

# Indirect Jumps Improve Instruction Sequence Performance

J.A. BERGSTRA<sup>1</sup>, C.A. MIDDELBURG<sup>1</sup>

## Abstract

Instruction sequences with direct and indirect jump instructions are as expressive as instruction sequences with direct jump instructions only. We show that, in the case where the number of instructions is not bounded, we are faced with increases of the maximal internal delays of instruction sequences on execution that are not bounded by a linear function if we strive for acceptable increases of the lengths of instruction sequences on elimination of indirect jump instructions.

**Keywords:** instruction sequence performance, indirect jump instruction, maximal internal delay

## 1 Introduction

Although instruction sequences with direct and indirect jump instructions are as expressive as instruction sequences with direct jump instructions only (see [2]), indirect jump instructions are widely used to implement certain features of contemporary high-level programming languages such as the switch statements and virtual method calls of Java [7] and C# [8]. Therefore, we consider a further analysis of indirect jump instructions relevant.

In this paper, we study the effect of eliminating indirect jump instructions from instruction sequences with direct and indirect jump instructions on the performance of instruction sequences with respect to the interaction with their environment on execution. It is implicit that the elimination of indirect jump instructions must preserve the behaviour of the instruction

---

<sup>1</sup>Informatics Institute, Faculty of Science, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, the Netherlands. Email: {J.A.Bergstra,C.A.Middelburg}@uva.nl.

sequence concerned on execution. We show that, in the case where the number of instructions is not bounded, there exist instruction sequences with direct and indirect jump instructions from which elimination of indirect jump instructions is possible without a super-linear increase of their maximal internal delay on execution only if their lengths increase super-linearly on elimination of indirect jump instructions.<sup>2</sup>

The work presented in this paper belongs to a line of research whose working hypothesis is that instruction sequence is a central notion of computer science. In this line of research, program algebra [1] is the setting used for investigating instruction sequences. The starting-point of program algebra is the perception of a program as a single-pass instruction sequence, i.e. a finite or infinite sequence of instructions of which each instruction is executed at most once and can be dropped after it has been executed or jumped over. This perception is simple, appealing, and links up with practice.

The program notation used in this paper to show that indirect jumps improve instruction sequence performance is  $\text{PGLB}_{ij}$ . This program notation is a minor variant of  $\text{PGLC}_{ij}$ , a program notation with indirect jumps instructions introduced in [2]. Both program notations are rooted in program algebra, are closer to existing assembly languages than the notation provided by program algebra, and have relative jump instructions. The main difference between them is that  $\text{PGLB}_{ij}$  has an explicit termination instruction and  $\text{PGLC}_{ij}$  has not. This difference makes the former program notation more convenient for the purpose of this paper.

The performance measure used in this paper is the maximal internal delay of an instruction sequence on execution. The maximal internal delay of an instruction sequence on execution is the largest possible delay that can take place between successively executed instructions whose effects are observable externally. Another conceivable performance measure is the largest possible sum of such delays on execution of the instruction sequence. In this paper, we do not consider the latter performance measure because it looks to be less adequate to the interactive performance of instruction sequences.<sup>3</sup>

The work presented in this paper can be looked at in a wider context. In the literature on computer architecture, hardly anything can be found that contributes to a sound understanding of the effects of the presence

---

<sup>2</sup>A super-linear increase means an increase that is not bounded by a linear function.

<sup>3</sup>Interactive performance means performance with respect to the interaction with the environment on execution.

of common kinds of instructions in the instruction set of a computer on important points such as instruction sequence size and instruction sequence performance. The work presented in this paper can be considered a first step towards such an understanding.

This paper is organized as follows. First, we give a survey of the program notation  $\text{PGLB}_{ij}$  (Section 2). Next, we introduce the notion of maximal internal delay of a  $\text{PGLB}_{ij}$  program (Section 3). After that, we present the above-mentioned result concerning the elimination of indirect jump instructions (Section 4). We also relate the work presented in this paper to the point of view which is the origin of the line of research to which it belongs (Section 5). Finally, we make some concluding remarks (Section 6).

## 2 PGLB with Indirect Jumps

In this section, we give a survey of the program notation  $\text{PGLB}_{ij}$ . This program notation is a variant of the program notation  $\text{PGLB}$ , which belongs to a hierarchy of program notations rooted in program algebra (see [1]).  $\text{PGLB}$  and  $\text{PGLB}_{ij}$  are closer to existing assembly languages than the notation provided by program algebra.

It is assumed that fixed but arbitrary numbers  $I$  and  $N$  have been given, which denote the number of registers available and the greatest natural number that can be contained in a register. Moreover, it is also assumed that fixed but arbitrary finite sets  $\mathcal{F}$  of *foci* and  $\mathcal{M}$  of *methods* have been given.

The set  $\mathfrak{A}$  of *basic instructions* is  $\{f.m \mid f \in \mathcal{F}, m \in \mathcal{M}\}$ . The view is that the execution environment of a  $\text{PGLB}_{ij}$  program provides a number of services, that each focus plays the role of a name of a service, that each method plays the role of a command that a service can be requested to process, and that the execution of a basic instruction  $f.m$  amounts to making a request to the service named  $f$  to process command  $m$ . The intuition is that the processing of the command  $m$  may modify the state of the service named  $f$  and that the service in question will produce  $\mathsf{T}$  or  $\mathsf{F}$  at its completion.

For example, a service may be able to process methods for setting the content of a Boolean cell to  $\mathsf{T}$  or  $\mathsf{F}$  and a method for getting the content of the Boolean cell. Processing of a setting method may modify the state of the service, because it may change the content of the Boolean cell with which

the service deals, and simply produces the final content at its completion. Processing of the getting method does not modify the state of the service, because it does not change the content of the Boolean cell with which the service deals, and produces the content of the Boolean cell at its completion.

PGLB<sub>ij</sub> has the following primitive instructions:

- for each  $a \in \mathfrak{A}$ , a *plain basic instruction*  $a$ ;
- for each  $a \in \mathfrak{A}$ , a *positive test instruction*  $+a$ ;
- for each  $a \in \mathfrak{A}$ , a *negative test instruction*  $-a$ ;
- for each  $l \in \mathbb{N}$ , a *direct forward jump instruction*  $\#l$ ;
- for each  $l \in \mathbb{N}$ , a *direct backward jump instruction*  $\backslash\#l$ ;
- for each  $i \in [1, I]$  and  $n \in [1, N]$ , a *register set instruction*  $\text{set}:i:n$ ;
- for each  $i \in [1, I]$ , an *indirect forward jump instruction*  $i\#i$ ;
- for each  $i \in [1, I]$ , an *indirect backward jump instruction*  $i\backslash\#i$ ;
- a *termination instruction*  $!$ .

PGLB<sub>ij</sub> programs are expressions of the form  $u_1 ; \dots ; u_k$ , where  $u_1, \dots, u_k$  are primitive instructions of PGLB<sub>ij</sub>. PGLB programs are PGLB<sub>ij</sub> programs in which register set instructions, indirect forward jump instructions and indirect backward jump instructions do not occur.

On execution of a PGLB<sub>ij</sub> program, the primitive instructions of PGLB<sub>ij</sub> have the following effects:

- the effect of a positive test instruction  $+a$  is that basic instruction  $a$  is executed and execution proceeds with the next primitive instruction if  $\top$  is produced and otherwise the next primitive instruction is skipped and execution proceeds with the primitive instruction following the skipped one – if there is no primitive instruction to proceed with, inaction occurs;
- the effect of a negative test instruction  $-a$  is the same as the effect of  $+a$ , but with the role of the value produced reversed;
- the effect of a plain basic instruction  $a$  is the same as the effect of  $+a$ , but execution always proceeds as if  $\top$  is produced;

- the effect of a direct forward jump instruction  $\#l$  is that execution proceeds with the  $l$ -th next instruction of the program concerned – if  $l$  equals 0 or there is no primitive instruction to proceed with, inaction occurs;
- the effect of a direct backward jump instruction  $\backslash\#l$  is that execution proceeds with the  $l$ -th previous instruction of the program concerned – if  $l$  equals 0 or there is no primitive instruction to proceed with, inaction occurs;
- the effect of a register set instruction  $\text{set}:i:n$  is that the content of register  $i$  is set to  $n$  and execution proceeds with the next primitive instruction – if there is no primitive instruction to proceed with, inaction occurs;
- the effect of an indirect forward jump instruction  $i\#i$  is the same as the effect of  $\#l$ , where  $l$  is the content of register  $i$ ;
- the effect of an indirect backward jump instruction  $i\backslash\#i$  is the same as the effect of  $\backslash\#l$ , where  $l$  is the content of register  $i$ ;
- the effect of the termination instruction  $!$  is that execution terminates.

If execution proceeds unbroken and forever with no other primitive instructions than jump instructions and register set instructions, this is identified with inaction.

$\text{PGLB}_{ij}$  is a minor variant of  $\text{PGLC}_{ij}$ , a program notation with indirect jumps instructions introduced in [2]. The differences between  $\text{PGLB}_{ij}$  and  $\text{PGLC}_{ij}$  are the following:

- in those cases where inaction occurs on execution of  $\text{PGLB}_{ij}$  programs because there is no primitive instructions to proceed with, termination takes place on execution of  $\text{PGLC}_{ij}$  programs;
- the termination instruction  $!$  is not available in  $\text{PGLC}_{ij}$ .

The meaning of  $\text{PGLC}_{ij}$  programs is formally described in [2] by means of a mapping of  $\text{PGLC}_{ij}$  programs to closed terms of program algebra. In that way, the behaviour of  $\text{PGLC}_{ij}$  programs on execution is described indirectly: the behaviour of the programs denoted by closed terms of program algebra on execution is formally described in several papers, including [2], using

basic thread algebra [1].<sup>4</sup> Because  $\text{PGLB}_{ij}$  is a minor variant of  $\text{PGLC}_{ij}$ , we refrain from describing the behaviour of  $\text{PGLB}_{ij}$  programs on execution formally in the current paper.

### 3 Internal Delays of $\text{PGLB}_{ij}$ Programs

In this section, we will define the notion of maximal internal delay of a  $\text{PGLB}_{ij}$  program.

It is assumed that a fixed but arbitrary set  $\mathfrak{X} \subset \mathfrak{A}$  of *auxiliary basic instructions* has been given. The view is that, in common with the effect of jump instructions and register set instructions, the effect of auxiliary basic instructions is wholly unobservable externally, but contributes to the realization of externally observable behaviour. Typical examples of auxiliary basic instructions are basic instructions for storing and fetching data of a temporary character. Typical examples of non-auxiliary basic instructions are basic instructions for reading input data from a keyboard, showing output data on a screen and writing data of a permanent character on a disk.

The maximal internal delay of a  $\text{PGLB}_{ij}$  program concerns the delay that takes place between successive non-auxiliary basic instructions on execution of the instruction sequence. Before we define the maximal internal delay of a  $\text{PGLB}_{ij}$  program, we define the execution traces of a  $\text{PGLB}_{ij}$  program.

Let  $u_1 ; \dots ; u_k$  be a  $\text{PGLB}_{ij}$  program. Then  $tr(\rho, j, u_1 ; \dots ; u_k)$ , where  $\rho : [1, I] \rightarrow [1, N]$  and  $j \in \mathbb{N}$ , is the set of all finite sequences over the set of all primitive instructions of  $\text{PGLB}_{ij}$  that may be encountered successively on execution of  $u_1 ; \dots ; u_k$  if execution starts with  $u_j$  with the registers used for indirect jumps set according to  $\rho$ .

Formally, for each  $\text{PGLB}_{ij}$  program  $u_1 ; \dots ; u_k$ , the different sets  $tr(\rho, j, u_1 ; \dots ; u_k)$  are defined by simultaneous induction on the structure of the finite sequences over the set of all primitive instructions of  $\text{PGLB}_{ij}$  by the following clauses:

1.  $\langle \rangle \in tr(\rho, j, u_1 ; \dots ; u_k)$ ;
2. if  $u_j \equiv a$  or  $u_j \equiv +a$  or  $u_j \equiv -a$ , and  $\sigma \in tr(\rho, j+1, u_1 ; \dots ; u_k)$ , then  $u_j\sigma \in tr(\rho, j, u_1 ; \dots ; u_k)$ ;
3. if  $u_j \equiv +a$  or  $u_j \equiv -a$ , and  $\sigma \in tr(\rho, j+2, u_1 ; \dots ; u_k)$ , then  $u_j\sigma \in tr(\rho, j, u_1 ; \dots ; u_k)$ ;

---

<sup>4</sup>In several early papers, basic thread algebra is presented under the name basic polarized process algebra.

4. if  $u_j \equiv \#l$  and  $\sigma \in tr(\rho, j+l, u_1; \dots; u_k)$ , then  $u_j\sigma \in tr(\rho, j, u_1; \dots; u_k)$ ;
5. if  $u_j \equiv \backslash\#l$  and  $\sigma \in tr(\rho, j \dot{-} l, u_1; \dots; u_k)$ , then  $u_j\sigma \in tr(\rho, j, u_1; \dots; u_k)$ ;<sup>5</sup>
6. if  $u_j \equiv \text{set}:i:n$  and  $\sigma \in tr(\rho', j+1, u_1; \dots; u_k)$ , then  $u_j\sigma \in tr(\rho, j, u_1; \dots; u_k)$ , where  $\rho'$  is defined by  $\rho'(i') = \rho(i')$  if  $i' \neq i$  and  $\rho'(i) = n$ ;
7. if  $u_j \equiv i\#i$  and  $\sigma \in tr(\rho, j + \rho(i), u_1; \dots; u_k)$ , then  $u_j\sigma \in tr(\rho, j, u_1; \dots; u_k)$ ;
8. if  $u_j \equiv i\backslash\#i$  and  $\sigma \in tr(\rho, j \dot{-} \rho(i), u_1; \dots; u_k)$ , then  $u_j\sigma \in tr(\rho, j, u_1; \dots; u_k)$ ;
9. if  $u_j \equiv !$ , then  $u_j \in tr(\rho, j, u_1; \dots; u_k)$ .

For example,

$$tr(\rho_0, 1, +a; \#3; \text{set}:1:3; \#2; \text{set}:1:1; i\#1; b; \#2; c; !),$$

where  $\rho_0$  is defined by  $\rho_0(i) = 0$  for all  $i \in [1, I]$ , contains

$$\begin{aligned} &+a \ \#3 \ \text{set}:1:1 \ i\#1 \ b \ \#2 \ !, \\ &+a \ \text{set}:1:3 \ \#2 \ i\#1 \ c \ !, \end{aligned}$$

and all prefixes of these two sequences, including the empty sequence.

The set of *execution traces* of a PGLB<sub>ij</sub> program  $P$ , written  $tr(P)$ , is  $tr(\rho_0, 1, P)$ , where  $\rho_0$  is defined by  $\rho_0(i) = 0$  for all  $i \in [1, I]$ .

The *maximal internal delay* of an PGLB<sub>ij</sub> program  $P$ , written  $MID(P)$ , is the largest  $n \in \mathbb{N}$  for which there exists an execution trace  $u_1 \dots u_k \in tr(P)$  and  $i_1, i_2 \in [1, k]$  with  $i_1 \leq i_2$  such that  $ID(u_j) \neq 0$  for all  $j \in [i_1, i_2]$  and  $\sum_{j \in [i_1, i_2]} ID(u_j) = n$ , where  $ID(u)$  is defined as follows:

$$\begin{aligned} ID(a) &= 0 & \text{if } a \notin \mathfrak{X}, & & ID(\#l) &= 1, \\ ID(a) &= 1 & \text{if } a \in \mathfrak{X}, & & ID(\backslash\#l) &= 1, \\ ID(+a) &= 0 & \text{if } a \notin \mathfrak{X}, & & ID(\text{set}:i:n) &= 1, \\ ID(+a) &= 1 & \text{if } a \in \mathfrak{X}, & & ID(i\#i) &= 2, \\ ID(-a) &= 0 & \text{if } a \notin \mathfrak{X}, & & ID(i\backslash\#i) &= 2, \\ ID(-a) &= 1 & \text{if } a \in \mathfrak{X}, & & ID(!) &= 0. \end{aligned}$$

<sup>5</sup>As usual, we write  $i \dot{-} j$  for the monus of  $i$  and  $j$ , i.e.  $i \dot{-} j = i - j$  if  $i \geq j$  and  $i \dot{-} j = 0$  otherwise.

Suppose that in the example given above  $a$ ,  $b$  and  $c$  are non-auxiliary basic instructions. Then

$$MID(+a; \#3; \text{set}:1:3; \#2; \text{set}:1:1; i\#1; b; \#2; c; !) = 4.$$

This delay takes place between the execution of  $a$  and the execution of  $b$  or  $c$ .

In [6], an extension of basic thread algebra is proposed which allows for internal delays to be described and analysed. We could give a formal description of the behaviour of PGLB<sub>ij</sub> programs on execution, internal delays included, using this extension of basic thread algebra. Founded on such a formal description, we could explain why the notion of maximal internal delay of a PGLB<sub>ij</sub> program defined above is the right one.

$MID(P)$  can be looked upon as the number of time units that elapse during the largest possible internal delay of  $P$ . Because the time that it takes to execute one basic instruction is taken for the time unit in the definition of  $MID(P)$ , it can alternatively be looked upon as the number of basic instructions that can be executed during the largest possible internal delay of  $P$ . As usual, the time that it takes to execute one basic instruction is called a *step*.

## 4 Indirect Jumps and Instruction Sequence Performance

In this section, we show that indirect jump instructions are needed for good instruction sequence performance.

It is assumed that  $\text{bool}:1, \text{bool}:2, \dots \in \mathcal{F}$  and  $\text{set}:T, \text{set}:F, \text{get} \in \mathcal{M}$ . The foci  $\text{bool}:1, \text{bool}:2, \dots$  serve as names of services that act as Boolean cells. The methods  $\text{set}:T, \text{set}:F$ , and  $\text{get}$  are accepted by services that act as Boolean cells and their processing by such a service goes as follows:

- $\text{set}:T$ : the contents of the Boolean cell is set to T, and T is produced;
- $\text{set}:F$ : the contents of the Boolean cell is set to F, and F is produced;
- $\text{get}$ : nothing is changed, but the contents of the Boolean cell is produced.

On execution of a PGLB<sub>ij</sub> program that interacts with a Boolean cell, the content of the Boolean cell may change at any time by external causes,

i.e. by causes that do not originate from the executed program. The use operators introduced in [5] permit encapsulation of services, as a result of which changes by external causes are excluded. On purpose, we do not apply these use operators here: the Boolean cell with which we are concerned serves as a means for a program to interact with the environment on execution.

Below, we will write  $\bullet_{i=1}^n P'_i$ , where  $P'_1, \dots, P'_n$  are PGLB<sub>ij</sub> programs, for the PGLB<sub>ij</sub> program  $P'_1; \dots; P'_n$ .

For each  $k \in \mathbb{N}$ , let  $P_k$  be the following PGLB<sub>ij</sub> program:

$$\begin{aligned} & \bullet_{i=1}^{2^k} (-\text{bool}:1.\text{get}; \#3; \text{set}:1:2 \cdot i + 1; \#(2^k - i) \cdot 4 + 2); !; \\ & \bullet_{i=1}^{2^k} (-\text{bool}:1.\text{get}; \#3; \text{set}:2:2 \cdot i + 1; \#(2^k - i) \cdot 4 + 2); !; \\ & i\#1; \bullet_{i=1}^{2^k} (a_i; \#(2^k - i) \cdot 2 + 1); i\#2; \bullet_{i=1}^{2^k} (a'_i; !), \end{aligned}$$

where  $a_1, \dots, a_{2^k}$  and  $a'_1, \dots, a'_{2^k}$  are mutually different basic instructions.

First,  $P_k$  repeatedly tests the Boolean cell `bool:1`. If T is not returned for  $2^k$  tests,  $P_k$  terminates. Otherwise, in case it takes  $i$  tests until T is returned, the content of register 1 is set to  $2 \cdot i + 1$ . If  $P_k$  has not yet terminated, it once again repeatedly tests the Boolean cell `bool:1`. If T is not returned for  $2^k$  tests,  $P_k$  terminates. Otherwise, in case it takes  $j$  tests until T is returned, the content of register 2 is set to  $2 \cdot j + 1$ . If  $P_k$  has not yet terminated, it performs  $a_i$  after an indirect jump and following this  $a'_j$  after another indirect jump. After that,  $P_k$  terminates. The length of  $P_k$  is  $12 \cdot 2^k + 4$  primitive instructions and the maximal internal delay of  $P_k$  is 4 steps.

The PGLB<sub>ij</sub> programs  $P_1, P_2, \dots$  defined above will be used in the proof of the result concerning the elimination of indirect jump instructions stated below.

In the proof concerned, we will further make use of the notion of a reachable occurrence of a basic instruction in a PGLB<sub>ij</sub> program. We assume that this notion is intuitively sufficiently clear to grasp the proof. Given a formal description of the behaviour of PGLB<sub>ij</sub> programs in the style used for PGLC<sub>ij</sub> programs in [2], reachability can easily be defined in the way followed in Section 3.3 of [9].

Below, we will write  $\ell(P)$ , where  $P$  is a PGLB<sub>ij</sub> program, for the length of  $P$ , i.e. its number of primitive instructions.

A mapping  $\varphi$  from the set of all PGLB<sub>ij</sub> programs to the set of all PGLB programs has a *linear upper bound on the increase in maximal internal delay* if for some  $c', c'' \in \mathbb{N}$ , for all PGLB<sub>ij</sub> programs  $P$ ,  $MID(\varphi(P)) \leq c' \cdot MID(P) + c''$ .

A mapping  $\varphi$  from a subset  $\mathcal{P}$  of the set of all PGLB<sub>ij</sub> programs to the set of all PGLB programs has a *quadratic lower bound on the increase in length* if for some  $c', c'' \in \mathbb{N}$  with  $c' \neq 0$ , for all  $P \in \mathcal{P}$ ,  $\ell(\varphi(P)) \geq c' \cdot \ell(P)^2 + c''$ .

It follows easily from the behaviour-preserving mapping from the set of all PGLC<sub>ij</sub> programs to the set of all PGLC programs given in [2] and an idea used in Section 5.1 of [5] that there exists a behaviour-preserving mapping from the set of all PGLB<sub>ij</sub> programs to the set of all PGLB programs with a linear upper bound on the increase in maximal internal delay.<sup>6</sup>

**Theorem 1** *Suppose  $\varphi$  is a behaviour-preserving mapping from the set of all PGLB<sub>ij</sub> programs to the set of all PGLB programs with a linear upper bound on the increase in maximal internal delay. Moreover, suppose that the number of basic instructions is not bounded. Then there exists a set  $\mathcal{P}$  of PGLB<sub>ij</sub> programs such that the restriction of  $\varphi$  to  $\mathcal{P}$  has a quadratic lower bound on its increase in length.*

**Proof:** For each  $k \in \mathbb{N}$ , let  $P_k$  be defined as above. We show that the restriction of  $\varphi$  to  $\{P_1, P_2, \dots\}$  has a quadratic lower bound on its increase in length. Take an arbitrary  $k \in \mathbb{N}$ . Because  $\varphi$  has a linear upper bound on the increase in maximal internal delay, we have  $MID(\varphi(P_k)) \leq c' \cdot MID(P_k) + c'' = c' \cdot 4 + c''$  for some  $c', c'' \in \mathbb{N}$ . Let  $c = c' \cdot 4 + c''$ . Suppose that  $k$  is much greater than  $c$ . This supposition requires that the number of basic instructions is not bounded. Because  $\varphi$  is a behaviour-preserving mapping and there is a possibility that  $\varphi(P_k)$  contains auxiliary basic instructions, there are at most  $2^c$  different occurrences of basic instructions in  $\varphi(P_k)$  reachable in  $c$  steps from any occurrence of a basic instruction in  $\varphi(P_k)$ . Let  $i \in [1, 2^k]$ . Then, because  $\varphi$  is a behaviour-preserving mapping, for each  $j \in [1, 2^k]$ , some occurrence of  $a'_j$  in  $\varphi(P_k)$  is reachable in  $c$  steps from each occurrence of  $a_i$  in  $\varphi(P_k)$ . Thus, there are at least  $2^k$  different occurrences of basic instructions to reach in  $c$  steps from one occurrence of  $a_i$ , although at most  $2^c$  occurrences of basic instructions are reachable in  $c$  steps. Therefore, there must be at least  $2^k / 2^c = 2^{k-c}$  different occurrences of  $a_i$  in  $\varphi(P_k)$ . Consequently,  $\ell(\varphi(P_k)) \geq 2^k \cdot 2^{k-c} = 2^{2 \cdot k - c}$ . Moreover,  $\ell(P_k) = 12 \cdot 2^k + 4$ . Hence, the restriction of  $\varphi$  to  $\{P_1, P_2, \dots\}$  has a quadratic lower bound on its increase in length.  $\square$

We conclude from Theorem 1 that we are faced with super-linear increases of maximal internal delays if we strive for acceptable increases

<sup>6</sup>Here behaviour-preserving means preserving the behaviour referred to at the end of Section 2, which behaviour does not include internal delays.

of program lengths on elimination of indirect jump instructions. In other words, indirect jump instructions are needed for good instruction sequence performance.

## 5 Projectionism

Semantically, we can eliminate indirect jump instructions by means of a behaviour-preserving mapping from  $\text{PGLB}_{ij}$  programs to  $\text{PGLB}$  programs, but we meet here two challenges of a point of view concerning programming language semantics of which such behaviour-preserving mappings form part.

In [3], the name projectionism is coined for the point of view referred to above and its main challenges are discussed. To put it briefly, projectionism is the point of view that:

- any program  $P$  first and for all represents a single-pass instruction sequence as considered in program algebra;
- this single-pass instruction sequence, found by behaviour-preserving mappings called projections, represents in a natural and preferred way what is supposed to take place on execution of  $P$ ;
- program algebra provides the preferred notation for single-pass instruction sequences.

Among the main challenges of projectionism identified in [3] are explosion of size, degradation of performance, complexity of description, and aesthetic degradation.

Projectionism is illustrated in [1] by means of a hierarchy of program notations rooted in program algebra. For each of the program notations that appear in the hierarchy, except program algebra, a mapping is given by means of which each program from that program notation is translated into a program from the first program notation lower in the hierarchy that produces the same behaviour on execution. Thus, the single-pass instruction sequence represented by a program in any program notation from the hierarchy can be found. The mappings concerned are examples of the projections referred to above in the outline of projectionism.

The behaviour-preserving mapping from Theorem 1 is a projection in the same sense. Theorem 1 concerns two of the above-mentioned challenges of projectionism, for it expresses that in the case of a projection from  $\text{PGLB}_{ij}$  programs to  $\text{PGLB}$  programs there is a trade-off between explosion of size and degradation of performance.

## 6 Conclusion

In the literature on computer architecture, hardly anything can be found that contributes to a sound understanding of the effects of the presence of common kinds of instructions in the instruction set of a computer on points such as instruction sequence size and instruction sequence performance. As a first step towards such an understanding, we have shown that, in the case where the number of instructions is not bounded, there exist instruction sequences with direct and indirect jump instructions from which elimination of indirect jump instructions is possible without a super-linear increase of their maximal internal delay on execution only if their lengths increase super-linearly on elimination of indirect jump instructions. It is an open problem whether this result goes through in the case where the number of instructions is bounded.

Instruction sequences with direct jump instructions, indirect jump instructions and register transfer instructions are as expressive as instruction sequences with direct jump instructions and indirect jump instructions without register transfer instructions. We surmise that a mapping that eliminates the register transfer instructions yields a result comparable to Theorem 1. However, we have not yet been able to provide a proof for that case. On the face of it, a proof for that case is much more difficult than the proof for the case treated in this paper.

For completeness, we mention that, in the line of research to which the work presented in this paper belongs, work that is mainly concerned with direct jump instructions includes the work presented in [4].

## Acknowledgements

We thank two anonymous referees for carefully reading a preliminary version of this paper and for suggesting improvements of the presentation of the paper.

## References

- [1] J. A. Bergstra and M. E. Loots. Program algebra for sequential code. *Journal of Logic and Algebraic Programming*, 51(2):125–156, 2002. doi: [10.1016/S1567-8326\(02\)00018-8](https://doi.org/10.1016/S1567-8326(02)00018-8).

- [2] J. A. Bergstra and C. A. Middelburg. Instruction sequences with indirect jumps. *Scientific Annals of Computer Science*, 17:19–46, 2007.
- [3] J. A. Bergstra and C. A. Middelburg. Instruction sequence notations with probabilistic instructions. [arXiv:0906.3083v1](https://arxiv.org/abs/0906.3083v1) [cs.PL], June 2009.
- [4] J. A. Bergstra and C. A. Middelburg. On the expressiveness of single-pass instruction sequences. *Theory of Computing Systems*, 50(2):313–328, 2012. [doi:10.1007/s00224-010-9301-8](https://doi.org/10.1007/s00224-010-9301-8).
- [5] J. A. Bergstra and A. Ponse. Combining programs and state machines. *Journal of Logic and Algebraic Programming*, 51(2):175–192, 2002. [doi:10.1016/S1567-8326\(02\)00020-6](https://doi.org/10.1016/S1567-8326(02)00020-6).
- [6] J. A. Bergstra and M. B. van der Zwaag. Mechanistic behavior of single-pass instruction sequences. [arXiv:0809.4635v1](https://arxiv.org/abs/0809.4635v1) [cs.PL], September 2008.
- [7] J. Gosling, B. Joy, G. Steele, and G. Bracha. *The Java Language Specification*. Addison-Wesley, Reading, MA, second edition, 2000.
- [8] A. Hejlsberg, S. Wiltamuth, and P. Golde. *C# Language Specification*. Addison-Wesley, Reading, MA, 2003.
- [9] S. H. P. Schroevers. Expressiveness and extensions of an instruction sequence semigroup. [arXiv:1003.1572v1](https://arxiv.org/abs/1003.1572v1) [cs.PL], March 2010.