtt{A}_1{}^\perp\,.\,\overline{\mathtt{B}_1}\,.\,\widehat{}\,\overline{\mathtt{B}_1'}$ and $g'=\mathtt{A}^\perp\,.\,\mathtt{A}_2{}^\perp\,.\,\overline{\mathtt{B}_2}\,.\,\widehat{}\,\overline{\mathtt{B}_2'}$, where the first id of $\mathtt{A}_1{}^\perp$ is different from the first id of $\mathtt{A}_2{}^\perp$, then

$$|g-g'| \stackrel{\text{def}}{=} \mathit{length}(\mathtt{A}_1{}^\perp\,.\,\overline{\mathtt{B}_1})+\mathit{length}(\mathtt{A}_2{}^\perp\,.\,\overline{\mathtt{B}_2})$$

– if $g=\mathtt{A}^\perp\,.\,\mathtt{A}_1{}^\perp\,.\,\widehat{}\,\mathtt{A}\,.\,\overline{\mathtt{B}}$ and $g'=\mathtt{A}^\perp\,.\,\mathtt{A}_2{}^\perp\,.\,\mathtt{A}\,.\,\overline{\mathtt{B}_1}\,.\,\widehat{}\,\overline{\mathtt{B}_2}$, where the first id of $\mathtt{A}_1{}^\perp$ is different from the first id of $\mathtt{A}_2{}^\perp$, then

$$|g-g'| \stackrel{\text{def}}{=} \mathit{length}(\mathtt{A}_1{}^\perp)+\mathit{length}(\mathtt{A}_2{}^\perp\,.\,\mathtt{A}^\perp\,.\,\overline{\mathtt{B}_1})$$

– if $g=\mathtt{A}^\perp\,.\,\mathtt{A}_1{}^\perp\,.\,\overline{\mathtt{B}_1}\,.\,\widehat{}\,\overline{\mathtt{B}_1'}$ and $g'=\mathtt{A}^\perp\,.\,\mathtt{A}_2{}^\perp\,.\,\overline{\mathtt{B}_2}\,.\,\widehat{}\,\overline{\mathtt{B}_2'}$, where the first id of $\mathtt{A}_1{}^\perp$ is different from the first id of $\mathtt{A}_2{}^\perp$, then

$$|g-g'| \stackrel{\text{def}}{=} \mathit{length}(\mathtt{A}_1{}^\perp\,.\,\overline{\mathtt{B}_1})+\mathit{length}(\mathtt{A}_2{}^\perp\,.\,\overline{\mathtt{B}_2})$$

The distance between two structures $\mathtt{S}$ and $\mathtt{S}'$ containing gates of the same types, noted $|\mathtt{S}-\mathtt{S}'|$, is

$$\sum_{g\in\mathtt{S},g'\in\mathtt{S}',\mathit{type}(g)=\mathit{type}(g')} |g-g'|.$$

PROPOSITION 4.4. *Let $\mathtt{S}$ be a coherent structure and let $\mathtt{S}\longrightarrow\mathtt{S}'\longrightarrow^*\mathtt{S}''$ be a minimal computation (according to Theorem 3.7). Then $|\mathtt{S}-\mathtt{S}''|>|\mathtt{S}'-\mathtt{S}''|$.*

PROOF. The proof is by contradiction and a case analysis on the reduction $\mathtt{S}\longrightarrow\mathtt{S}'$. One shows that, for every type of reduction, if $|\mathtt{S}-\mathtt{S}''|\le|\mathtt{S}'-\mathtt{S}''|$ (actually, the equality is not possible) then $\mathtt{S}'\longrightarrow^*\mathtt{S}''$ must revert the reduction $\mathtt{S}\longrightarrow\mathtt{S}'$ thus contradicting the assumption of minimality. □

The algorithm takes two coherent structures $\mathtt{S}$ and $\mathtt{S}'$ such that, for every gate in $\mathtt{S}$ there is a corresponding one in $\mathtt{S}'$ with the same type, and conversely. We assume that the structures are *lexicographically ordered* using the ids of the signal and the first ids of the type of gates (by coherence, the first ids are sufficient to discriminate gates). The reachability algorithm is specified as follows:

1. If $\mathtt{S}=\mathtt{S}'$ then the algorithm terminates with success;
2. otherwise, a gate g in $\mathtt{S}$ is chosen with non-null distance from the corresponding one g' in $\mathtt{S}'$ and such that it may be reduced in $\mathtt{S}$ by decreasing its distance from g'. Let $\mathtt{S}\longrightarrow\mathtt{S}''$ be such reduction (by construction $|\mathtt{S}-\mathtt{S}'|>|\mathtt{S}''-\mathtt{S}'|$).
   (a) if no such reduction is possible the algorithm terminates with failure;
   (b) otherwise the algorithm returns to 1, replacing $\mathtt{S}$ with $\mathtt{S}'$.

The data structures of the algorithm are two arrays. The first one stores the gates and is addressed using the first id of their type (by coherence, the first ids are sufficient to discriminate gates). The second array stores signals. The elements are accessed through the co-name of the signal. Every element is a boolean array that is accessed through the id and containing `true` or `false` according to whether the corresponding signal is present or absent, respectively. Let n be the number of gates in $\mathtt{S}$ and let k be the maximal length of a gate in $\mathtt{S}$. The step 2 of the algorithm may require (i) a complete visit of the array of gates, that costs n, and, for each element, (ii) a gate analysis for determining the distance and the possible reduction that costs k. Since in the worst case, gates may be at distance $2k$, the algorithm may iterate $2k\times n$ times. Then its computational complexity is $O(2k^2\times n^2)$. It is worth to remark that the computational complexity of the reachability problem in

(weak coherent) reversible structures reduces to the reachability marking problem in bounded place-transition Petri nets, which is EXPSPACE complete [12, 6] and we are not aware of any better algorithm for not coherent structures.

5. THE ENCODING OF ASYNCHRONOUS RCCS

Coherent structures can encode a process calculus with a reversible transition relation: the *asynchronous* RCCS [7]. This allows one to precisely assess the expressive power of coherent structures and to establish properties of asynchronous RCCS using those of coherent structures, such as Theorem 4.3, which has been proved for RCCS in [7], or the above algorithm of reachability, which is original.

The syntax of asynchronous RCCS uses an infinite set of *names*, ranged over by $a, b, c, \cdots$, and a disjoint set of *co-names*, ranged over by $\overline{a}, \overline{b}, \overline{c}, \cdots$. Names and co-names are ranged over by $\alpha, \beta, \cdots$ and are generically called *actions*. *Processes* P are defined by the following grammar (we are assuming that I are finite sets):

$$P \ ::= \ \mathbf{0} \quad | \quad \textstyle\sum_{i\in I} \alpha_i.P_i \quad | \quad \prod_{i\in I} P_i \quad | \quad (\mathtt{new}\ a)\ P$$

The term $\mathbf{0}$ defines the terminated process; $\sum_{i\in I} \alpha_i.P_i$ defines a process that may perform one action α_i and continues as P_i; $\prod_{i\in I} P_i$ defines the parallel composition of processes P_i; finally the term $(\mathtt{new}\ a)\ P$ defines a name with scope P. Processes meet the following well-formed conditions:

- in $\overline{a}.P$, the process P is $\mathbf{0}$ (continuations of co-names are empty);
- in $\prod_{i\in I} P_i$ the processes P_i are guarded choices.

The semantics of asynchronous RCCS is defined in terms of a *transition relation* that uses

- *memories* m:

$$m \ ::= \ \langle\,\rangle \ \mid \ \langle i\rangle_n \bullet m \ \mid \ \langle m, \alpha, Q\rangle \bullet m$$

- *run-time processes* R:

$$R \ ::= \ m \triangleright P \ \mid \ R \mid R \ \mid \ (\mathtt{new}\ a)\ R$$

- *structural congruence* $\equiv$, defined in the standard way (see Section 2), plus the rules

$$m \triangleright (\textstyle\prod_{i\in 1..n} P_i) \ \equiv \ \prod_{i\in 1..n} \langle i\rangle_n \bullet m \triangleright P_i$$

$$m \triangleright (\mathtt{new}\ a)\ P \ \equiv \ (\mathtt{new}\ a)\ (m \triangleright P) \qquad (a \notin \mathsf{fn}(m))$$

The reduction relation $\longrightarrow$ is the least relation on run-time processes satisfying the axioms:

- $m \triangleright (a.P + Q) \mid m' \triangleright (\overline{a} + R) \longrightarrow \langle m', a, Q\rangle \bullet m \triangleright P \mid \langle m, \overline{a}, R\rangle \bullet m' \triangleright \mathbf{0}$,
- $\langle m', a, Q\rangle \bullet m \triangleright P \mid \langle m, \overline{a}, R\rangle \bullet m' \triangleright \mathbf{0} \longrightarrow m \triangleright (a.P + Q) \mid m' \triangleright (\overline{a} + R)$,

and closed under the contextual rules for parallel, new and structural congruence.

Let us have a short discussion that illustrates the main ideas of our encoding of RCCS before going into the details. Consider the process $a.\overline{b} + \overline{a}$ that may progress either as $\overline{b}$ or terminate according to whether the external environment offers an output or an input on a, respectively. The structure encoding this process is

$$(\mathtt{new}\ c')\ ((\hat{}c\,.\,a\,.\,u{:}\overline{c'} \ \mid \ \hat{}c'\,.\,u'{:}\overline{b}) \ \mid \ \hat{}c\,.\,v{:}\overline{a})$$

We assume that the environment may emit at most one signal with co-name $\overline{c}$. When such a signal arrives, one of the gates $\hat{}c\,.\,a\,.\,u{:}\overline{c'}$ and $\hat{}c\,.\,v{:}\overline{a}\,.\,w{:}\overline{c''}$ will react, let it be the second. Then the structure becomes

$$(\mathtt{new}\ c')\ ((\hat{}c\,.\,a\,.\,u{:}\overline{c'} \ \mid \ \hat{}c'\,.\,u'{:}\overline{b}) \ \mid \ u'{:}c\,.\,\hat{}v{:}\overline{a})$$

that emits a signal $v{:}\overline{a}$. It is crucial for the correctness of the encoding that $v{:}\overline{a}$ cannot interact with any other branch of the choice, *i.e.* with the gate $\hat{}c\,.\,a\,.\,u{:}\overline{c'}$. At this stage, it is possible that the context offers a signal $v'{:}\overline{a}$ rather than accepting signals $v{:}\overline{a}$. That is, the local choice of the process does not match the choice of the context. Reversibility plays a crucial role at this point. In fact, the above reductions are reverted; the signal $u'{:}\overline{c}$ is re-emitted, and the left branch of the above choice is chosen, thus obtaining the structure

$$(\mathtt{new}\ c')\ ((u'{:}c\,.\,\hat{}a\,.\,u{:}\overline{c'} \ \mid \ \hat{}c'\,.\,u'{:}\overline{b}) \ \mid \ \hat{}c\,.\,v{:}\overline{a})$$

that may accept the signal $v'{:}\overline{a}$. Notice that RCCS memories are implemented by inactive processes that are in parallel with the active ones. No ad-hoc memory management operation is used.

Our encoding uses environments Γ that map memories to signals $u{:}\overline{c}$ such that:

1. (Γ is injective) if $m \neq m'$ then the signals $\Gamma(m)$ and $\Gamma(m')$ are different (they have different ids and co-names);
2. (Γ is prefix closed) if $s \bullet m \in \mathtt{dom}(\Gamma)$ then $m \in \mathtt{dom}(\Gamma)$ and
   - if $s = \langle i\rangle_n$ then $\langle 1\rangle_n \bullet m, \cdots, \langle n\rangle_n \bullet m \in \mathtt{dom}(\Gamma)$;
   - if $s = \langle m', \alpha, P\rangle$ then $m' \in \mathtt{dom}(\Gamma)$.

Let $\mathsf{fn}(\Gamma) \stackrel{\text{def}}{=} \{a \mid \exists\, m, u.\ \Gamma(m) = u{:}\overline{a}\}$.

Let Γ be an environment such that, for every $i \in I$, $\Gamma(m_i) = u_i{:}\overline{c_i}$, for some u_i. The encoding $[\![\cdot]\!]^\Gamma$ is defined by

$$[\![\textstyle\prod_{i\in I} m_i \triangleright P_i]\!]^\Gamma \ = \ (\mathtt{new}\ \mathsf{fn}(\Gamma))\ \Big(\ \textstyle\prod_{i\in I}([\![P_i]\!]_{c_i} \ \mid \ \Gamma(m_i)) \ \mid \ \mathcal{U}(\{m_i \mid i \in I\}, \Gamma)\Big)$$

where the auxiliary functions $[\![P]\!]_c$ and $\mathcal{U}(M, \Gamma)$ are defined in Figure 2. As a consequence of the definition, the function $[\![\cdot]\!]^\Gamma$ always returns a coherent structure.

LEMMA 5.1. *If $R \longrightarrow R'$ then, for suitable Γ and Γ', $[\![R]\!]^\Gamma \longrightarrow^* [\![R']\!]^{\Gamma'}$, with $id(\mu_i) \cap id(\Gamma(m) \mid \Gamma(m')) \neq \emptyset$.*

PROOF. Without loss of generality, let

$$R = m \triangleright \sum_{i\in I} a_i.P_i + \sum_{j\in J} \overline{a_j} \ \mid \ m' \triangleright \sum_{i\in I'} b_i.Q_i + \sum_{j\in J'} \overline{b_j} \ \mid \ m'' \triangleright P''$$

and let $\overline{a_h} = b_k = \overline{a}$ and $P' = \sum_{i\in I\setminus\{h\}} a_i.P_i + \sum_{j\in J} \overline{a_j}$ and $Q' = \sum_{i\in I'} b_i.Q_i + \sum_{j\in J'\setminus\{k\}} \overline{b_j}$. Then $R \longrightarrow \langle m', a, P'\rangle \bullet m \triangleright P_h \ \mid \ \langle m, \overline{a}, Q'\rangle \bullet m' \triangleright \mathbf{0} \ \mid \ m'' \triangleright P''$.

$$
\begin{array}{rcl}
[\![\mathbf{0}]\!]_c & = & \hat{}c \\
[\![\sum_{i\in I} a_i.P_i + \sum_{j\in J} \overline{a_j}]\!]_c & = & (\mathtt{new}\ c_i^{\,i\in I})\ (\prod_{i\in I}(\hat{}c\,.\,a_i\,.\,u_i{:}\overline{c_i}\ \mid\ \prod_{i\in I}[\![P_i]\!]_{c_i})\ \mid\ \prod_{j\in J}\hat{}c\,.\,u_j{:}\overline{a_j}) \\
 & & \qquad\qquad \text{where, for every } \ell\in I\cup J,\ u_\ell, u'_\ell \text{ are fresh} \\
[\![\prod_{i\in 1..n} P_i]\!]_c & = & (\mathtt{new}\ c_i^{\,i\in 1..n})\ (\hat{}c\,.\,u_1{:}\overline{c_1}\,.\,\cdots\,.\,u_n{:}\overline{c_n}\ \mid\ \prod_{i\in 1..n}[\![P]\!]_{c_i}) \qquad \text{where } u_1,\cdots,u_n \text{ are fresh} \\
[\![(\mathtt{new}\ a)\ P]\!]_c & = & (\mathtt{new}\ a)\ ([\![P]\!]_c)
\end{array}
$$

$$
\begin{array}{rcl}
\mathcal{U}(\{\langle 1\rangle_n \bullet m,\cdots,\langle n\rangle_n \bullet m\}\uplus M,\Gamma) & = & u{:}c\,.\,v_1{:}\overline{c_1}\,.\,\cdots\,.\,v_n{:}\overline{c_n}\hat{}\ \mid\ \mathcal{U}(\{m\}\uplus M,\Gamma) \\
 & & \text{where } \Gamma(m) = u{:}\overline{c} \text{ and } \Gamma(\langle i\rangle_n \bullet m) = v_i{:}\overline{c_i} \\
\mathcal{U}(\{\langle m',a,P\rangle \bullet m, \langle m,\overline{a},Q\rangle \bullet m'\}\uplus M,\Gamma) & = & u{:}c\,.\,w{:}a\,.\,v{:}\overline{c_1}\hat{}\ \mid\ [\![P]\!]_c\ \mid\ u'{:}c'\,.\,w{:}\overline{a}\hat{}\ \mid\ [\![Q]\!]_{c'}\ \mid\ \mathcal{U}(\{m,m'\}\uplus M,\Gamma) \\
 & & \text{where } \Gamma(m) = u{:}\overline{c} \text{ and } \Gamma(m') = u'{:}\overline{c'} \text{ and } \Gamma(\langle m',a,P\rangle \bullet m) = v{:}\overline{c_1} \\
 & & \text{and } w \text{ fresh}
\end{array}
$$

Figure 2: The functions $[\![P]\!]_c$ and $\mathcal{U}(M,\Gamma)$

Let Γ be such that $\Gamma(m) = w{:}\overline{c}$, $\Gamma(m') = w'{:}\overline{c'}$ and $\Gamma(m'') = w''{:}\overline{c''}$. Then

$$
\begin{array}{l}
[\![R]\!]^\Gamma \\
\longrightarrow\ (\mathtt{new}\ \widetilde{c})\ (\quad w{:}c\,.\,\hat{}a\,.\,u_k{:}\overline{c_k}\ \mid\ [\![P_k]\!]_{c_k}\ \mid\ \prod_{i\in I\cup J\setminus\{h\}}[\![P_i]\!]_c \\
\qquad\qquad \mid\ w'{:}\overline{c'}\ \mid\ \prod_{i\in I'\cup J'}[\![Q_i]\!]_{c'} \\
\qquad\qquad \mid\ w''{:}\overline{c''}\ \mid\ [\![P'']\!]_{c''}\ \mid\ \mathcal{U}(\{m,m',m''\},\Gamma)\quad) \\
\longrightarrow\ (\mathtt{new}\ \widetilde{c})\ (\quad w{:}c\,.\,\hat{}a\,.\,u_k{:}\overline{c_k}\ \mid\ [\![P_k]\!]_{c_k}\ \mid\ \prod_{i\in I\cup J\setminus\{h\}}[\![P_i]\!]_c \\
\qquad\qquad \mid\ w'{:}c'\,.\,\hat{}v'_h{:}\overline{a}\ \mid\ \prod_{i\in I'\cup J'\setminus\{h\}}[\![Q_i]\!]_{c'} \\
\qquad\qquad \mid\ w''{:}\overline{c''}\ \mid\ [\![P'']\!]_{c''}\ \mid\ \mathcal{U}(\{m,m',m''\},\Gamma)\quad) \\
\longrightarrow\ (\mathtt{new}\ \widetilde{c})\ (\quad w{:}c\,.\,\hat{}a\,.\,u_k{:}\overline{c_k}\ \mid\ [\![P_k]\!]_{c_k}\ \mid\ \prod_{i\in I\cup J\setminus\{h\}}[\![P_i]\!]_c \\
\qquad\qquad \mid\ v'_h{:}\overline{a}\ \mid\ w'{:}c'\,.\,v'_h{:}\overline{a}\,.\,\hat{}\ \mid\ \prod_{i\in I'\cup J'\setminus\{h\}}[\![Q_i]\!]_{c'} \\
\qquad\qquad \mid\ w''{:}\overline{c''}\ \mid\ [\![P'']\!]_{c''}\ \mid\ \mathcal{U}(\{m,m',m''\},\Gamma)\quad) \\
\longrightarrow\ (\mathtt{new}\ \widetilde{c})\ (\quad w{:}c\,.\,v'_h{:}a\,.\,\hat{}u_k{:}\overline{c_k}\ \mid\ [\![P_k]\!]_{c_k}\ \mid\ \prod_{i\in I\cup J\setminus\{h\}}[\![P_i]\!]_c \\
\qquad\qquad \mid\ w'{:}c'\,.\,v'_h{:}\overline{a}\,.\,\hat{}\ \mid\ \prod_{i\in I'\cup J'\setminus\{h\}}[\![Q_i]\!]_{c'} \\
\qquad\qquad \mid\ w''{:}\overline{c''}\ \mid\ [\![P'']\!]_{c''}\ \mid\ \mathcal{U}(\{m,m',m''\},\Gamma)\quad) \\
\longrightarrow\ (\mathtt{new}\ \widetilde{c})\ (\quad u_k{:}\overline{c_k}\ \mid\ [\![P_k]\!]_{c_k}\ \mid\ w{:}c\,.\,v'_h{:}a\,.\,u_k{:}\overline{c_k}\hat{}\ \\
\qquad\qquad \mid\ \prod_{i\in I\cup J\setminus\{h\}}[\![P_i]\!]_c \\
\qquad\qquad \mid\ w'{:}c'\,.\,v'_h{:}\overline{a}\,.\,\hat{}\ \mid\ \prod_{i\in I'\cup J'\setminus\{h\}}[\![Q_i]\!]_{c'} \\
\qquad\qquad \mid\ w''{:}\overline{c''}\ \mid\ [\![P'']\!]_{c''}\ \mid\ \mathcal{U}(\{m,m',m''\},\Gamma)\quad) \\
= [\![R']\!]^{\Gamma'}
\end{array}
$$

where

$$\Gamma' = \Gamma[\langle m',a,P'\rangle \bullet m \mapsto u_k{:}\overline{c_k}\ ;\ \langle m,\overline{a},Q'\rangle \bullet m' \mapsto v''_h{:}\overline{c_h}]$$

Since our structures and asynchronous RCCS are both reversible, the above computation also demonstrates that $R' \longrightarrow R$ implies $[\![R']\!]^{\Gamma'} \longrightarrow^* [\![R]\!]^\Gamma$. $\square$

It is possible to define the reverse encoding of $[\![\cdot]\!]^\Gamma$. Given a coherent reversible structure $\prod_{i\in 1..m} g_i \mid \prod_{i\in m+1..n} u_j{:}\overline{a_j}$, the corresponding asynchronous RCCS process is $\prod_{i\in 1..m}\langle i\rangle_n\cdot \breve{g_i} \mid \prod_{i\in m+1..n}\langle i\rangle_n \cdot \overline{a_j}$, where $\breve{g_i}$ are gates without ids, and a mapping (similar to Γ) associates memories to ids. It is not difficult to demonstrate a correspondence between the two terms similar to Lemma 5.1.

6. IRREVERSIBLE COMBINATORS

Having developed a theory for reversible structures, we now analyze how to extend our calculus in order to integrate (irreversible) strand algebra [4].

Let us add an *irreversible co-name*, noted $\overline{\mathtt{d}}$, to the set of co-names. There is no name corresponding to $\overline{\mathtt{d}}$ and we assume that occurrences of $\mathtt{d}$ in gates, if any, are in the last position. For example, $\hat{}a\,.\,u{:}\overline{b}\,.\,v{:}\overline{\mathtt{d}}$ and $w{:}a\,.\,u{:}\overline{b}\,.\,v{:}\overline{\mathtt{d}}\hat{}$ are valid gates. We also assume that structures never retain signals with co-name $\overline{\mathtt{d}}$.

The reduction relation of Definition 2.2 is then expanded with

$$\mathtt{A}^\perp\,.\,\overline{\mathtt{B}}\,.\,\hat{}u{:}\overline{\mathtt{d}} \longrightarrow \mathtt{A}^\perp\,.\,\overline{\mathtt{B}}\,.\,u{:}\overline{\mathtt{d}}\hat{}$$

that, contrary to the other output release rules, does not emit any signal $u{:}\overline{\mathtt{d}}$ and has no associated converse reduction. The intended meaning is that, while reversibility is admitted up-to the signal preceding $\overline{\mathtt{d}}$, it is forbidden once $\overline{\mathtt{d}}$ has been consumed.

This extension of reversible structures is expressive enough to encode

- the (irreversible) strand algebra in [4]. We only illustrate the encoding $[\![\cdot]\!]$ of an irreversible strand $a\,.\,\overline{b}$ (v and w are fresh ids):

$$[\![a\,.\,\overline{b}]\!] \stackrel{\text{def}}{=} a\,.\,v{:}\overline{b}\,.\,w{:}\overline{\mathtt{d}}\,.$$

- the (irreversible) guarded choice in asynchronous CCS. We only illustrate the encoding $\lceil\cdot\rceil$ of $a.P+\overline{b}$ (u,u',v,v' are fresh ids):

$$\begin{array}{rcl}\lceil a.P+\overline{b}\rceil \stackrel{\text{def}}{=} & (\hat{}c\,.\,a\,.\,u{:}\overline{c'}\,.\,u'{:}\overline{\mathtt{d}}\ \mid\ \lceil P\rceil_{c'}) \\ & \mid\ (\hat{}c\,.\,v{:}\overline{b}\,.\,v'{:}\overline{c''}\ \mid\ \lceil\mathbf{0}\rceil_{c''})\end{array}$$

  where the irreversible reduction is performed once the input continuation has been triggered. The encoding of output has no tailing $\overline{\mathtt{d}}$ combinator.

Figure 3 defines the translation of the gate $a\,.\,v{:}\overline{b}\,.\,w{:}\overline{\mathtt{d}}$ in the DSD language. According to the semantics of the DSD

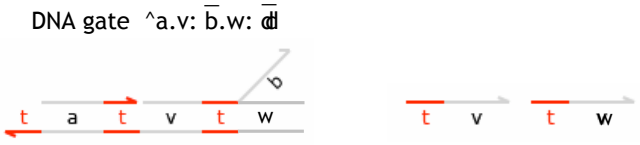
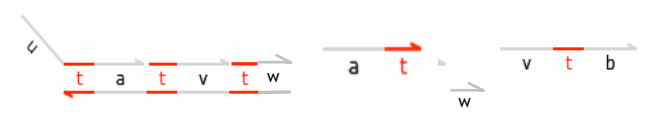


Figure 3: The encoding of $a \bullet v{:}\bar{c} \bullet w{:}\bar{d}$ in the dsd language

language, if a strand `<u t~ a>` arrives we will obtain the DSD term



that cannot be reverted to the initial one. We ignore the minor issue of the inert "w" segment that is left over by the DSD reduction.

7. CONCLUSIONS

We have developed a reversible concurrent calculus that is amenable to biological implementations in terms of DNA circuits and is expressive enough to encode a reversible process calculus such as asynchronous RCCS.

This study can be extended in several directions. The encoding of RCCS is given in terms of coherent structures. For this reason asynchronous RCCS bears Theorem 4.3 (that has been already proved for RCCS in [7]), and an efficient algorithm of reachability. However coherence – a solution must contain exactly one molecule of every species – is very hard to achieve in nature, even if it will become easier in the future. So, biology prompts a thorough study of reversible concurrent calculi where processes have multiplicities and the causal dependencies between copies may be exchanged. Section 3 is a preliminary study of this matter.

Another direction is about implementations. In this paper we have discussed the implementation of a concurrent language in biomolecules. The presence of irreversible combinators makes this implementation more interesting because it paves the way for modelling standard (irreversible) constructs of programming languages. Comparing biological *in vivo* implementations and standard *in silico* implementations of programming languages is an exciting research direction both for biology and computer science.

Our study about reachability has been inspired by biology and retains an easy solution in reversible structures because of their simplicity. Studying other biological relevant problems, such as detecting the absence of molecules/processes, stable concentrations of materials, etc., and designing efficient algorithms are other directions that need to be investigated in reversible structures and may bear simple solutions in this model.

8. REFERENCES

[1] C. H. Bennett. Logical reversibility of computation. *IBM J. Res. Dev.*, 17(6):525–532, 1973.

[2] G. Berry and G. Boudol. The chemical abstract machine. In *Proceedings of POPL'90*, pages 81–94. ACM, 1990.

[3] G. Boudol and I. Castellani. Permutation of transitions: An event structure semantics for CCS and SCCS. In *Linear Time, Branching Time and Partial Order in Logics and Models for Concurrency*, volume 354 of *Lecture Notes in Computer Science*, pages 411–427. Springer, 1989.

[4] L. Cardelli. Strand algebras for DNA computing. In *DNA 2009*, volume 5877 of *Lecture Notes in Computer Science*, pages 12–24, 2009.

[5] L. Cardelli. Two-domain DNA strand displacement. In *Developments in Computational Models (DCM 2010)*, volume 25 of *EPTCS*, pages 33–47, 2010.

[6] E. Cardoza, R. J. Lipton, and A. R. Meyer. Exponential space complete problems for Petri Nets and commutative semigroups: Preliminary report. In *Eighth Annual ACM Symposium on Theory of Computing*, pages 50–54. ACM, 1976.

[7] V. Danos and J. Krivine. Reversible communicating systems. In *CONCUR 2004*, volume 3170 of *Lecture Notes in Computer Science*, pages 292–307, 2004.

[8] P. Degano, J. Meseguer, and U. Montanari. Axiomatizing net computations and processes. In *LICS'89*, pages 175–185. IEEE Computer Society, 1989.

[9] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem*, 81:2340–2361, 1977.

[10] I. Lanese, C. A. Mezzina, and J.-B. Stefani. Reversing higher-order pi. In *Proceedings of CONCUR 2010*, volume 6269 of *Lecture Notes in Computer Science*, pages 478–493. Springer, 2010.

[11] J.-J. Lévy. An algebraic interpretation of the *lambda-beta-k*-calculus; and an application of a labelled *lambda* -calculus. *Theor. Comput. Sci.*, 2(1):97–114, 1976.

[12] E. W. Mayr and A. R. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Adv. in Math.*, 46(3):305–329, 1982.

[13] A. Phillips and L. Cardelli. A programming language for composable DNA circuits. *Journal of the Royal Society Interface*, 6(S4), 2009.

[14] I. Phillips and I. Ulidowski. Reversibility and models for concurrency. In *Proceedings of SOS 2007*, volume 192 of *ENTCS*, pages 93–108, 2007.

[15] I. Phillips and I. Ulidowski. Reversing algebraic process calculi. *J. Log. Algebr. Program.*, 73(1-2):70–96, 2007.