A graph-based tool to embed the π-calculus into a computational DPO framework

Andersen, Jakob L.; Hellmuth, Marc; Merkle, Daniel; Nøjgaard, Nikolai; Peressotti, Marco

Published in:
CEUR Workshop Proceedings

Publication date:
2020

Document version
Final published version

Document license
CC BY

Citation for published version (APA):

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:

• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk
A Graph-Based Tool to Embed the π-Calculus into a Computational DPO Framework

Jakob L. Andersen\textsuperscript{1}, Marc Hellmuth\textsuperscript{2}, Daniel Merkle\textsuperscript{1,3}, Nikolai Nøjgaard\textsuperscript{1,2}, and Marco Peressotti\textsuperscript{1}

\textsuperscript{1} Department of Mathematics and Computer Science, University of Southern Denmark, Odense, DK
\{jlandersen,daniel,nojgaard,peressotti\}@imada.sdu.dk
\textsuperscript{2} Institute of Mathematics and Computer Science, University of Greifswald, DE
mhellmuth@mailbox.org
\textsuperscript{3} Harvard Medical School, Department of Systems Biology, Boston, MA, US

Abstract Graph transformation approaches have been successfully used to analyse and design chemical and biological systems. Here we build on top of a DPO framework, in which molecules are modelled as typed attributed graphs and chemical reactions are modelled as graph transformations. Edges and vertexes can be labelled with first-order terms, which can be used to encode, e.g., steric information of molecules. While targeted to chemical settings, the computational framework is intended to be very generic and applicable to the exploration of arbitrary spaces derived via iterative application of rewrite rules, such as process calculi like Milner’s π-calculus. To illustrate the generality of the framework, we introduce \textit{EpiM}: a tool for computing execution spaces of π-calculus processes. \textit{EpiM} encodes π-calculus processes as typed attributed graphs and then exploits the existing DPO framework to compute their dynamics in the form of graphs where nodes are π-calculus processes and edges are reduction steps. \textit{EpiM} takes advantage of the graph-based representation and facilities offered by the framework, like efficient isomorphism checking to prune the space without resorting to explicit structural equivalences. \textit{EpiM} is available as an online Python-based tool.

Keywords: Double Pushout, Process Calculi, Typed Attributed Graphs, Graph Isomorphism

1 Introduction

Graph transformation approaches have been shown to provide formalisms that elegantly facilitate the construction of reaction rules in organic chemistry and biology [2,10,14]. Many of these frameworks, while aimed at modelling organic chemistry and biology, are constructed as generic foundations which can be used to model in a wider variety of domains.

The framework focused on in this paper is \textit{MØD} and is traditionally used to model organic chemistry [2,8]. Here, graphs represent molecules, transformation rules specify how molecules can interact, and direct derivations represent...
concrete chemical reactions. Specifically, graph transformations are in \textsc{mod} performed with the double-pushout (DPO) approach, with injective graph morphisms \cite{11,13}. A higher-level “strategy framework” then allows for programming the sequence of rules to apply on sets of graphs, while performing the necessary graph isomorphism checks \cite{1}. Additionally, graphs in \textsc{mod} are labelled with first-order terms which allows for the specification of even more abstract rules with named variables as attributes.

In this paper we will use the foundational framework of \textsc{mod} to implement a tool for computing execution spaces of \(\pi\)-calculus processes. \(\pi\)-calculus is a process algebra that, like all process algebras, is concerned with the problem of formally modelling concurrent systems \cite{18,19}. A process in \(\pi\)-calculus represents a “unit” of computation. In a concurrent system several of such processes can run concurrently and communication between processes occurs via channels on which channel names are sent. Given a set of processes, all running concurrently, the execution space is then all possible state transitions up to structural congruence.

The modelling of \(\pi\)-calculus as a graph transformation system is not a new concept \cite{12,17}. Similarly, tools already exist that simulates \(\pi\)-calculus \cite{7,20}. Few tools, however, simulates the execution space of a process up to structural congruence, and no tools, to our knowledge, does this by using established graph transformation concepts such as the DPO approach.

Here, we present the tool \textsc{epim}, which is a Python-based library that embeds the language of \(\pi\)-calculus into the framework of \textsc{mod}. It allows the modelling of processes as simple Python expressions which are then encoded into graphs. The execution space of processes is modelled as a set of transformation rules applied on their corresponding graph encodings. Structural congruence is checked using efficient graph isomorphism checking of graphs labelled with first-order terms provided by \textsc{mod}. The encoding and the simulation of process communication via graph transformations are based on the results established in \cite{12}, with the notable modification that our encoding obtains simple labelled graphs, while their model is concerned with directed typed hypergraphs. A web front-end of \textsc{epim} is provided at \url{http://cheminf.imada.sdu.dk/epim}.

2 Preliminaries

2.1 The \(\pi\)-calculus

The \(\pi\)-calculus is a mathematical model of concurrent interacting processes \((P, Q, R, \ldots)\) that communicate using names \((x, y, z, \ldots)\). A name represents a communication channel or session, and can be sent and received as a message \cite{18,19}. Process terms are given by the following grammar:

\[
P, Q ::= \pi[y].P \quad \text{output name } y \text{ on channel } x \text{ and continue as } P \\
| x(y).P \quad \text{input a name on } x, \text{ bind it to } y \text{ and continue as } P \\
| P|Q \quad \text{run processes } P \text{ and } Q \text{ in parallel (parallel composition)} \\
| P + Q \quad \text{run either } P \text{ or } Q \text{ (choice)} \\
| (\nu x)P \quad \text{bind } x \text{ in } P \text{ (restriction)} \\
| 0 \quad \text{terminated process}
\]
The sets \( fn(P) \) and \( bn(P) \) of free and bound names in a process \( P \) are defined as expected as well as \( \alpha \)-conversion.

Intuitively, process terms that differ solely on the order of parallel composition \( | \), sums +, and restrictions \( \nu \) represent the same process. This intuition is formalised by structural equivalence, i.e., the least relation \( \equiv \) on process terms that is a congruence w.r.t. the grammar above and closed under the Abelian laws for \( \diamond \in \{ |, + \} \) under 0:

\[
P \diamond Q \equiv Q \diamond P \quad (P \diamond Q) \diamond R \equiv P \diamond (Q \diamond R) \quad P \diamond 0 \equiv 0
\]

and under the distributivity laws for restriction:

\[
(\nu x)(\nu y)P \equiv (\nu y)(\nu x)P \quad (\nu x)0 \equiv 0 \quad (\nu x)(P | Q) \equiv (\nu y)P | Q \text{ for } x \notin fn(Q).
\]

For the sake of simplicity, we adopt the common assumption that choices are always guarded (i.e. every branch in a sum + is either a choice or a communication): in the sequel we assume, similarly to [12], the following grammar:

\[
P, Q ::= M | P | Q
\]

\[
M, N ::= 0 | x(y).P | [y].P | M + N
\]

The semantics of process terms is given by the reduction relation \( \rightarrow \), i.e., the least relation closed under the rules below:

\[
\frac{P \equiv P'}{P \rightarrow \text{COM
}} \quad \frac{P' \rightarrow Q'}{P \rightarrow Q}
\]

\[
\frac{P \rightarrow Q}{\text{STR}} \quad \frac{P \rightarrow Q}{\text{RES}} \quad \frac{P | Q \rightarrow P'| Q}{\text{PAR}}
\]

Rule Com models synchronous communication between two processes, possibly under a non-deterministic context (subterm \( M \) may simply be 0). Rule Str ensures structurally equivalent processes have the same behaviour. Rule Par models the (interleaved) execution of parallel components. Rule Res allows execution under a restriction.

### 2.2 Graphs, Transformation Rules, and the MØD framework

In this section we will be giving a very brief overview of the MØD framework and refer to [2] for a full overview. Encodings of processes will be represented as simple labelled graphs, where vertices and edges of graphs are labelled with first-order terms. We use the common definitions of graph and monomorphisms for labelled graphs, where we require a most general unifier to exist for the set of term mappings induced by the underlying graph morphisms. For full details see Companion Technical Report [6].

Graph transformation rules are modelled in the traditional DPO framework with injective morphisms. See [13] of an overview of the DPO approach. The span of a rule, \( L \leftarrow K \rightarrow R \) is illustrated in Fig. [4]. Using first-order terms as labels is useful, since it allows us to specify abstract graph transformation rules
such as in Fig. 4, where any term leading with an underscore, e.g., \(_X\), specifies a variable term.

The application of graph transformation rules on some input graph, is done by using a strategy framework, where a specific order in which the graph transformation rules should be applied can be specified. The direct derivations obtained from applying a given strategy on an input graph is illustrated by a derivation graph. A derivation graph is a directed graph, where vertices are graphs obtained from rule applications and where an edge \((G, H)\), represents a direct derivation \(G \Rightarrow H\). Any pair of graphs in the derivation graph are non-isomorphic, and as a result the closure up to isomorphism of a given strategy can be automatically computed by MOD. See Fig. 2 for an example.

### 3 Encoding \(\pi\)-calculus

Suppose we wanted to model the process \(P = x(z).\pi[w] | \pi[y]\) as a graph. To encode \(P\) in EpIM we can write the following:

```plaintext
1 x, y, z, w = names("x y z w")
2 p1 = Process().input(x, z).output(z, w)
3 p2 = Process().output(x, y)
4 P = p1 | p2
5 G = P.encode()
6 G.print()
```

Line 1 defines the names to be used in \(P\). Line 2–3 defines the behaviour of each subprocess in \(P\), while line 4 defines \(P\) as the parallel composition between them. Finally, line 5 encodes \(P\) as a graph that is printed in line 6. The resulting graph is illustrated in Fig. 1a. The encoding is based on [12] modified to work on simple labelled graphs instead of typed directed hypergraphs.

We denote the graph encoded by a process \(P\) as \([P]\). The overall topology of \([P]\) can be thought of as tree-like, describing the behaviour of \(P\), just like a syntax tree for a context-free grammar, but where names are shared between operators. Since MOD works on undirected graphs, \([P]\) is equipped with a special vertex with the label \(go\), representing the “root” of the tree. This will be useful later, to be able to specify the “top” of the graph in transformation rules.

All other vertices in \([P]\) is wrapped in one of two functions \(v(\_X)\) or \(t(\_X)\). Any vertex equipped with the function \(v\) represents a name with the value of \(_X\). On the other hand, vertices equipped with \(t\) describes the behaviour of the corresponding process or represents an implementation detail that will prove useful later. More precisely, the argument of \(t\) can represent input/output operators \(t(in)/t(out)\), or parallel/sum composition operators \(t(p)/t(s)\).

We allow vertices with the terms \(t(p)\) or \(t(s)\) to have a degree less than 3 as seen in Fig. 1a. In this case, the vertices do not actually represent a corresponding parallel or sum composition, but are useful implementation details when implementing the reduction mechanism with graph transformation. Any input or output operator is followed by a vertex with the term \(t(p)\), regardless of \(t(p)\) representing a parallel composition or not. Moreover, every vertex with the term \(t(p)\), is either a leaf or immediately followed by a vertex with the label \(t(s)\).
Any prefix operator points to a name corresponding to the received or sent channel. Since the graphs considered here are not embedded, and hence do not have an order on the neighbourhood of a vertex, we label an edge emanating from a prefix operator with “sync” or “arg” whether the corresponding name is used for synchronization or as the argument for the operator. If a name is both the channel and argument of an operator, then the corresponding edge is equipped with “arg-sync” since MOD does not support parallel edges.

The vertices corresponding to names of the process are shared among vertices corresponding to input and output operators. In this way, if an input and output operator synchronize on the same name, it is illustrated in the resulting graph by both of them pointing to the vertex corresponding to that name. For example, the outermost input and output operator of the process encoded in Fig. 1a both synchronize on the free name $x$, and hence they both contain an edge to the vertex with the term $v(x)$.

Many of the vertices found in $[P]$ are implementation details, that are useful for simulating process reductions, but makes it difficult to interpret the behaviour of $P$ simply by looking at $[P]$. In this regard, by default, EpiM filters away such implementation details, and instead depicts a simplified version of the encoding illustrated in Fig. 1b.

4 Computing Execution Spaces

Suppose we were given the process $P = x(z).\pi[w] | (\pi[y] + \pi[y])$ and we are able to encode it into its graph equivalent $[P]$. Supplied with $[P]$, we want to compute the execution space of $P$. In EpiM we can write the following:

```plaintext
1 x, y, z, w = names("x y z w")
2 p1 = Process().input(x, z).output(z, w)
3 p2 = Process().output(x, y) + Process().output(x, y)
4 P = (p1 | p2)
5 exec_space = ReductionDG(P)
```
Line 1–4 specifies the process \( P \) as explained in the previous section. Line 5–6 computes all possible reductions, while the result is printed in line 7. The result is illustrated in Fig. 2. The execution space of \( P \) is illustrated as a directed graph \( G \) where each vertex in \( G \) corresponds to a process encoding, \([Q]\), derived from reductions on \( P \), while each edge, \(([Q], [R])\), represents the reduction \( Q \rightarrow R \), for some processes \( Q \) and \( R \). For instance, Fig. 3 contains the transformation that encodes the reduction \( x(z).\overline{y}[w] | (\overline{x}[y] + \overline{y}[y]) \rightarrow \overline{y}[w] \).

In practice, the reduction step is simulated as three distinct strategies. The first strategy, \( R_\pi \), takes care of the actual reduction \( (\rightarrow) \), while the two other strategies, \( R_{gc}, R_m \), functions as “house-cleaning”, ensuring that the transformed graphs represent process encodings.

The strategy \( R_\pi \) consists of two rules, one of which is depicted in Fig. 4. The rule in Fig. 4 simulates the reduction \((x(y).P + M) | (\overline{x}[y] + Q + R) \rightarrow P\{y/w\} | Q \). The match of the rule searches for the parallel composition on the outermost prefix operators, i.e., a vertex labelled \( t(p) \) attached to the root. Recall that every vertex corresponding to a parallel composition is followed by the sum composition vertex with the label \( t(s) \). One of these branches of the sum composition vertex is matched such that we find an input and output operator vertex both synchronizing on the same name.

When a match is found, the reduction is then simulated by deleting the prefix operators and marking their continuations (which are always followed by a parallel composition vertex) to be coalesced into the “top” parallel composition operator. Additionally, the name \( y \) is now bound to \( w \), which is simulated by merging the corresponding vertices.

An example of applying \( R_\pi \) to an encoded process is shown in Fig. 3 as the direct derivation \([P] \Rightarrow G'\), where \([P]\) describes the process defined in the start of this section. Note, that the rules are constructed such that when applying a rule any resulting connected component not containing the root vertex will never be part of the corresponding process encoding and hence such components are not depicted here. The input operator in \( P \) can synchronize on the output operator in both branches of \( P \), and similarly there exists two matches of the rule depicted in Fig. 4 when applied to \([P]\). Either match of the rule, however,
A Graph-Based Tool to Embed the π-Calculus.

Figure 3: The application of each strategy, $\mathcal{R}_\pi$, $\mathcal{R}_m$, $\mathcal{R}_{gc}$, to simulate the reduction $P \rightarrow Q$.

leads to a graph isomorphic to $G'$. The graph $G'$ does not represent an actual encoding of a process. The reason for this is two-fold: First, we have simply marked vertices that must be coalesced, but not actually coalesced them yet, due to the DPO approach in $\texttt{M\&D}$ only supporting injective morphisms. Second, the encoding of any branch not chosen, in this case one of the output operators of $P$, must be deleted. The coalescing of vertices is done by $\mathcal{R}_m$, obtaining the graph $G''$ from $G'$ in Fig. 3 while the deletion of non-chosen summation branches is done by $\mathcal{R}_{gc}$, finally resulting in the graph $[Q]$ obtained from $G''$ in Fig. 3. For more details on $\mathcal{R}_m$ and $\mathcal{R}_{gc}$ see Companion Technical Report [6].

It was shown in [12] that for the resulting graph $[Q]$ there exists some process $Q$ with the encoding $[Q]$. Let $[P] \xrightarrow{\mathcal{R}_\pi} [Q]$ be the sequence of derivations obtained by applying $\mathcal{R}_\pi$, $\mathcal{R}_m$, and $\mathcal{R}_{gc}$ on $[P]$ as illustrated in Fig. 3 transforming the encoded process $[P]$ into some encoded process $[Q]$. Then, it was additionally shown that given two processes $P$ and $Q$, the reduction $P \rightarrow Q$ exists iff $[P] \xrightarrow{\mathcal{R}_\pi} [Q]$ exists.

The actual derivation graph obtained from computing every possible transition $[P] \xrightarrow{\mathcal{R}_\pi} [Q]$, involves many direct derivations, as is evident from Fig. 3 that does not correspond to any specific state of a process, but are necessary implementation details to transition from one process state to another.

Equipped with the derivations illustrated in Fig. 3, however, we can identify all direct derivations between actual graphs corresponding to processes. We can
Figure 4: The rule for synchronization in $R_\pi$ when channel and argument differ, represented as the span $L \leftarrow K \rightarrow R$. Terms starting with an underscore represents variables, e.g., the match in $L$ is searching for a name “v(_X)”, where the actual name is unspecified and contained in “_X”. The variable can then be used in $R$. Matched vertices and edges coloured blue in $L$ will be modified by the rule while vertices and edges coloured black will be in the context $K$ of the rule. We see, e.g., that the vertex with the term “go” remains unchanged, while the vertex with the term “t(in)” in $L$ is deleted in $R$. Vertices whose terms will be relabelled will be coloured purple in the context. We see that vertices with the term “t(s)” in $L$ will be renamed to “gc” in $R$, represented in the context with “⟨t(s), gc⟩”. Finally, vertices marked for coalescing will be wrapped in a term called “merge” and connected to the vertex they should be coalesced into. This is seen by the renaming of the name “v(_Z)” to “merge(v(_Z))” and connecting it to “v(_Y)”, representing the name “_Y” being bound to “_Z” in the simulated reduction.

then use such direct derivations to create a new abbreviated derivation graph, representing our execution space, where any sequence of derivations $[P] \xrightarrow{R_\pi} [Q]$ is modelled as an edge. The result is an execution space as illustrated in Fig. 2 as first introduced in the start of this section.

5 A Final Example

Suppose we want to model the following behaviour:
- a patient stumbles into a hospital with two doctors; Dr. Jekyll and Mr. Hyde;
- the patient will be treated by the first doctor the patient comes across;
- Dr. Jekyll will cure the patient while Mr. Hyde will kill the patient.

The behaviour can be simulated with the expression $Hospital = P | J | H$, for the process $\phi$ defined by:

$$P = \phi \text{stumble}[name].name(d).P' \quad P' = \phi \text{kill}(x) + \text{cure}(x).P$$

$$J = \phi \text{stumble}(n).\pi[j \text{ekyll}].\pi[j \text{cly}].P$$

$$H = \phi \text{stumble}(n).\pi[\text{hyde}].\pi[h \text{ill}][h]$$
We refer to the Companion Technical Report [6] for details about the encoding of recursive processes and how to compute the execution spaces including recursive processes.

The arguments are not given for the recursive processes, as they are just the set of free names of the respective process. A patient walks into the hospital and gives their name to either Jekyll or Hyde via the stumble channel. The doctor then sends his own name back to the patient as a greeting via the patient’s name as a channel (note, this serves no practical purpose for the model). Finally, the patient is either cured or killed, depending on the doctor received. If cured, \( P \) is called recursively simulating that the patient can get sick and stumble into the hospital again. If the patient died, we reached a deadlock as there is no cure for death. Note, many of the arguments sent over the channels are not used, but are simply there for illustrative purposes.

We can model the above example in EpiM as follows:

```plaintext
s, n, pn, j, h, cu, ki, d, x = names("s n pn j h cu ki d x")
free_names = [s, n, ki, cu]

rp = RecursiveProcess()
P = Process().output(s, n).input(n, d).call("Pp", free_names)
rp.add("P", free_names, P)

Pp = Process().input(ki, x) + Process().input(cu, x).call("P", free_names)
rp.add("Pp", free_names, Pp)

J = Process().input(s, pn).output(pn, j).output(cu, j).call("J", free_names)
rp.add("J", free_names, J)

H = Process().input(s, pn).output(pn, h).output(ki, h).call("H", free_names)
rp.add("H", free_names, H)

Hospital = P | J | H
exec_space = ReductionDG(Hospital, rp)
exec_space.calc()
exec_space.print()
```

Note, we have shortened the names to simple single letter names for illustrative purposes. The resulting execution space, representing the simulation of \( Hospital \), is illustrated in Fig. 5. Although, it might be difficult to parse the graph, the expected behaviour should be clear from the graph: if the patient is treated by Dr. Jekyll, \( P \) is called recursively and as a result the behaviour is modelled as the only cycle in the derivation graph. If on the other hand the patient was killed off then \( P \) is terminated. Hence the behaviour is modelled in the graph as a path that end in a vertex with no outgoing edges (reached a deadlock).

6 Conclusion

We have introduced EpiM, available as a web front-end at [http://cheminf.imada.sdu.dk/epim](http://cheminf.imada.sdu.dk/epim); a tool using graph transformation for computing execution spaces of \( \pi \)-calculus processes. In this regard we presented a brief overview
of the encodings and transformations involved. Practically, execution spaces can be used for basic analysis of processes such as determining liveness. Since Epim is directly embedded into \texttt{MOD}, it would be possible to use the range of features provided by the \texttt{MOD} framework, notably \[5\]. Since structural congruence can be determined directly from graph isomorphism of the encoded processes, canonicization of processes is given for free. \texttt{MOD} provides a framework for stochastic simulations \[4\], suggesting the possibilities of using a version of the encoding of \(\pi\)-calculus presented here to simulate stochastic \(\pi\)-calculus processes \[9\].

The type of DPO supported by \texttt{MOD} does not allow rewriting rules that perform bulk duplication or deletion of subgraphs. This means that duplicating or deleting processes cannot be implemented with a single rewrite operation. On one hand, this elicits the cost of process duplication or recursion often ignored in process calculi, on the other, it suggests to explore the use of \texttt{MOD} with “resource aware calculi” like linear variations of the \(\pi\)-calculus \[8\][15][16].

**Acknowledgements.** This work was partially supported by the Independent Research Fund Denmark, grant no. DFF-7014-00041.
References


