A Process Algebra for Optimization for Parallel Programs

Ichiro Satoh

Department of Information Sciences, Ochanomizu University 2-1-1 Otsuka Bunkyo-ku
Tokyo 112-8610 Japan
Tel: +81-3-5978-5388 Fax: +81-3-5978-5390
E-mail: ichiro@is.ocha.ac.jp

Abstract

We propose a theoretical framework for the performance analysis and optimization of parallel programs through an algebraic relation on expressions in a time-extended process calculus. The relation is an extension of bisimulation and is characterized by having the ability to order behaviorally equivalent communicating processes with respect to their relative speeds, and have some useful algebraic properties. This abstract outlines the calculus and the relation and gives some views on the framework.

1 Introduction

In parallel computing systems, including distributed ones, interactions among processes, such as communication and synchronization strongly affect the performance of these systems. Therefore, to minimize the overheads of such interactions and to optimize the timings of them are important and necessary. Therefore, there have been many optimization techniques for communication and synchronization. However, unfortunately most of them often lack qualitative and quantitative investigation about their effectiveness and validation. Moreover, parallel and distributed computation is far more complex than sequential one. We need a theoretical framework to reason about optimization techniques for parallel computation.

On the other hand, over the last few years, many researchers have explored time-extended process calculi for example see [3, 4, 9, 13]. The calculi provides widely studied frameworks for modeling and verifying real-time systems. Such theories typically consist of a simple language with time-dependent behaviors that have a well-defined operational semantics given in terms of labeled transition systems. They also have time-sensitive equivalence relations that are used to relate implementations and specifications, which are both given as terms in the language. However, the relations equate two real-time processes when both their functionally behavioral properties and their temporal properties are completely matched with each other. However, the main goal of the relations is to provide the verification of real time systems, instead of analyzing optimization techniques for parallel computation.

This extended abstract outlines our basic ideas in constructing a theoretical framework for qualitatively and quantitatively verifying timing optimization for communications among processes. The framework is formulated through a new process calculus with an algebraic order relation over processes. The relation order behaviorally equivalent processes with respect to their relative speeds.

We here present the organization of this extended abstract: 1 In the next section we briefly present our basic ideas concerning the process calculi and then define their syntax and semantics. In Section 3 we define a speed-sensitive order relation on communicating processes and study their basic properties. In Section 4 we compare with related work and give some concluding remarks.

2 Description Language

This section constructs a process calculus for reasoning about temporal costs in computing and communication, for example execution time and synchronization time. The calculus is defined by incorporating many common features of existing process calculi and is characterized by having the ability to express time.

Before defining the syntax of the calculus, we give the basic idea on time and the notations of time values. We assume that time is interval between events instead of any absolute time, and is continuous, instead of discrete time.

Definition 2.1 Let $T$ denote the set of the positive real numbers including 0.

---

1This abstract summarizes our previous work presented in [11] but we plan to present some notable results of our current work at the meeting.
Time values are often denoted as positive real numbers including zero.
We define symbols to present the events of processes.

**Definition 2.2**

- Let $\mathcal{A}$ be an infinite set of names denoting communication actions. Its elements are denoted as $a, b, \ldots$
- Let $\mathcal{A}$ be an infinite set of co-names. Its elements are denoted as $\bar{a}, \bar{b}, \ldots$
- Let $\mathcal{L} \equiv \mathcal{A} \cup \mathcal{A}$ be a set of communication action names. Elements of the set are written as $\ell, \ell', \ldots$
- Let $\tau$ denote an internal action.
- Let $P$ be the set of actions corresponding the amount of the passage of time. Elements of the set are denoted as $\langle t_1 \rangle, \langle t_2 \rangle, \ldots$, where $t_1, t_2, \ldots \in \mathbb{T}$.
- Let $\mathcal{Act} \equiv \mathcal{L} \cup \{ \tau \}$ be the set of operational actions. Its elements are denoted as $\alpha, \beta, \ldots$

$\pi$ is the complementary action of $a$. $\tau$-action represents all handshake communications and is considered to be unobservable from outside environments.

In the calculus, we describe communications of parallel programs by means of the language defined below. The syntax of the calculus is coincide with all the construction of some existing process calculi, for example CCS[6], except for a new prefix operator whose contents are dependent on the passage of time, called *delay operator*. This operator suspends a process for a specified period, written as $\langle t \rangle$, where $t$ is the amount of the suspension. For instance, $\langle t \rangle . P$ means a process which is idle for $t$ time units and then behaves as $P$. On the other hand, to preserve the pleasant properties of the original process calculi, all communication and internal actions are assumed to be instantaneous. Instead, we introduce the new operator for expressing temporal costs of computation and communication, for example execution time and communication latency.

**Definition 2.3** The set $\mathcal{P}$ of TSCS expressions ranged over by $P, P_1, P_2, \ldots$ is defined recursively by the following abstract syntax:

\[
\begin{align*}
P &::= 0 \quad \text{(Terminate Process)} \\
&| \; \alpha P \quad \text{(Action Prefix)} \\
&| P_1 + P_2 \quad \text{(Summation)} \\
&| P_1 | P_2 \quad \text{(Composition)} \\
&| P \setminus L \quad \text{(Action Restriction)} \\
&| \langle t \rangle . P \quad \text{(Delay Prefix)} \\
&| A \mathcal{=} P \quad \text{(Recursive Definition)}
\end{align*}
\]

where $t$ is an element of $\mathbb{T}$ and $L$ is a subset of $\mathcal{L}$. $A$ is a process variable in set $\mathcal{K}$. We assume that in $A \mathcal{=} P$, $P$ is always closed, and each occurrence of $A$ in $P$ is only within some subexpressions $\alpha A$ where $\alpha$ is not empty, or $\langle t \rangle . A$ where $t > 0$. $\square$

The informal meaning of each process constructor is as follows:

- $0$ is a terminate process that can perform no internal nor communication action.
- $\alpha . P$ is a process to perform action $\alpha$ and then behaves like $P$, where $\alpha$ is an input, output, or internal action.
- $P_1 + P_2$ represents a process which may behave as either $P_1$ or $P_2$.
- $P_1 | P_2$ represents that process $P_1$ and $P_2$ may run in parallel.
- $P \setminus L$ behaves like $P$ but it is prohibited to communicate with external processes at actions in $L \cup \bar{L}$.
- $A \mathcal{=} P$ means that $A$ is defined as $P$, where $P$ may include $A$.
- $\langle t \rangle . P$ represents a process which is suspended for $t$ time units and then behaves like $P$.

We need the notion of action sort later.

**Definition 2.4** The syntactic sort of each process, $\mathcal{L}(P)$, is defined inductively by:

\[
\begin{align*}
\mathcal{L}(0) &= \emptyset \\
\mathcal{L}(\alpha P) &= \{ \alpha \} \cup \mathcal{L}(P) \\
\mathcal{L}(\bar{\alpha} P) &= \{ \bar{\alpha} \} \cup \mathcal{L}(P) \\
\mathcal{L}(\tau P) &= \mathcal{L}(P) \\
\mathcal{L}(P_1 + P_2) &= \mathcal{L}(P_1) \cup \mathcal{L}(P_2) \\
\mathcal{L}(P_1 | P_2) &= \mathcal{L}(P_1) \cup \mathcal{L}(P_2) \\
\mathcal{L}(P \setminus L) &= \mathcal{L}(P) - (L \cup \bar{L}) \\
\mathcal{L}(\langle t \rangle . P) &= \mathcal{L}(P)
\end{align*}
\]

when $A \mathcal{=} P$, we have $\mathcal{L}(P) \subseteq \mathcal{L}(A)$. We often call sort simply. $\square$

Next, we give the semantics of the calculus. However, before doing this, we here should explain our basic ideas on the passage of time. Time passes in all processes at the same speed. Also, all processes follow the same clock, or different ones but well-synchronized clocks.

The operational semantics of the calculus can computationally encompass that of CCS and embody the notion of time. The semantics is defined as two tiers of labeled transition rules. One of them defines the semantics of functional behaviors of processes, called *behavioral transition*, written as $\alpha \rightarrow (\cdots \subseteq \mathcal{P} \times \mathcal{Act} \times \mathcal{P})$ and the other defines the passage of time on processes, called *temporal transition*, written as $\langle t \rangle \rightarrow (\cdots \subseteq \mathcal{P} \times \mathcal{P} \times \mathcal{P})$. The
advance of time is modeled as the latter transition. It is labeled by a quantity of time to indicate the amount of the advance, for example \( P \xrightarrow{(t)} P' \). It means that process \( P \) become \( P' \) after \( t \) time units.

**Definition 2.5** The calculus is a labeled transition system \( \langle \mathcal{P}, \text{Act} \cup \mathcal{I}, \{ \xrightarrow{\alpha} \subseteq \mathcal{P} \times \mathcal{E} \mid \alpha \in \text{Act} \cup \mathcal{I} \} \rangle \). The transition relation \( \xrightarrow{\alpha} \) is defined by two kinds of structural induction rules given in Figure 1 and 2.

In giving the rules, we adopt the convention that the transition below the horizontal line may be inferred from the transitions above the line.

\[
\begin{align*}
\alpha.P & \xrightarrow{\alpha} P' & P \xrightarrow{\alpha} P', \alpha \in \text{Act}, L \subseteq \mathcal{P} \\
R_1 & \xrightarrow{\alpha} R_1' & \ell.P \xrightarrow{\alpha} \ell.P \\
P_1 + P_2 & \xrightarrow{\alpha} P_1' + P_2' & P_1 \parallel P_2 \xrightarrow{\alpha} P_1' \parallel P_2' \\
P_1 \mid P_2 & \xrightarrow{\alpha} P_1' \mid P_2' & P_1 \parallel P_2 \xrightarrow{\alpha} P_1' \parallel P_2' \\
\mathcal{P} \xrightarrow{\alpha} \mathcal{P}' (\alpha \in \mathcal{I}) & \mathcal{P} \xrightarrow{\alpha} \mathcal{P}'
\end{align*}
\]

Figure 1: Inference Rules for Behavioral Transition

\[
\begin{align*}
0 & \xrightarrow{(t)} 0 & 0 & \xrightarrow{(t)} 0 \\
\ell.P & \xrightarrow{(t)} \ell.P & \ell.P & \xrightarrow{(t)} \ell.P \\
P_1 & \xrightarrow{(t)} P_1' & P_1 \xrightarrow{(t)} P_1' \\
P_1 + P_2 & \xrightarrow{(t)} P_1' + P_2' & P_1 \parallel P_2 \xrightarrow{(t)} P_1' \parallel P_2' \\
P_1 | P_2 & \xrightarrow{(t)} P_1' | P_2' & P_1 \parallel P_2 \xrightarrow{(t)} P_1' \parallel P_2' \\
\mathcal{P} \xrightarrow{(t)} \mathcal{P}' & \mathcal{P} \xrightarrow{(t)} \mathcal{P}' \\
\mathcal{P} \parallel L & \xrightarrow{(t)} \mathcal{P} \parallel L & \mathcal{P} \parallel L & \xrightarrow{(t)} \mathcal{P} \parallel L \\
\mathcal{P} \parallel \mathcal{P}' & \xrightarrow{(t)} \mathcal{P}' & \mathcal{P} \parallel \mathcal{P}' & \xrightarrow{(t)} \mathcal{P}' \\
(t + \ell').E & \xrightarrow{(t + \ell')} (t + \ell').E & (t + \ell') E & \xrightarrow{(t + \ell')} (t + \ell').E
\end{align*}
\]

Figure 2: Inference Rules for Temporal Transition

We briefly explain some important transitions shown in Definition 2.5. The semantics assumes that when an internal or communication action is enabled, processes must perform the action immediately without imposing unnecessary idling. This assumption is the same as the notion of maximal progress shown in [4, 13]. It lets us exactly measure necessary time for synchronization among parallel processes, and enables the calculus to preserve the observation properties of many existing non-timed process calculus.

**Example 2.6** We show some basic examples of processes in \( \mathcal{P} \), as follows:

1. \( \pi.P_1 \) is a process which performs output action \( \pi \) after 2 time units and then behaves like \( P_1 \).
2. \( \pi.P_1 \xrightarrow{(2)} \pi.P_1 \xrightarrow{(t)} P_1 \)
3. \( (a. P_1 + b. P_3) \) is a process which can receive either input action \( a \) or \( b \) after 3 time units, and then behaves like \( P_2 \) or \( P_3 \).
4. \( (a. P_1 + b. P_3) \) performs a communicate between \( \pi.P_1 \) and \( (a. P_2 + b. P_3) \) at action name \( a \).

\[
\begin{align*}
(2). \pi.P_1 & \xrightarrow{(2)} \pi.P_1 \xrightarrow{(t)} P_1 \\
(3). (a. P_2 + b. P_3) & \xrightarrow{(3)} \pi.P_1 \parallel (a. P_2 + b. P_3) \\
& \xrightarrow{(t)} P_1 \parallel P_3 
\end{align*}
\]

The transition relation \( \xrightarrow{\text{def}} \) does not distinguish between observable and unobservable actions. We define two transition relations due to the non-observationality of \( \tau \).

**Definition 2.7**

1. \( P \xrightarrow{(t)} P' \) is defined as \( P \xrightarrow{\tau} P' \).
2. \( P \xrightarrow{\tau \rightarrow} P' \) is defined as \( P \xrightarrow{\tau} P' \).
3. \( P \xrightarrow{(t_1)} P' \xrightarrow{(t_1)} P' \) is defined as \( P \xrightarrow{\tau} P' \xrightarrow{(t_1)} \ldots \xrightarrow{(t_n)} P' \) where \( + \) is a mathematical addition over two numbers.

In the following section, we present an algebraic inequality over process expressions in the calculus. However, in order to give a rational theory, we need to impose some certain syntactic restrictions on processes.

**Definition 2.8**

1. \( (\alpha_1 \cdots \alpha_n). P \) is defined as \( \sum_{i \leq n} \alpha_i \cdot P \), where \( \alpha_i.P \equiv \alpha P \). \( (\alpha_1 \cdots \alpha_n). P \) is called confluent summation.
(2) $P_1 \parallel_L P_2$ is defined as $(P_1 | P_2) \setminus L$. $P_1 \parallel_L P_2$ is confluent composition if $\mathcal{L}(P_1) \cap \mathcal{L}(P_1) = \emptyset$ and $\mathcal{L}(P_1) \cap \mathcal{L}(P_1) \subseteq L \cup \mathcal{T}$.

\[ \square \]

**Definition 2.9** $P \in \mathcal{P}$ is time-stable if $P$ is built using only terminate process, action prefix, delay prefix, action restriction, confluent composition, confluent summation, and recursion given in Definition 2.3.

Interprocess communication in real communicating systems is often realized by means of asynchronous style as well as synchronous one. The word *asynchronous* here means that sender processes can send messages without synchronizing any processes. However, most of existing process calculi, including this calculus are formulated based on synchronous communication. However, the calculus can essentially have the expressive power of asynchronous communication. A way to express asynchronous communication is to restrict the formation of a term, $\pi.P$, in the original calculus to the case where $P$ is a terminate process, written as $0$. That is, an asynchronous output, $\pi$, followed by a process, $P$, is the same as the parallel composition $\pi.0|P$. We leave this extension to another paper [11].

### 3 Speed-Sensitive Prebisimulation

Based on time-extended process calculi, several researchers have explored time-sensitive equivalence relations that are based on trace equivalence, failure equivalence, testing equivalence, and bisimulation equivalence like ours, for example see [4, 9, 13]. These equivalence relations equate two processes if they cannot be distinguished from each other in their temporal properties as well as their behavioral one. However, the relations may often be too strict in the analysis of most time-dependent systems, including non-strict real-time systems. This is because most systems have various temporal uncertainties, for example unpredictable transmission delays in communication, and unexpected interruptions in processors. Therefore, the temporal properties of implementations in the real world are never the same as those of their specification exactly. Also, we can often say that an implementation is able to satisfy its specification, only when the implementation can perform the behavioral properties given in its specification at earlier timings than those given in the specification. It is convenient to construct a framework that can decide whether two processes can perform the same behaviors and whether one of them (e.g. an implementation of a system) can perform the behaviors faster than the other (e.g. the specification of the system).

This section develops such an algebraic order relation on processes with respect to their speeds based on the bisimulation concept.

**Definition 3.1** A binary relation $\mathcal{R} \subseteq (\mathcal{P} \times \mathcal{P}) \times T$ is a $t$-prebisimulation over communicating processes if $(P_1, P_2) \in \mathcal{R}_t$ $(t \geq 0)$ implies, for all $\alpha \in \text{Act}$:

(i) $\forall d \forall P_1'$: $P_1 \xrightarrow{(d)} \alpha P_1' \text{ then } \exists d \exists P_2'$:

$P_2 \xrightarrow{(d)} \alpha P_2' \text{ and } (P_1', P_2') \in \mathcal{R}_{t+d}$

(ii) $\forall d \forall P_1'$:

$P_1 \xrightarrow{(d)} (t) \alpha P_1' \text{ then } \exists P_1'$:

$P_1 \xrightarrow{(d)} (t) \alpha P_1' \text{ and } (P_1', P_2') \in \mathcal{R}_0$

In the above definition, $\mathcal{R}_t$ is a family of relations indexed by a non-negative time value $t$. Intuitively, $t$ is the relative difference between the time of $P_1$ and that of $P_2$; that is, it means that $P_1$ precedes $P_2$ by $t$ time units. The following order relation starts with a prebisimulation indexed by $t$ (i.e., $\mathcal{R}^t$) and can change $t$ as the bisimulation proceeds only if $t \geq 0$.

**Definition 3.2** We let $P_1 \leq^t P_2$ if there exists some $t$-prebisimulation such that $(P_1, P_2) \in \mathcal{R}_t$. We call $\leq^t$ speed-sensitive order on communicating processes. We shall often abbreviate $\leq^0$ as $\leq$.

Hereafter, we usually consider $\leq$ only. We show several algebraic properties of the order relation below.

**Proposition 3.3** Let $P, P_1, P_2, P_3 \in \mathcal{P}$. Then,

1. $P \leq P$
2. $P_1 \leq P_2 \text{ and } P_2 \leq P_3 \text{ then } P_1 \leq P_3$

From these results, we see that $\leq$ is a preorder relation. We also have $P_1 \leq^t \leq^t P_2$ if $P_1 \leq^t P_2$ and $P_2 \leq^t P_3$.

**Proposition 3.4** Let $P_1, P_2 \in \mathcal{P}, t_1, t_2 \in \mathcal{T}$ such that $t_1 \leq t_2$. Then,

$(t_1).P \leq (t_2).P$

The above proposition shows an important characteristic of $\leq$.

**Example 3.5** We show some basic examples of $\leq$ as follows:

1. $a.P \leq (1).a.P$

This means that the performance of $P_1$ is at most $t$ time units faster than that of $P_2$.
(2) \( (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow P \leq (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow P \leq (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow P \leq P_2 \).

\( \text{Proposition 3.6} \) Let \( P_1, P_2, Q \in P \) such that

\( (1) \alpha, P_1 \leq \alpha, P_2 \)
\( (2) P_1 \setminus L \leq P_2 \setminus L \)
\( (3) \langle \rangle P_1 \leq \langle \rangle P_2 \)

It is convenient to develop a precongruence with respect to speeds in order to guarantee the substitutability between two ordered processes in any parallel context. However, there is an undesirable problem in defining such a pre-congruence with temporal inequality. Suppose three objects: \( A_1 \equiv (\pi_0 \pi_0) \), \( A_2 \equiv (\pi_0 \pi_0) \), and \( B \equiv (\pi_0 \pi_0) \). We clearly have \( A_1 \leq A_2 \) but cannot expect that

\( A_1 \otimes B \geq A_2 \otimes B \), because \( A_1 \otimes (\pi_0 \pi_0) \equiv (\pi_0 \pi_0) \) and \( A_2 \otimes (\pi_0 \pi_0) \equiv (\pi_0 \pi_0) \). This anomaly is traced to contexts that restrict the capability to execute a particular computation due to the passage of time, for example \( \text{timeout} \) handling in \( B \). However, when we restricted processes to be included in the time-stable process set given in the previous section, they can perform executable actions in any order and thus the order relation is preserved in parallel context.

\( \text{Proposition 3.7} \) Let \( P_1, P_2, Q \in P \) be time-stable processes. Then,

\[ P_1 \leq P_2 \quad \text{then} \quad P_1 \otimes Q \leq P_2 \otimes Q \]

In order to prove the above result, we need some lemmas, including a fact that any time-stable processes are confluent. However, for lack of space, we leave its detail proof to another paper.

Intuitively, the above result tells that a parallel composition between the faster processes can really perform faster than one between the slower ones. That is, a system when embedding the faster processes can still perform faster than when embedding the slower ones.

4 Discussion

Related Work

We briefly survey related work. There have indeed been many process calculi for reasoning about temporal properties of communicating systems, for example see [3, 4, 7, 9, 12, 13]. Most of the calculi have been equipped with time-sensitive equivalence relations as verification methods. However, only a few of them intend to analyze and compare temporal costs of communicating processes, for example [5, 7, 10, 12].

Among them, Moller and Toft in [7] proposed a preorder relation over processes with respect to their relative speeds, based on the bisimulation technique. Unlike ours, their calculus assumes to permit an executable communication to be suspended for arbitrary periods of time. As a result, the relation shows only that a process may possibly execute faster than the other. Recently, Vogler in [12] and Jenner and Vogler in [5] presented speed-sensitive preorder relations based on testing equivalence. The relation can relate asynchronously communicating processes according to their relative speeds, but its semantics is formulated based on causality between events on the assumption that actions are not instantaneous, unlike ours. Also, some researchers have explored the performance analysis by means of process algebra, for example see [2]. However, most of them are based on non-instantaneous actions. The other assumes that every process proceeds in lockstep and at every instant performs a single action. Arun-Kumar and Hennessy in [1] Natarajan and Cleaveland [8] propose approaches to relate processes with respect to their relative efficiencies according to the number of necessary internal actions, \( \tau \)-actions through the same communication. However, it is very difficult to reflect the execution cost of real systems upon the number of \( \tau \)-actions exactly in the description of the systems.

Concluding Remarks

This abstract outlines a theoretical framework for the performance analysis and optimization of communicating processes, based on the process calculus and its algebraic theories. It gives only a starting point for formulating such a framework. There are many issues that we leave in this abstract. This paper studied a speed-sensitive order relation for end-to-end synchronous communication. However, in real communicating systems, interprocess communication is often realized by means of asynchronous style as well as synchronous one. In asynchronous communication settings, the sender of a message cannot know when the message is actually consumed as opposed to synchronous ones. We are interested in formulating a speed-sensitive order relation for asynchronously communicating processes. In synchronous communication settings, processes must be blocked until their partner processes are ready to communicate. The order relation presented in this paper can order two syn-
chronously communicating processes when a conceptual observer cannot distinguish between them in their communications, and when the timings of the communications with one of them are earlier than those with the other. On the other hand, in asynchronous communication settings, the observer cannot exactly know when the messages that it sends are received by processes. This difference between synchrony and asynchrony in communication means that a suitable speed-sensitive order relation corresponding to asynchronous communication is needed. The relation has to be able to know only the arrival timings of return messages. An observer sends arbitrary messages to processes and waits for return messages from them. It orders the two processes when the return messages cannot be distinguished from each other, and when the arrival timings of the messages from one of them are earlier than those from the other. The relation can reveal essential differences between synchrony and asynchrony in interactions among processes in time-sensitive contexts.

References


---

\footnote{You can some results on such a speed-sensitive order relation for asynchronously communicating processes in [11].}