# The Algebra of Finite State Processes

Peter Michael Sewell

Doctor of Philosophy
University of Edinburgh
1995

Graduation date: November 1995

# Abstract


This thesis is concerned with the algebraic theory of finite state processes. The processes we focus on are those given by a signature with prefix, summation and recursion, considered modulo strong bisimulation. We investigate their equational and implicational theories.

We first consider the existence of finite equational axiomatisations. In order to express an interesting class of equational axioms we embed the processes into a simply typed lambda calculus, allowing equation schemes with metasubstitutions to be expressed by pure equations. Two equivalences over the lambda terms are defined, an extensional equality and a higher order bisimulation. Under a restriction to first order variables these are shown to coincide and an examination of the coincidence shows that no finite equational axiomatisation of strong bisimulation can exist. We then encode the processes of Basic Process Algebra with iteration and zero ($\mathrm{BPA}^*_\delta$) into this lambda calculus and show that it too is not finitely equationally axiomatisable, in sharp contrast to the extant positive result for the fragment without zero.

For the implicational theory, we show the existence of finite computable complete sets of unifiers for finite sets of equations between processes (with zero order variables). It follows that the soundness of sequents over these is decidable.

Some applications to the theories of higher order process calculi and non-well-founded sets are made explicit.

# Acknowledgements

I would like to thank my supervisor, Robin Milner, for his teaching in the practice and purpose of research, and for his patience and enthusiasm while this thesis mutated into its current form. Stuart Anderson provided intellectual support at an important time and also a period of educational employment. Section 2.3 was improved by discussions with Zoltán Ésik.

The Laboratory for the Foundations of Computer Science has provided a broad and stimulating environment in which to work. I thank particularly the members of the $\pi$ and concurrency clubs.

This thesis was written with the aid of a voice recognition system, largely funded by the SERC. Thanks are also due to the Department and particularly to John Butler for providing the underlying hardware. I have been supported by SERC studentship 90311819 and the ESPRIT BRA 6454 (CONFER). Paul Taylor's diagram and proof tree macro packages were used.

I have enjoyed my time in Scotland. This is partly due to the companionship of my long-suffering office mates Chen, Stephen and Luc, and Alex, Anthony, David, Dilip, Ian, John, Neil, Roberto, Savi and Steve.

Unfailing support has been provided by my parents and family.

# Declaration

I declare that this thesis was composed by myself and that the work presented is my own except where otherwise stated. The main results of Chapter 3 and §4.1 have been reported in [Sew94].

Peter Sewell

This is a revised version, incorporating the suggestions of my examiners, Colin Stirling and David Walker.

# Table of Contents

# Chapter 1

# Introduction

Finite state machines have been the subject of a great deal of work in theoretical computer science, particularly by the language theory and process algebra communities. In their various formalisations they are the basis for models or specifications of many computational phenomena. A common formalisation is the labelled transition system consisting of a (finite) set equipped with an indexed family of binary relations over it. Typically the set is thought of as the possible states that a modelled system may be in, with the relations as the allowable changes of state. This is extremely general — for a particular application additional structure (such as a termination predicate on states) may be required and the model may be too fine, containing many states which should be identified. Many equivalences have been proposed, differing in their treatment of nondeterministic choice and termination. We shall largely be concerned with two, language equivalence and bisimulation, that are reasonably canonical among them. The introduction of an equivalence immediately raises a question of decidability. These are both decidable over finite state systems and indeed efficient decision procedures have been widely used in practice.

Direct presentations of finite state machines as sets and relations are awkward to work with. Accordingly, syntactic forms have been introduced to represent them, including a variety of process calculi and regular expressions. Any such raise questions of expressiveness, congruence and axiomatisability (all with respect to a chosen equivalence). The first two are generally straightforward. It is desirable for the syntax to be expressively complete, i.e. for all finite state machines to be representable up to the equivalence. If any manipulation of syntactic terms, such as equational reasoning, is to be done then the equivalence must also be a congruence and preferably also substitutive. There are several motivations for seeking axiomatisability results. The most obvious is that any sound system may be useful for human or machine manipulation of terms, particularly but not necessarily if it is complete. A number of nontrivial verification problems arising from practice can be addressed using such systems. For this thesis a more important motivation is that axiomatisability results (and especially their proofs) shed light on the nature of the equivalences involved and on their interaction with the syntax of axioms. If positive, axiomatisability results enable the semantic maps to be presented as the unique maps from term algebras, which is sometimes convenient. Lastly they permit a comparison of different equivalences and with the alternative view that takes a set of axioms as primary.

Given a syntax that admits substitution and a choice of equivalence there is a further natural question of the existence of solutions to (or unifiers for) equations between terms. The existence of finite computable complete sets of unifiers gives the decidability of implications between equations.

We will largely be concerned with a single equivalence — strong bisimulation — over the terms of a simple syntax — the $\mu$-expressions, with prefix, summation and a binding operator for recursion. We consider the existence of finite axiomatisations and of finite computable complete sets of unifiers.

The existence of a finite axiomatisation obviously depends upon the strength of the metalanguage in which axioms are written. The weakest interesting choice is to consider equational axioms. For a syntax with binding there are further choices of the entities that variables in axioms may range over. We embed the $\mu$-expressions into a fragment of a simply typed lambda calculus, allowing many non-trivial properties to be expressed as equations. Our first main result is that no finite set of these equations can be a complete axiomatisation. There are then almost immediate non-finite-axiomatisability results for bisimulation of certain regular expressions and for equality of certain non-well-founded set expressions. A key tool is a notion of higher order bisimulation which we relate to the normal bisimulation of the higher order $\pi$ calculus.

Our second main result is the construction of finite complete sets of unifiers for finite sets of equations between $\mu$-expressions.

## 1.1 Overview

In Chapter 2 the basic definitions of syntax and semantics are introduced and the problems of axiomatisability that arise are discussed, firstly in general terms with reference to their motivation and then with reference to some previous work in the field.

In Chapter 3 our main non-axiomatisability result, that there is no finite equational axiomatisation for bisimulation over $\mu$-expressions, is formulated and proved.

Chapter 4 consists of a number of applications and developments of the theory of higher order bisimulation. We consider the axiomatisability of bisimulation over $*$-expressions, relate higher order bisimulation to the higher order $\pi$ calculus and cast our results into the language of non-well-founded set theory. We also give a definition of weak higher order congruence and show that weak congruence is finitely axiomatisable relative to bisimulation.

In Chapter 5 there is an investigation of the implicational theory of finite state processes (up to bisimulation). We show the existence of finite computable complete sets of unifiers for sets of equations between $\mu$-expressions and hence the decidability of implications.

There are only limited formal dependencies between chapters. The definitions in §2.1 are used throughout and the sections of Chapter 4 should be read after Chapter 3 (but are independent of each other).

## 1.2 Notation

The notation used is by and large standard, a few points are mentioned here for reference.

Theorems, propositions, lemmas and corollaries are numbered in two sequences, one of theorems and one of the others. The statements of a few lemmas are repeated in appendices with the original numbers. Subsidiary lemmas and definitions within proofs are sometimes indented.

If $\longrightarrow$ is any binary relation on a set we write $\longrightarrow^*$, $\longrightarrow^+$, $\longrightarrow^n$, $\longrightarrow^{\leq n}$ for its reflexive transitive closure, its transitive closure, the $n$-ary self composition and the union of $\longrightarrow^m$ for $m \leq n$. We often work with a family $\xrightarrow{a} \mid a \in Act$ of binary relations indexed by actions and write $\longrightarrow$ for $\cup_{a \in Act} \xrightarrow{a}$. Universal quantifiers over actions are freely omitted.

Substitutions are written postfix with $E[F/X]$ standing for $E$ in which all occurrences of $X$ are replaced by copies of $F$. The composition of substitutions $\rho$ and $\rho'$ is written as $\rho \circ \rho'$. The substitution which is as $\rho$ except where overridden by $\rho'$ is written $\rho \oplus \rho'$

Tuples are written with tildes, e.g. writing $\tilde{E}$ for a tuple $E_1 \ldots E_n$, or in angle brackets, e.g. $\langle A, B, C \rangle$. Concatenation and append are both written simply as juxtapositions.

The set of natural numbers is written $\omega$. We sometimes identify the number $n$ with the set $\{0, \ldots, n-1\}$.

# Chapter 2

# Background

In this chapter the basic definitions of syntax and semantics are introduced and the problems of axiomatisability that arise are discussed, firstly in general terms with reference to their motivation and then with reference to some previous work in the field.

## 2.1 Processes and Automata

The current applicability and history of the study of finite state machines are both too extensive to be treated here in any depth — a very brief historical sketch will be given. The basic notion is that of a finite discrete system together with an informal idea of its behaviour in discrete time. This is explicitly present in the work of Turing [Tur37] and of McCulloch and Pitts [MP43], although it may well have existed earlier. It is there applied respectively to the modelling of abstract (human) computation and concrete neural networks. In the first, machines (i.e. the finite state controls of Turing machines) are given by explicit descriptions of the sets of states and transition relations. In the second they are given structurally as networks of fundamental elements, with expressivity results relating structure and behaviour. A more syntactic description, the *regular*[1] or $*$-*expressions*, was introduced by Kleene [Kle56] and later simplified by Copi, Elgot and Wright [CEW58].

[1]We avoid the use of 'regular'. Most entities that we deal with could be qualified by it, to no useful end.

They have an explicit iteration operator together with sequential composition and summation. The paper of Kleene introduced in addition the study of the algebra of these expressions, giving some sound equations for his interpretation (which in the simplified formulation of [CEW58] is close to the standard language interpretation). Finer interpretations were apparently first introduced by Milner [Mil80] to give a more refined account of the interactions between a machine and an experimenter, leading to the *bisimulation equivalence* of Park [Par81]. To express all finite state machines up to bisimulation a richer syntax is required, such as the $\mu$*-expressions* which have variables and an explicit recursion operator together with prefix and summation. The algebra of the four combinations of syntax and model has been the subject of a large body of work, some of which is discussed in §2.3. We first give precise definitions of the two models, followed by the definitions of syntax and the basic results. Detailed references and proofs are omitted — the latter are all either straightforward or essentially from [Kle56] or [Mil84].

### 2.1.1 Models

The models are parameterised by a set $Act$ of atomic actions, ranged over by $a, b$. This is generally supposed to be nonempty but we require no other structure and do not require it to be infinite. The set of finite sequences of actions will be written $Act^*$ with the empty sequence as $\epsilon$ and concatenation as juxtaposition. Universal quantifiers over $Act$ will generally be omitted.

We start not with labelled transition systems but with charts. These contain additional information in the form of a *visibility predicate* $\triangleright$ between states and a set $Var$ of variables. This allows $\mu$-expressions containing free variables to be modelled and permits a close tie-up between bisimulation and language equivalence of $*$-expressions. The definitions are justified by Proposition 11 below relating them to the closed term case. When dealing with $\mu$-expressions we suppose $Var$ to be infinite, to permit alpha conversion.

**Definition** A *chart* $S$ is a tuple $\langle S, \longrightarrow, \triangleright, s_0 \rangle$ where $S$ is a set of states, $s_0 \in S$ is an initial state, $\longrightarrow \subseteq S \times Act \times S$ is a transition relation and $\triangleright \subseteq S \times Var$ is a visibility predicate. We will sometimes ignore the initial state.

The finest equivalence on charts that we deal with is bisimulation, at the top of the linear-branching time hierarchy. It takes full account of the nondeterministic branching structure of the transition relations.

**Definition** A relation $R \subseteq S \times S$ over the state set of a chart is a *bisimulation* if $s\ R\ t$ implies:

- If $s \xrightarrow{a} s'$ then $\exists t'\ .\ t \xrightarrow{a} t' \wedge s'\ R\ t'$.

- If $s \rhd X$ then $t \rhd X$.

and symmetrically. Two states are *bisimilar*, written $s \sim s'$, if there exists a bisimulation relating them.

**Proposition 1** *The relation $\sim$ is itself a bisimulation.*

At the bottom of the linear-branching time hierarchy are various forms of trace or language equivalence. The following definition allows the chart and language semantics of $*$-expressions to be closely related (by Proposition 12 below).

**Definition** The *extended language* of a state $s$ of a chart is the subset of $Act^* \times Var$ containing $a_1, \ldots, a_n, X$ iff there exist states $s_1, \ldots, s_n$ such that $s \xrightarrow{a_1} s_1 \ldots \xrightarrow{a_n} s_n \rhd X$. Two states are *language equivalent* if they have the same extended languages.

The trace congruence of [Rab93] will also be referred to.

**Definition** The *trace set* of a state $s$ is the subset of $Act^*$ containing $a_1, \ldots, a_n$ iff there exist states $s_1, \ldots, s_n$ such that $s \xrightarrow{a_1} s_1 \cdots \xrightarrow{a_n} s_n$. Two states are *trace congruent* if they have the same trace sets and the same extended languages.

In many applications it is necessary to abstract from certain actions which are considered to take place 'internally'. One equivalence that does this is the weak (or observational) congruence of [Mil89]. This is defined in terms of a distinguished action $\tau$ and the derived transition relations

$$\begin{array}{rcll} \stackrel{a}{\Longrightarrow} & \stackrel{\text{def}}{=} & \xrightarrow{\tau}{}^{*} \xrightarrow{a} \xrightarrow{\tau}{}^{*} & \\ \stackrel{\hat{a}}{\Longrightarrow} & \stackrel{\text{def}}{=} & \xrightarrow{\tau}{}^{*} \xrightarrow{a} \xrightarrow{\tau}{}^{*}, & \text{if } a \in Act - \tau \\ & \stackrel{\text{def}}{=} & \xrightarrow{\tau}{}^{*}, & \text{if } a = \tau. \end{array}$$

**Definition** A relation $R \subseteq S \times S$ over the state set of a chart is a *weak bisimulation* if $s\ R\ t$ implies:

- If $s \xrightarrow{a} s'$ then $\exists t'\ .\ t \stackrel{\hat{a}}{\Longrightarrow} t' \wedge s'\ R\ t'$.

- If $s \triangleright X$ then $t \xrightarrow{\tau}^{*} \triangleright X$.

and symmetrically. Two states are *weak bisimilar*, written $s \approx s'$, if there exists a bisimulation relating them. Two states are *weak congruent*, written $s \approx^{c} s'$, if there is a weak bisimulation $R$ with $s\ \mathcal{W}(R)\ t$, where $s\ \mathcal{W}(R)\ t$ if

- If $s \xrightarrow{a} s'$ then $\exists t'\ .\ t \stackrel{a}{\Longrightarrow} t' \wedge s'\ R\ t'$.

- If $s \triangleright X$ then $t \xrightarrow{\tau}^{*} \triangleright X$.

and symmetrically.

**Proposition 2** *Bisimulation is strictly finer than weak congruence and trace congruence. The latter is strictly finer than language equivalence.*

Bisimulation and the other equivalences are lifted to relations between charts by applying the above definitions to a disjoint union and considering the initial states.

### 2.1.2 Syntax: $\mu$-expressions

**Definition** The $\mu$-*expressions* are those of the grammar

$$E ::= 0 \mid X \mid aE \mid E + E \mid \mu X E$$

where $X$ and $a$ are drawn from sets $Var, Act$ of variables and actions and $\mu$ is a binding operator. We adopt standard notions of free and bound variables, substitution and alpha conversion. The scope of a binder is generally as far to the right as possible. Sum is taken to have lower precedence than prefix so $aE + F$ is $(aE) + (F)$.

We will have occasion to refer to the *infinite term equivalence* $=_{\text{infinite term}}$ induced by unwinding the recursions in $\mu$-expressions to give finite or infinite trees. A formal definition will not be given, however.

**Definition** Take the relations $\xrightarrow{a} \mid a \in Act$ and $\rhd$ to be the least over $\mu$-expressions such that

$$\frac{}{aE \xrightarrow{a} E} \qquad\qquad \frac{}{X \rhd X}$$

$$\frac{E \xrightarrow{a} E'}{E + F \xrightarrow{a} E'} \text{ and sym.} \qquad \frac{E \rhd X}{E + F \rhd X} \text{ and sym.}$$

$$\frac{E \xrightarrow{a} E'}{\mu Y E \xrightarrow{a} E'[\mu Y E \,/ Y]} \qquad \frac{E \rhd X \quad X \neq Y}{\mu Y E \rhd X}\,.$$

Any $\mu$-expression $E$ can thus be regarded as a chart, with initial state $E$ itself. The rule for $\mu$ differs from the more usual

$$\frac{E[\mu Y E \,/ Y] \xrightarrow{a} E'}{\mu Y E \xrightarrow{a} E'}$$

but is slightly more convenient. We check that it is equipotent. Let $\stackrel{a}{\rightsquigarrow}$ be the least relation over $\mu$-expressions satisfying the rules for prefix, sum and the second rule for $\mu$.

**Proposition 3** *The relations $\xrightarrow{a}$ and $\stackrel{a}{\rightsquigarrow}$ coincide.*

PROOF The following lemmas are used, both of which are provable by induction on derivations. Closely related properties will be heavily used later (see Lemmas 35 and 52).

> **Lemma 4** *If $E \stackrel{a}{\rightsquigarrow} F$ then $E[G/Y] \stackrel{a}{\rightsquigarrow} F[G/Y]$.*
>
> **Lemma 5** *If $E[F/Y] \xrightarrow{a} G$ then either $E \xrightarrow{a} E'$ and $E'[F/Y] = G$ or $E \rhd Y$ and $F \xrightarrow{a} G$.*

The proposition can now be proved by an induction on derivations for each direction. The only interesting cases are those in which the conclusion of a derivation is the relevant $\mu$ rule. Lemma 4 is required for the inclusion $\xrightarrow{a} \subseteq \stackrel{a}{\rightsquigarrow}$ and Lemma 5 for the converse. $\square$

The $\mu$-expressions suffice to express all finite charts up to bisimulation and therefore also up to language equivalence.

**Definition** A chart $\langle S, \longrightarrow, \rhd, s_0 \rangle$ is *finite* if $S$, $\longrightarrow$ and $\rhd$ are all finite.

**Proposition 6** *A chart is bisimilar to a finite chart iff it is bisimilar to some $\mu$-expression.*

The equivalences are congruences.

**Proposition 7** *Bisimulation of $\mu$-expressions is a congruence, i.e. it is an equivalence satisfying*

$$\frac{E \sim F}{aE \sim aF} \qquad \frac{E \sim E' \quad F \sim F'}{E + F \sim E' + F'} \qquad \frac{E \sim F}{\mu X E \sim \mu X F}$$

*and is substitutive, i.e. it satisfies*

$$\frac{E \sim F}{E[G/X] \sim F[G/X]}.$$

**Proposition 8** *Language equivalence of $\mu$-expressions is a substitutive congruence.*

**Proposition 9** *Trace congruence of $\mu$-expressions is a substitutive congruence.*

**Proposition 10** *Weak congruence of $\mu$-expressions is a substitutive congruence.*

In a given application one might well take as primary bisimulation over labelled transition systems or *closed* $\mu$-expressions. The extension to open terms given above can be justified as the coarsest reasonable such.

**Proposition 11** *Bisimulation of $\mu$-expressions is the largest congruence that, when restricted to the closed $\mu$-expressions, is at least as fine as bisimulation there.*

### 2.1.3 Syntax: $*$-expressions

Finite state systems have also been described using calculi with a unary or binary iteration operator in place of explicit recursion, such as the $*$-*expressions* given by

$$E ::= a \mid 0 \mid 1 \mid E + E \mid E \cdot E \mid E^* \mid E^\star E$$

where $a \in Act$. There are several points to note:

- Actions are now taken to be nullary and there is a sequential composition.

- Two iteration operators are included. A binary iteration $E^\star F$ represents zero or more iterations of $E$ followed by one of $F$. This was introduced in [Kle56] and simplified in [CEW58] to the ubiquitous unary iteration, with $E^*$ representing zero or more iterations of $E$. In the presence of $1$ they will be interdefinable in our semantics, with

$$E^\star F = E^* \cdot F \qquad \text{and} \qquad E^* = E^\star 1.$$

- Both a $0$ and a $1$ are included. In process calculus terms they will represent the unsuccessfully and successfully terminated processes respectively. Language theoretically, they will represent the empty language and the language containing only the empty word. In both cases they will be units and zeros for choice and sequential composition as below.

$$\begin{array}{rcccl} 0 + E & = & E & = & E + 0 \\ 1 \cdot E & = & E & = & E \cdot 1 \end{array} \qquad\qquad 0 \cdot E \;=\; 0$$

  Language theoretically we will also have $E \cdot 0 = 0$. In the presence of $0$ and $*$ (but not $0$ and $^\star$ alone) $1$ is definable with $1 = 0^*$. Calculi without $1$ have recently been studied in the setting of ACP, where the interaction of $1$ and the parallel operators may be subtle. Binary iteration has received renewed attention in these, for example in [BBP94].

We give chart semantics using an additional judgment $\surd$ to record successful termination. Formally we regard $\surd$ as a distinguished element of $Var$ and write $E\surd$ as shorthand for $E \rhd \surd$. The previous definitions of bisimulation and language can therefore be applied.

**Definition** Take the relations $\xrightarrow{a} \mid a \in Act$ and predicate $\surd$ to be the least over the $*$-expressions such that

$$\frac{}{a \xrightarrow{a} 1} \qquad \frac{}{1\surd}$$

$$\frac{E \xrightarrow{a} E'}{E+F \xrightarrow{a} E'} \text{ and sym.} \qquad \frac{E\surd}{E+F\surd} \text{ and sym.}$$

$$\frac{E \xrightarrow{a} E'}{E \cdot F \xrightarrow{a} E' \cdot F} \qquad \frac{E\surd \quad F \xrightarrow{a} F'}{E \cdot F \xrightarrow{a} F'} \qquad \frac{E\surd \quad F\surd}{E \cdot F\surd}$$

$$\frac{E \xrightarrow{a} E'}{E^* \xrightarrow{a} E' \cdot E^*} \qquad \frac{}{E^*\surd}$$

$$\frac{E \xrightarrow{a} E'}{E \star F \xrightarrow{a} E' \cdot (E \star F)} \qquad \frac{F \xrightarrow{a} F'}{E \star F \xrightarrow{a} F'} \qquad \frac{F\surd}{E \star F\surd}.$$

Note that there are no rules for $0$.

The $*$-expressions are expressively complete for language equivalence. We first recall the standard definition of the language denoted by a regular expression.

**Definition** The language denoted by a $*$-expression $E$ is $[\![E]\!]$ where

$$\begin{aligned}
[\![a]\!] &\stackrel{\text{def}}{=} \{a\} \\
[\![0]\!] &\stackrel{\text{def}}{=} \{\} \\
[\![1]\!] &\stackrel{\text{def}}{=} \{\epsilon\} \\
[\![E+F]\!] &\stackrel{\text{def}}{=} [\![E]\!] \cup [\![F]\!] \\
[\![E \cdot F]\!] &\stackrel{\text{def}}{=} \{\, st \mid s \in [\![E]\!] \wedge t \in [\![F]\!] \,\} \\
[\![E^*]\!] &\stackrel{\text{def}}{=} \{\, s_1 \ldots s_n \mid n \geq 0 \wedge \forall i \in 1..n \,.\, s_i \in [\![E]\!] \,\} \\
[\![E \star F]\!] &\stackrel{\text{def}}{=} \{\, s_1 \ldots s_n t \mid t \in [\![F]\!] \wedge n \geq 0 \wedge \forall i \in 1..n \,.\, s_i \in [\![E]\!] \,\}.
\end{aligned}$$

This factors through the chart semantics.

**Proposition 12** *A word $w$ is in the language denoted by a $*$-expression iff $w, \surd$ is in the extended language of the chart associated with the expression.*

**Proposition 13** *A language $L \subseteq Act^*$ is denotable by a $*$-expression iff $\{\, w, \surd \mid w \in L \,\}$ is the extended language of a finite chart with $Var = \{\surd\}$.*

Thus they are as expressive as the $\mu$-expressions, up to language equivalence, however they express fewer bisimulation classes:

**Proposition 14** *There exists a finite chart with $Var = \{\surd\}$ that is not bisimilar to any $*$-expression.*

Both equivalences are congruences.

**Proposition 15** *Bisimulation of $*$-expressions is a congruence, i.e. it is an equivalence satisfying*

$$\frac{E \sim E' \quad F \sim F'}{E + F \sim E' + F'} \qquad \frac{E \sim E' \quad F \sim F'}{E \cdot F \sim E' \cdot F'} \qquad \frac{E \sim F}{E^* \sim F^*} \qquad \frac{E \sim E' \quad F \sim F'}{E \star F \sim E' \star F'}$$

*and is substitutive (w.r.t. substitution of terms for actions), i.e. it satisfies*

$$\frac{E \sim F}{E[G/a] \sim F[G/a]}.$$

**Proposition 16** *Language equivalence of $*$-expressions is a substitutive congruence.*

Note that the definition of the $*$-expressions does not include variables — actions play a dual role, appearing both in the definitions of the equivalences and in the substitution rule of axiomatisations. This is standard in the literature. It is shown to be innocuous by the previous two propositions.

A variety of subcalculi have been discussed in the literature with differing notation. For reference we include a little table:

| $a$ | $0$ | $1$ | $+$ | $\cdot$ | $\_^*$ | $\_{}^\star\_$ | |
|---|---|---|---|---|---|---|---|
| $a$ | $\Lambda$ | | $\vee$ | $\cdot$ | $\_^*$ | | [CEW58] |
| $a$ | $0$ | $1$ | $+$ | $\cdot$ | $\_^*$ | | [Con71,Koz94] |
| $a$ | $\phi$ | | $+$ | $\cdot$ | $\_^*$ | | [Sal66] |
| $a$ | | $\epsilon$ | $+$ | $\cdot$ | | | BPA$^\epsilon$ as in [Mol89] |
| $a$ | | | $+$ | $\cdot$ | | $\_{}^*\_$ | BPA$^*$ as in [BBP94,FZ94] |
| $a$ | $\delta$ | | $+$ | $\cdot$ | | $\_{}^*\_$ | BPA$^*_\delta$ as in [BBP94,FZ94,Fok94] |

The cited work is variously concerned with algebras satisfying certain axioms or with particular models. We therefore need to state carefully exactly what the above correspondences are. For the first three lines the common expressions denote the same language in the standard interpretation, except that in [CEW58] $E^*$ does not necessarily contain the empty word. For the last three lines bisimulation as defined below agrees with the definitions in the cited work, as follows.

For terms of $\{1, a, +, \cdot\}$ bisimulation is that of [Mol89, §6.3.1] for BPA$^\epsilon$ (identifying $1$ and $\epsilon$). As discussed there it differs from the original BPA$^\epsilon$ semantics of [Vra86]. Our transition system differs also from the semantics of [BBP94,FZ94] for terms of $\{0, a, +, \cdot, {}^\star\}$. There the rules above involving $1, \_^*$ or $\surd$ are replaced by the following

$$\frac{}{a \xrightarrow{a} \surd}$$

$$\frac{E \xrightarrow{a} \surd}{E + F \xrightarrow{a} \surd} \text{ and sym.}$$

$$\frac{E \xrightarrow{a} \surd}{E \cdot F \xrightarrow{a} \surd}$$

$$\frac{E \xrightarrow{a} \surd}{E^\star F \xrightarrow{a} E^\star F} \qquad \frac{F \xrightarrow{a} \surd}{E^\star F \xrightarrow{a} \surd}$$

with the bisimulation $\leftrightarrow$ defined by replacing the condition

- If $E' \triangleright \surd$ then $F' \triangleright \surd$.

by

- If $E' \xrightarrow{a} \surd$ then $F' \xrightarrow{a} \surd$.

The bisimulations coincide, however.

**Proposition 17** *For terms of $\{0, a, +, \cdot, {}^\star\}$ bisimulation $\sim$ coincides with bisimulation $\leftrightarrow$ over BPA$^*_\delta$ (identifying $0$ and $\delta$) as defined in [BBP94].*

PROOF Let $P, Q$ range over terms of $\{0, a, +, \cdot, {}^\star\}$ and $E, F, G$ range over terms of $\{1, 0, a, +, \cdot, {}^\star\}$. If $=_1$ is the least equivalence over the latter such that $1 \cdot E =_1 E$ and

$$\frac{E =_1 F}{E \cdot G =_1 F \cdot G}$$

then it is straightforward to show that

$$\{ P, Q \mid \exists E, F \,.\, P =_1 E \sim F =_1 Q \}$$

is contained in $\leftrightarrow$ and that

$$\{ E, F \mid \exists P, Q \,.\, E =_1 P \leftrightarrow Q =_1 F \} \cup \{ E, F \mid E =_1 1 =_1 F \}$$

is contained in $\sim$. $\square$

## 2.2 Axiomatisation — some general remarks

The bulk of this thesis addresses questions of axiomatisability of equivalences over the $\mu$- and $*$-expressions. As it stands these questions are poorly defined — there is a range of possible definitions of what an axiomatisation is and the motivation of Chapter 1 does not provide technical criteria that pick out a unique one. In general the insight gained from an axiomatisation is inversely related to the expressiveness of the metalanguage in which the axioms and rules are presented. We briefly discuss the possible spectrum. Detailed references to and discussion of previous work are left to the next section.

We first consider infinite equational axiomatisations. The sets of all equivalent pairs of expressions are trivially complete (and indeed also recursive, as bisimulation and language equivalence of finite charts are decidable). More interesting are sets generated in a uniform way from a small number of schemata (using some ad hoc metanotation) such as the group or commutative identities reproduced in §2.3, or the axioms $C_{mn}$ of §2.3.4. These expose some of the structure of the equivalence concerned. They are compelling if they do this well and are simple — an informal criterion. Tighter questions can be asked about the existence of finite axiomatisations. To make these precise and in particular to state nonexistence results we must define classes of allowable axiomatisations. We give a variety of notions of 'pure' system, i.e. systems in which all terms that appear are of the object language.

For an object language without binding operators, such as the $*$-expressions, the situation is unproblematic. We take a finite pure equational axiomatisation to consist of a finite set $\mathcal{A}$ of pairs of object language terms together with the rules in Figure 2–1 for congruence and the following:

$$\frac{}{E = F}\ E = F \in \mathcal{A} \qquad\qquad \frac{E = F}{E[G/a] = F[G/a]}.$$

Moving to a richer logic, a finite pure horn clause axiomatisation consists of a finite set of expressions of the form

$$E_1 = F_1 \wedge \cdots \wedge E_m = F_m \rightarrow E' = F'$$

$$\frac{}{E = E}\ ref$$

$$\frac{E = F}{F = E}\ sym \qquad \frac{E = F \quad F = G}{E = G}\ tran$$

$$\frac{E = F \quad E' = F'}{E + E' = F + F'}\ {+cong} \qquad \frac{E = F \quad E' = F'}{E \cdot E' = F \cdot F'}\ {\cdot\, cong}$$

$$\frac{E = F}{E^* = F^*}\ {*cong} \qquad \frac{E = F \quad E' = F'}{E^\star E' = F^\star F'}\ {^\star cong}$$

**Figure 2–1:** Congruence rules for $*$-expressions

or

$$E_1 = F_1 \wedge \cdots \wedge E_m = F_m \rightarrow \perp,$$

where $E, F$ are object language terms and $m \geq 0$, together with rules for congruence, instantiation and manipulation of $\wedge$, $\rightarrow$ and $\perp$.

In the presence of binding operators the situation is not quite so simple, as equations between object language terms are then very inexpressive. For example alpha conversion or the unfolding of recursions would each require an infinite set of equations. Instead one might write down schemata such as

$$\begin{aligned} E[\mu X E/X] &= \mu X E \\ \mu Y E[Y/X] &= \mu X E\ , \text{ if } Y \text{ not free in } \mu X E, \end{aligned}$$

expressing unfolding and alpha conversion. It would be possible to give a definition (perhaps several) of a natural class of such schemata and, by formalising the manner in which they are instantiated, a definition of a class of axiomatisations containing a finite set of them. We do not because it would be rather complex — the schemata above contain metavariables over both object variables ($X, Y$) and terms ($E$), substitutions and a simple side condition. Instead we embed the $\mu$-expressions into a fragment of the simply typed lambda calculus. This is described in detail in Chapter 3. For now we note only that it lets us take a finite pure equational axiomatisation to be simply a finite set of pairs of lambda terms (of appropriate types) together with standard rules for $\beta\eta$ equality. The first schema above can be written simply as

$$e : P \to P \triangleright \mathit{fix}\ e = e(\mathit{fix}\ e) : P$$

and the second is subsumed by lambda calculus alpha conversion. A finite pure horn clause axiomatisation would then be as before except that $E, F$ now range over typed lambda terms.

A number of finite horn clause axiomatisations have been given which are essentially impure in that they contain rules which cannot be put into this form. We mention two, deferring full definitions of the notation involved to later in this chapter. Firstly the unique fixed point rule schema from [Mil84]:

$$E = F[E/X] \wedge X \text{ guarded in } F \quad \to \quad E = \mu X F$$

in which the 'guarded' condition is not equational and secondly the functorial implication from [BÉ94]:

$$\forall i \in m\ .\ E_i[Y_{\rho(j)}/X_j]_{j \in m} = F_{\rho(i)} \quad \to \quad \forall i \in m\ .\ \tilde{\mu}_i \tilde{X}\ \tilde{E} = \tilde{\mu}_{\rho(i)} \tilde{Y}\ \tilde{F}$$

which has an instance for each pair $m, n \geq 1$ of natural numbers and surjective function $\rho : m \to n$.

In summary, for both $*$-expressions and $\mu$-expressions we have a hierarchy of classes of axiomatisations:

- Infinite (but elegant) equational
- Finite impure horn clause
- Finite pure horn clause
- Finite pure equational

of which the first two are not formally defined. For $\mu$-expressions the last two will be further subdivided by placing restrictions on the types allowable in axioms.

We now return to the general motivation for axiomatisability results of Chapter 1 and discuss it with respect to this hierarchy.

#### Human and Machine reasoning

Sound systems are in principle useful for reasoning about the equivalence of expressions. In practice, however, they are only useful for human reasoning about relatively small expressions. For that there is no advantage in restricting to a weak metalanguage — indeed it may be useful to apply (say) both unique fixed point reasoning from the system of [Mil84] and the schema from the infinite system of §2.3.1 in the same work. For larger expressions model-based techniques which apply directly to the chart denoted by an expression seem to be preferable, as efficient algorithms have been found. We should note that for more expressive calculi the situation may be very different. There it may be possible to use (human) understanding of the (term) structure in a particular equivalence problem to quickly find a good proof in a suitable formal system. It may also be possible to orient the equations of an equational axiomatisation to give a useful term rewriting system.

#### Insight

As stated in Chapter 1 a more important motivation for this work is the insight gained from proofs — of positive (completeness) or negative (non-axiomatisability) results. In general we would like positive results as far down the hierarchy as possible since an inexpressive metalanguage requires a better understanding. (In the extreme case of a strong logic the definitions of the equivalences could be directly but pointlessly written down.) Conversely, negative results should be as far up as possible although obviously their statements require formal definitions of classes of axioms.

#### Expressivity and Comparison

Positive results can be used to compare different equivalences, particularly if there is a clean extension of an axiomatisation for one which is complete for another. A more general approach is to consider questions of *relative axiomatisability*, i.e. to consider whether an equivalence $\sim_1$ is axiomatisable in one of the senses above augmented with a rule

$$E \sim_2 F \quad \rightarrow \quad E = F$$

for another equivalence $\sim_2$. A different form of comparison is to keep the equivalence fixed and vary the expressions. The existence of an axiomatisation can be seen as depending on the metalanguage being expressive with respect to the object language. This can be investigated by looking for axiomatisable sub-languages of a non-axiomatisable object language.

#### Other models

Throughout this section and in the greater part of the thesis we have considered axiomatisation of the equivalences induced by a few particular, intended, models of the expressions. An alternative approach is to consider the category of all models equipped with suitable homomorphisms. Pure equational and horn clause axiomatisations (not necessarily finite) are desirable in that they permit the results of universal algebra to be applied. For the $*$-expressions the notion of model is standard and the results (from [MT92], say) are immediately applicable. In particular any pure equational (resp. horn clause) axiomatisation defines a variety (resp. quasi variety) containing an initial model. For the $\mu$-expressions there are several possible notions of model. Two, the preiteration algebras and strong preiteration algebras, are given in [BÉ94] together with an analogue of the variety theorem. They are specialized from work on iteration theories. These provide a general setting for the investigation of the equational logic of fixed point operators. We cannot do justice to the work here, but instead refer the reader to [BÉ93b].

One can also look for axiomatisations which are sound in several intended models, as is done for example in [Koz94] for the $*$-expressions. In particular, axioms involving a 'guarded' condition tend not to be valid in coarser models.

Domain theoretic models of the $\mu$-expressions will be briefly discussed in §2.3.5 along with the infinitary rules of $\omega$-induction and the approximation induction principle.

### Completeness and $\omega$-completeness

We have several times referred in an ambiguous way to completeness results. There can be a significant difference between completeness for closed and open terms, as shown for example in [Mol89] for a calculus with parallel composition. We recall the definitions from §2.3.2 of that work.

Typically one might have an equivalence $\simeq$ defined over closed terms and a proof system deriving judgments $\vdash E = F$ between arbitrary terms. The proof system is then said to be complete if

- for all closed $E, F$ if $E \simeq F$ then $\vdash E = F$,

and $\omega$-complete if

- for all $E$ and $F$, if $E\rho \simeq F\rho$ for all closed substitutions $\rho$ then $\vdash E = F$.

For $\mu$-expressions $\omega$-completeness does not seem to be a problem. The negative result of Chapter 3 (Theorem 6) shows that there is no complete system and so immediately that there is no $\omega$-complete system. The positive results mentioned in the next section (Theorems 1, 2, 3 and 4) are all with respect to all $\mu$-expressions and so could be considered as $\omega$-completeness results.

On the other hand, one could take the pure axiomatisations obtained from the first two and last of these by embedding into the lambda calculus. They might then be considered to be $\omega$-complete iff all the sound identities between terms containing higher type variables are derivable (i.e., with notation from Chapter 3: iff they are complete for $\sim_{ext}$ over $T^P_\omega$, not just over $T^P_0$). We do not have a completeness result over $T^P_\omega$, or even over $T^P_1$ — such a result would be of some limited interest.

## 2.3 An overview of previous work

There is a substantial literature dealing with axiomatisation of equivalences over calculi denoting finite state machines. A part of it is summarised in Figure 2–2, classified by the equivalence, calculus and strength of logic addressed and labelled $\surd$ (resp. $\times$) if finite complete systems are given (resp. shown not to exist). Results labelled [Sew] are contained in this thesis and have also been announced in [Sew94]. Care must be taken when interpreting the figure as there are differ-

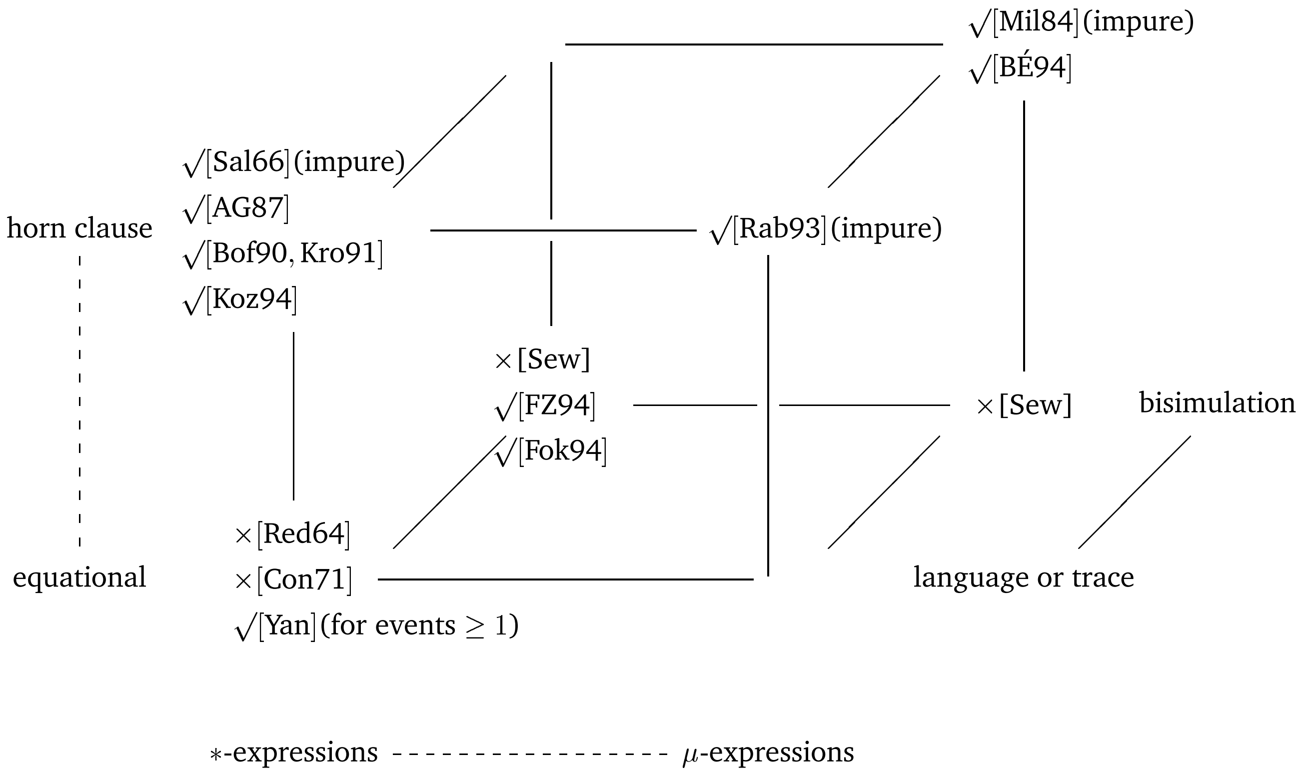



**Figure 2–2:** Finite axiomatisability results

ing definitions, in particular of the syntax of the $*$-expressions and of language equivalence over $\mu$-expressions.

The figure is not intended to imply that all vertices have equal interest — in fact the expressivity results of Propositions 6, 13, 14 suggest that the language/$*$-expression and bisimulation/$\mu$-expression edges are more significant than the others. Further, there are many interesting equivalences which are not shown. There are trivial implications only along the logic (vertical) axis — relative axiomatisability results would give implications along the equivalence axis but we have none applicable.

We will first discuss the $\mu$-expression problems in some detail, not attempting to be exhaustive or historical but rather focusing on the techniques used to show completeness. We then discuss the $*$-expression work more briefly and finally mention some more distantly related work. Soundness proofs will be omitted.

In the absence of recursion there is no difficulty. The axioms

$$\begin{array}{lrcl} A1 & E+(F+G) & = & (E+F)+G \\ A2 & E+F & = & F+E \\ A3 & E+0 & = & E \\ A4 & E+E & = & E \end{array}$$

together with rules for congruence are complete for bisimulation over the expressions given by

$$E ::= 0 \mid aE \mid E+E$$

(from [HM80]) and adding $P1$ or $P1, P2$

$$\begin{array}{lrcl} P1 & aE+aF & = & a(E+F) \\ P2 & a0 & = & 0 \end{array}$$

gives completeness for trace congruence or language equivalence respectively. (We present these and subsequent axioms informally using metavariables $E, F$. They could easily be put in a pure form in a lambda calculus similar to that of Chapter 3.) Axiomatisations for a number of equivalences in the linear-branching time spectrum are collected in [Gla90]. Bisimulation over arbitrary well-founded GSOS systems is finitely axiomatised in [ABV92].

### 2.3.1 Infinite systems for $\mu$-expressions

We now give an infinite system for bisimulation over all $\mu$-expressions. That this is possible is of course trivial and indeed the completeness proof, although new, is not terribly difficult. It is included for several reasons. Firstly, it provides some insight and suggests interesting conjectures about the completeness of subsystems. Indeed the nonexistence proof of Chapter 3 arose from considering whether restrictions of a closely related system were complete. Secondly the proof is closely related to those for certain finite impure horn clause systems considered in the next section. Finally the result gives a relative axiomatisability result as an immediate corollary and may be useful for completeness proofs of other systems — e.g. the pure horn clause system of §2.3.3 — as it is only necessary to show that all instances of these equations are derivable.

Together with $A1$–$4$ there are axioms for alpha conversion, recursion unfolding, absorbing unguarded variables and a form of Bekič's lemma.

$$\begin{array}{llcl}
A5 & \mu Y E[Y/X] & = & \mu X E\ ,\ \text{ if } Y \text{ not free in } \mu X E \\
B1 & E[\mu X E\,/X] & = & \mu X E \\
B2 & \mu X X + E & = & \mu X E \\
B3 & \mu X E[\mu Y F\,/Y] & = & \mu X E[\mu Y F[\mu X E\,/X]\,/Y]
\end{array}$$

In addition we have a set of axioms $C_{mn}$ indexed by pairs of non-zero natural numbers. Loosely, $C_{mn}$ expresses the equality of any $m$-state transition system with one in which each state has been split into $n$ bisimilar states. To state these precisely we need some additional notation.

A natural number, say $n$, will be identified with the set $\{0, \ldots, n-1\}$. We introduce sequences of formal equations, i.e. nonempty sequences of pairs of a variable (all of which must be distinct) and a $\mu$-expression. A variety of notations will be used — the following all denote the same sequence.

$$\begin{array}{c}
\langle X_0 = E_0, \ldots, X_{n-1} = E_{n-1}\rangle \\
\langle X_i = E_i \mid i \in n\rangle \\
\langle \tilde{X} = \tilde{E}\rangle
\end{array}
\qquad
\left\langle \begin{array}{c} X_0 = E_0 \\ \vdots \\ X_{n-1} = E_{n-1} \end{array} \right\rangle$$

The fixed point of a sequence is a $\mu$-expression defined by induction:

$$\begin{array}{rcl}
\mu\,\langle X = E\rangle & \stackrel{\text{def}}{=} & \mu X E \\
\mu\,\langle \tilde{X} = \tilde{E}, X = E\rangle & \stackrel{\text{def}}{=} & \mu\,\langle \tilde{X} = \tilde{E}[\mu X E\,/X]\rangle .
\end{array}$$

The simultaneous substitution of $n$ expressions $E_i \mid i \in n$ for $n$ distinct variables $X_i \mid i \in n$ will be written as $[E_i/X_i]_{i \in n}$. We let $i$ and $k$ range over $m$, $j$ over $n$ and $f$ over the functions from $n$ to $n$. Pairs of numbers are given the product ordering. Finally, the axiom is

$$\begin{array}{lll}
C_{mn} & & \mu\,\langle X_i = E_i[X_k/Z_{kf}]_{k \in m, f \in n^n} \mid i \in m\rangle \\
 & = & \mu\,\langle Y_{ij} = E_i[Y_{k,f(j)}/Z_{kf}]_{k \in m, f \in n^n} \mid i \in m, j \in n\rangle ,
\end{array}$$

in which we suppose that none of the $X_k$ or $Y_{kj}$ appear in any of the expressions $E_i$. Note that the substitutions appearing have as domain all $mn^n$ of the $Z_{kf}$.

The idea is rather simpler than the notation. Consider a non-binding occurrence of a variable (say $X_k$) in a sequence of $m$ equations. If each state is split into $n$, giving a sequence of $mn$ equations, it becomes $n$ occurrences. Each might be any of the $Y_{kj} \mid j \in n$. To allow for any such 'behaviour' of the original occurrence we begin with $n^n$ formal variables $Z_{kf}$ for each $X_k$.

For example here is $C_{12}$, omitting the first index of the $Y$'s (which is always 0):

$$\mu \left\langle \; X_0 = E_0[X_0X_0X_0X_0] \; \right\rangle = \mu \left\langle \begin{array}{l} Y_0 = E_0[Y_0Y_0Y_1Y_1] \\ Y_1 = E_0[Y_0Y_1Y_0Y_1] \end{array} \right\rangle$$

and $C_{22}$:

$$\mu \left\langle \begin{array}{l} X_0 = E_0[X_0X_0X_0X_0X_1X_1X_1X_1] \\ X_1 = E_1[X_0X_0X_0X_0X_1X_1X_1X_1] \end{array} \right\rangle = \mu \left\langle \begin{array}{l} Y_{00} = E_0[Y_{00}Y_{00}Y_{01}Y_{01}Y_{10}Y_{10}Y_{11}Y_{11}] \\ Y_{01} = E_0[Y_{00}Y_{01}Y_{00}Y_{01}Y_{10}Y_{11}Y_{10}Y_{11}] \\ Y_{10} = E_1[Y_{00}Y_{00}Y_{01}Y_{01}Y_{10}Y_{10}Y_{11}Y_{11}] \\ Y_{11} = E_1[Y_{00}Y_{01}Y_{00}Y_{01}Y_{10}Y_{11}Y_{10}Y_{11}] \end{array} \right\rangle .$$

We show how these axioms can be used to derive the equation

$$\mu XaX = \mu Xa^nX$$

(for any $n \geq 1$) in a simple way. Given $n$, consider the instance of $C_{1n}$ obtained by taking $E_0 = aZ_{0,f}$ where $f : n \to n$ is defined by

$$f(x) \stackrel{\text{def}}{=} (x+1) \bmod n.$$

This instance of $C_{1n}$ is

$$\mu X_0 a X_0 = \mu Y_0 a \mu Y_1 a \ldots \mu Y_{n-1} a Y_0$$

from which the equation follows by use of $B1$ and $A5$.

The completeness of these axioms is reasonably straightforward. We show that any expression is provably equal (using axioms $A, B$ only) to the fixed point of a sequence of equations in a standard form. Given such sequences for two bisimilar expressions we construct a 'product' sequence of equations and use instances of $C$ to show that all three fixed points are provably equal. The proofs of Lemmas 18, 19, 20 and 21 are banished to Appendix A.

**Notation** If $E = F$ is provable from the axioms $Q$ together with $A1$–$5$ and rules for congruence then we write $Q \vdash E = F$. We work up to the equivalence induced by $A1$–$5$ throughout.

First we give two simple results. Provable equality lifts from expressions to the fixed points of sequences and the binary $B3$ implies an $m$-ary form (palliating the occasional ambiguity in the order of sequences).

**Lemma 18** *For axioms $Q$, if $Q \vdash E_i = F_i$ for all $i$ then $Q \vdash \mu \langle \tilde{X} = \tilde{E} \rangle = \mu \langle \tilde{X} = \tilde{F} \rangle$.*

**Lemma 19** *If $\pi : m \to m$ is a permutation with $\pi(0) = 0$ then $B3 \vdash \mu \langle X_i = E_i \mid i \in m \rangle = \mu \langle X_{\pi(i)} = E_{\pi(i)} \mid i \in m \rangle$.*

Expressions are provably equal to the fixed point of a sequence of equations that is in a standard form.

**Definition** A sequence of equations $\langle X_i = E_i \mid i \in m \rangle$ is *standard* if each $E_i$ is of the form

$$E ::= 0 \mid W \mid aX \mid E + E$$

(where $X$ ranges over $\{ X_i \mid i \in m \}$ and $W$ ranges over all other variables) and does not contain a free occurrence of $X_0$. A standard sequence $\langle X_i = E_i \mid i \in m \rangle$ can be regarded as a finite chart with states $m$, transitions $i \xrightarrow{a} j$ iff $aX_j$ is a summand of $E_i$, visibilities $i \rhd W$ iff $W$ is a summand of $E_i$ and initial state $0$. It is *accessible* if $\forall i \in m \,.\, 0 \longrightarrow^* i$.

**Lemma 20** *For any $\mu$-expression $E$ there is a standard accessible sequence $\langle \tilde{X} = \tilde{E} \rangle$ such that $B \vdash E = \mu \langle \tilde{X} = \tilde{E} \rangle$.*

**Lemma 21** *If $\langle \tilde{X} = \tilde{E} \rangle$ is standard then $\langle \tilde{X} = \tilde{E} \rangle \sim \mu \langle \tilde{X} = \tilde{E} \rangle$.*

There is a natural 'product' definable over bisimilar standard sequences.

**Definition** If $\langle X_i = E_i \mid i \in m \rangle$, $\langle Y_j = F_j \mid j \in n \rangle$ are standard, accessible and bisimilar then we define a standard accessible sequence $\langle \tilde{X} = \tilde{E} \rangle \otimes \langle \tilde{Y} = \tilde{F} \rangle$ as follows. Let $R \subseteq m \times n$ be the largest bisimulation between the sequences considered as charts. It is clear that if $i \; R \; j$ then

- if $W$ is a summand of $E_i$ then $W$ is a summand of $F_j$,
- if $aX_{i'}$ is a summand of $E_i$ then there is some $j'$ such that $aY_{j'}$ is a summand of $F_j$ and $i' \; R \; j'$

(and symmetrically) and further that $\mathrm{dom}(R) = m$, $\mathrm{ran}(R) = n$ and $0 \; R \; 0$. Now

$$\langle \tilde{X} = \tilde{E} \rangle \otimes \langle \tilde{Y} = \tilde{F} \rangle \stackrel{\text{def}}{=} \langle Y_{ij} = G_{ij} \mid i \; R \; j \rangle,$$

where

$$\begin{array}{rcl} G_{ij} & \stackrel{\text{def}}{=} & \sum\{\, W \mid W \text{ is a summand of } E_i \text{ and } F_j \,\} \\ & + & \sum\{\, aY_{i'j'} \mid aX_{i'} \text{ is a summand of } E_i,\ aY_{j'} \text{ is a summand of } F_j \\ & & \text{and } i' \; R \; j' \,\}. \end{array}$$

**Lemma 22** *If $\langle \tilde{X} = \tilde{E} \rangle$, $\langle \tilde{Y} = \tilde{F} \rangle$ are standard, accessible and bisimilar then $C \vdash \mu \, \langle \tilde{X} = \tilde{E} \rangle = \mu \, (\langle \tilde{X} = \tilde{E} \rangle \otimes \langle \tilde{Y} = \tilde{F} \rangle)$ and symmetrically.*

PROOF We show $C \vdash \mu \, \langle \tilde{X} = \tilde{E} \rangle = \mu \, \langle Y_{ij} = G_{ij} \mid i \; R \; j \rangle$, with notation as in the previous definition. Take $E'_i \mid i \in m$ to be

$$\begin{array}{rcl} E'_i & \stackrel{\text{def}}{=} & \sum\{\, W \mid W \text{ is a summand of } E_i \,\} \\ & + & \sum\{\, aZ_{i'f} \mid \forall j \; . \; i \; R \; j \Rightarrow aY_{i'f(j)} \text{ is a summand of } G_{ij} \,\}. \end{array}$$

It is straightforward to check that the equalities

$$\begin{array}{rcl} E_i & = & E'_i[X_k/Z_{kf}]_{k \in m, f \in n^n} \\ G_{ij} & = & E'_i[Y_{k,f(j)}/Z_{kf}]_{k \in m, f \in n^n} \end{array}$$

are provable from $A1$–$5$ so by Lemma 18 so are

$$\begin{array}{rcl} \mu \, \langle \tilde{X} = \tilde{E} \rangle & = & \mu \, \langle \tilde{X} = \tilde{E}'[X_k/Z_{kf}]_{k \in m, f \in n^n} \rangle \\ \mu \, \langle Y_{ij} = G_{ij} \mid i \; R \; j \rangle & = & \mu \, \langle Y_{ij} = E'_i[Y_{k,f(j)}/Z_{kf}]_{k \in m, f \in n^n} \mid i \; R \; j \rangle. \end{array}$$

It remains to note that

$$C \vdash \mu \, \langle \tilde{X} = \tilde{E}'[X_k/Z_{kf}]_{k \in m, f \in n^n} \rangle \;=\; \mu \, \langle Y_{ij} = E'_i[Y_{k,f(j)}/Z_{kf}]_{k \in m, f \in n^n} \mid i \; R \; j \rangle.$$

□

**Theorem 1** *The axioms $A1$–$5$, $B1$–$3$ and $\{\, C_{mn} \mid m, n \in \omega \times \omega \,\}$ are sound and complete for bisimulation over $\mu$-expressions.*

PROOF Completeness is immediate from Lemmas 20, 21 and 22. □

### Complete subsets

This theorem can be sharpened somewhat. In the presence of $C_{12}$ and $A1$–$5$ axiom $B1$ can be simplified to the following.

$$B1' \qquad \mu X E \;=\; E \;, \text{ if } X \text{ not free in } E.$$

Further, it is clear that if $m' > m$ and $n' > n$ then any instance of $C_{mn}$ is an instance of $C_{m'n'}$ so it suffices to take enough instances of $C_{mn}$ to cover $\omega \times \omega$ (with respect to the product ordering). One can then ask whether this is a necessary condition. We conjecture that it is but we know only that instances with arbitrarily large $n$ are required (from the results of Chapter 3).

### Relative axiomatisability

The axioms $C_{mn}$ (and also $A5, B1, B3$) are all sound for infinite term equality of $\mu$-expressions. It is therefore an immediate corollary of Theorem 1 that bisimulation is finitely equationally axiomatisable relative to infinite term equality, i.e. that $A1$–$4$, $B2$ and the implication

$$E =_{\text{infinite term}} F \quad \to \quad E = F$$

are complete for bisimulation. In §4.4 we show that weak congruence is finitely equationally axiomatisable relative to bisimulation, and hence also relative to infinite term equality.

One could ask whether adding $P1$ gives completeness for trace congruence. We conjecture that it does not. More generally one could ask whether it or any other equivalence in the linear-branching time spectrum is finitely equationally axiomatisable relative to infinite term equality. Answers to these questions would perhaps be of technical interest. Positive results would isolate the 'non-finitely-equational' part of the equivalences to the $C_{mn}$. Negative results would of course be stronger than simple non-finite-axiomatisability results for the same equivalence. They might nonetheless be easier to prove, being rather more focussed.

### The commutative identities

A closely related infinite system has been given by Bloom and Ésik in [BÉ94]. We will simply state the result here, without doing justice to the bulk of their work (which is in a more general setting). We first need to define another $n$-ary fixed point operator.

**Definition** If $\tilde{X}, \tilde{E}$ are $n$-tuples of variables and terms then $\tilde{\mu}\tilde{X}\ \tilde{E}$ is an $n$-tuple of terms defined inductively:

$$\begin{aligned} \tilde{\mu}XE &\stackrel{\text{def}}{=} \mu XE \\ \tilde{\mu}\tilde{X}X\ \tilde{E}E &\stackrel{\text{def}}{=} (\tilde{\mu}\tilde{X}\tilde{E}[\mu XH\,/X]), \mu XH \end{aligned}$$

where $H \stackrel{\text{def}}{=} E[\tilde{\mu}\tilde{X}\tilde{E}\,/\tilde{X}]$. The $i$th component of the tuple $\tilde{\mu}\tilde{X}\tilde{E}$ will be written $\tilde{\mu}_i\tilde{X}\tilde{E}$.

There is a *commutative identity* for each $m, n \geq 1$ and family of functions $\rho_{ij} : n \to n \mid i \in m, j \in n$. As before, each identity loosely expresses the equality of an $m$-state transition system with one in which each state has been split into $n$ bisimilar states. We let $i, i', k$ range over $m$ and $j, j', l$ range over $n$. The commutative identity is then

$$\begin{aligned} \forall i' \in m, j' \in n\ .\ \ &\tilde{\mu}_{i'}\ \langle X_i = E_i[X_k/Z_{kl}]_{k\in m, l\in n} \mid i \in m\rangle \\ = \ \ &\tilde{\mu}_{i'j'}\ \langle Y_{ij} = E_i[Y_{k,\rho_{ij}(l)}/Z_{kl}]_{k\in m, l\in n} \mid i \in m, j \in n\rangle \end{aligned}$$

in which the variables $X_k$ and $Y_{kl}$ are presumed not to occur free in any of the terms $E_i$. For the reader familiar with [BÉ94] this is just

$$\mu x.(t \parallel [R/x]) = (\mu y.t[y_\rho/x])_\rho$$

in notation which may be easier to understand at first sight, although perhaps harder to work with.

The following two axioms are also required. For the first it is assumed that $Z$ does not occur free in $E$.

$$\begin{array}{lrcl} E1 & \mu ZE[ZZ/XY] &=& \mu X\mu YE \\ E2 & \mu XE[F/X] &=& E[\mu XF[E/X]\,/X] \end{array}$$

These are the *double iteration identity* and the *composition identity*.

**Theorem 2 ([BÉ94, Theorem 6.6])** *The commutative identities together with axioms $A1$–$5$, $B2$, $E1$–$2$ are sound and complete for bisimulation over $\mu$-expressions.*

We conjecture that in the presence of a finite number of axioms (e.g. $A5$,$B1$–$3$,$E1$–$2$) the commutative identities for $m, n^n$ imply all the consequences of $C_{mn}$, which in turn imply all the consequences of the commutative identities for $m, n$.

### 2.3.2 Finite impure horn clause systems

A number of complete systems have been given that contain an impure horn clause expressing the fact that certain equations have unique solutions (together with a finite set of equational axioms). The first seems to be that for language equivalence of $*$-expressions by Salomaa [Sal66]. For $\mu$-expressions there are complete systems for bisimulation [Mil84], weak congruence [Mil89], branching bisimulation congruence [Gla93a], divergence bisimulation [Gla93b] and trace congruence [Rab93]. We will reproduce the system of [Mil84] for bisimulation and sketch a completeness proof.

The axioms are $A1$–5,$B1$–2 together with rules for congruence, substitution and $D$:

$$D \qquad E = F[E/X] \wedge X \text{ guarded in } F \;\; \rightarrow \;\; E = \mu X F$$

where $X$ is guarded in $F$ if every free occurrence of $X$ in $F$ is contained in a subexpression $aG$. We write $\vdash E = F$ if $E = F$ is provable from these axioms.

The completeness proof below is a minor rearrangement of that given in [Mil84]. It is similar in structure to that of the previous subsection (in fact the latter was based on this). The definitions of standard sequences of equations and of the product of two bisimilar standard sequences are common. Now, however, instead of showing that expressions are provably equal to the fixed points of standard sequences they are shown to be provable solutions of standard sequences. The axiom $D$ is then shown to imply an $m$-ary form, that certain sequences have provably unique solutions (this is the only place where $D$ is required).

**Definition** A expression $E$ is a *provable solution* of the equations $\langle X_i = E_i \mid i \in m\rangle$ if there exist expressions $G_i \mid i \in m$ such that $\vdash E = G_0$ and $\forall i \in m . \vdash G_i = E_i[\tilde{G}/\tilde{X}]$. It is a *provably unique* solution if for any $E'$ that is also a provable solution $\vdash E = E'$.

**Lemma 23 ([Mil84, Theorem 5.9])** *For any expression $E$ there is a standard accessible sequence $\langle \tilde{X} = \tilde{E}\rangle$ for which $E$ is a provable solution.*

**Lemma 24** *If $E$ is a provable solution of a standard $\langle \tilde{X} = \tilde{E}\rangle$ then $E \sim \langle \tilde{X} = \tilde{E}\rangle$.*

PROOF Straightforward. □

**Lemma 25 ([Mil84, Proof of Theorem 5.10])** *If $\langle\tilde{X} = \tilde{E}\rangle, \langle\tilde{Y} = \tilde{F}\rangle$ are standard, accessible and bisimilar with provable solutions $E, F$ then $E, F$ are both provable solutions of $\langle\tilde{X} = \tilde{E}\rangle \otimes \langle\tilde{Y} = \tilde{F}\rangle$.*

**Definition** A sequence of equations $\langle X_i = E_i \mid i \in m\rangle$ is *guarded* if each $X_i$ is guarded in all the $E_j$. Clearly any standard sequence is guarded.

**Lemma 26 ([Mil84, Theorem 5.7])** *Any guarded sequence of equations has a provably unique solution.*

**Theorem 3 ([Mil84, Theorem 5.10])** *The axioms $A1$–$5$, $B1$, $B2$, $D$ are sound and complete for bisimulation over $\mu$-expressions.*

PROOF Completeness is immediate from the above lemmas. □

### 2.3.3 Finite pure horn clause systems

The previously cited [BÉ94] by Bloom and Ésik also contains a finite pure horn clause system. We present a minor variant which has axioms $A1$–$5$, $B2$ and $E1$–$3$ together with rules for congruence and substitution. $E3$, also known as the *GA implication*, is given below. It is assumed that $Z$ is not free in $E$ or $F$.

$$E3 \quad \mu Z E[ZZ/XY] = \mu Z F[ZZ/XY] \;\;\rightarrow\;\; \mu Z E[ZZ/XY] = \mu X F[\mu Y E/Y]$$

We will give an outline of the completeness proof. It uses another implication — the functorial implication — which cannot be written as a single pure horn clause.

**Definition** The *functorial implication*for a surjective function $\rho : m \rightarrow n$, where $m, n \geq 1$, is

$$\forall i \in m \;.\; E_i[Y_{\rho(j)}/X_j]_{j \in m} = F_{\rho(i)} \;\;\rightarrow\;\; \forall i \in m \;.\; \tilde{\mu}_i \tilde{X}\; \tilde{E} = \tilde{\mu}_{\rho(i)} \tilde{Y}\; \tilde{F}$$

where $\tilde{X}, \tilde{E}$ (resp. $\tilde{Y}, \tilde{F}$) are $m$-tuples (resp. $n$-tuples) of variables and terms and it is supposed that no $Y_k$ is free in any $E_i$.

The core of the completeness proof is Lemma 28 below which can be shown by induction on $n$.

**Lemma 27 ([BÉ94, Lemma 7.5])** *The axioms $A5$, $E1$–$3$ imply all instances of the functorial implication for functions $\rho : m \to 1$.*

**Lemma 28 ([BÉ94, Lemma 7.4])** *The axioms $A5$, $E1$–$2$ and the functorial implication for all $\rho : m \to 1$ $(m \geq 1)$ imply all instances of the functorial implication for surjective $\rho : m \to n$.*

**Lemma 29 ([BÉ94, Proof of Prop 7.1])** *The axioms $A5$, $E1$–$2$ and the functorial implication for all $\rho : m \to n$ imply all instances of the commutative identity.*

**Theorem 4 ([BÉ94, Theorem 6.6])** *The axioms $A1$–$5$, $B2$, $E1$–$3$ are sound and complete for bisimulation over $\mu$-expressions.*

PROOF Completeness is immediate from the above lemmas and Theorem 2. □

We conjecture that $E3$ is also sound for weak congruence. If so then it follows that the axioms $F1$–$3$, which are introduced in §4.4.1 in our proof that weak congruence is equational relative to bisimulation, are, together with $A1$–$5$, $B1$–$3$, $E1$–$3$, a finite pure horn clause axiomatisation of weak congruence over $\mu$-expressions.

### 2.3.4 Systems for $*$-expressions

Unless stated otherwise all work mentioned in this subsection is with respect to language equivalence.

It has been shown that language equivalence of $*$-expressions is not finitely equational. Three proofs of this are sketched in [Con71]. The first was apparently given by Redko [Red64] and Salomaa in an incomplete form and later completed by Pilling. We will reproduce a sketch of the second, from [Con71, p.106], here to permit an easy comparison with our proof that bisimulation of $\mu$-expressions is not finitely equational. Both intuitively say that any finite set of equations does not permit the introduction of arbitrary prime factors into the lengths of loops or iterates. Whether this can be made precise by giving a single proof for both cases is an interesting open question.

**Theorem 5 (from [Con71])** *Language equivalence of $*$-expressions is not finitely equationally axiomatisable.*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $C1$ | $A + 0$ | $=$ | $A$ | $C8$ | $A \cdot (B + C)$ | $=$ | $(A \cdot B) + (A \cdot C)$ |
| $C2$ | $A + B$ | $=$ | $B + A$ | $C9$ | $(B + C) \cdot A$ | $=$ | $(B \cdot A) + (C \cdot A)$ |
| $C3$ | $(A + B) + C$ | $=$ | $A + (B + C)$ | $C10$ | $(A \cdot B) \cdot C$ | $=$ | $A \cdot (B \cdot C)$ |
| $C4$ | $A \cdot 0$ | $=$ | $0$ | $C11$ | $(A + B)^*$ | $=$ | $(A^* \cdot B)^* \cdot A^*$ |
| $C5$ | $0 \cdot A$ | $=$ | $0$ | $C12$ | $(A \cdot B)^*$ | $=$ | $1 + A \cdot (B \cdot A)^* \cdot B$ |
| $C6$ | $A \cdot 1$ | $=$ | $A$ | $C13$ | $(A^*)^*$ | $=$ | $A^*$ |
| $C7$ | $1 \cdot A$ | $=$ | $A$ | | | | |

**Figure 2–3:** Axioms $C1$–13

PROOF For any prime $p$ there is a model $A_p$ of the $*$-expressions in which all tautologies of less than a certain size are satisfied but others are not. The underlying set of the model is a set of $2^p{+}1$ $*$-expressions over a distinguished action $x \in Act$:

$$|A_p| \stackrel{\text{def}}{=} \Big\{ \sum_{i \in I} x^i \;\Big|\; I \subseteq \{0, \ldots, p-1\} \Big\} \cup \{x^*\} .$$

Using axioms $C1$–13 as in Figure 2–3 and the assumption $x^p = 1$ it can be shown that any $*$-expression over $x$ is provably equal to a unique element of $|A_p|$, defining the operations of $A_p$. If $\sigma$ is an environment, i.e. a function $\sigma : Act \to |A_p|$, then we write $[\![E]\!]\sigma$ for the interpretation of a term $E$ in $A_p$.

Let $\equiv$ be the equivalence relation over $|A_p|$ that identifies $x^*$ and $\sum_{i \in p} x^i$. It can be shown that $\equiv$ is a congruence and that $A_p/\equiv$ satisfies all tautologies. On the other hand $A_p$ clearly does not satisfy the tautology

$$C14.p \qquad A^* \;=\; (A^p)^* \cdot \textstyle\sum_{i \in p} A^i$$

at $A = x$. It remains to find, for any finite set of tautologies, a prime $p$ such that all are satisfied by $A_p$.

Let the canonical $*$-expressions be the sums of terms, each of which is $0$, $1$ or contains no occurrence of $0$, $1$ or $+$. The length of such is the number of nonzero subterms. Any $*$-expression is provably equal to a canonical one, using $C1$–13. Further, any expression obtained from a canonical one by replacing actions by $0$ or $1$ is provably equal to a canonical one of the same or smaller length.

We show that any tautology $E = F$ with $E, F$ both canonical and of length less than $p$ holds in $A_p$. Suppose not (for a contradiction), i.e. for some $\sigma : Act \to |A_p|$ that $[\![E]\!]\sigma \neq [\![F]\!]\sigma$. We can assume w.l.g. that for all actions $y$ occurring in $E, F$

that $\sigma(y) \notin \{0,1\}$ and further that $[\![E]\!]\sigma = \sum_{i\in p} x^i$ and $[\![F]\!]\sigma = x^*$ (as $A_p/\equiv$ satisfies all tautologies). By soundness any action $y$ occurring under a $*$ in $E$ or $F$ must occur under a $*$ in both. $\sigma(y)$ must then be a single power $x^i$, otherwise $[\![E]\!]\sigma = x^*$. Let $E', F'$ be $E, F$ with all variables that do not occur under a $*$ replaced by $1$. It can be shown that $[\![E']\!]\sigma \neq x^* = [\![F']\!]\sigma$ so $[\![E']\!]\sigma = \sum_{i\in p} x^i$. There must therefore be $p$ subterms of $E'$ (as each is a single power) which is a contradiction. □

Note that $C14.p$ is not in general sound for bisimulation, so this proof cannot be trivially adapted to show the non-axiomatisability of bisimulation over $*$-expressions.

Turning to positive results, complete infinite systems have been given by Krob in [Kro91] and Bloom and Ésik in [BÉ93a].

The latter contains another set of *commutative identities* which are slightly awkward to state without introducing a lot of specialized notation. We need to consider matrices of $*$-expressions, defining sequential composition in the obvious fashion and the $*$ of a square matrix inductively on its size:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^* \stackrel{\text{def}}{=} \begin{bmatrix} (A + B\cdot D^* \cdot C)^* & (A + B \cdot D^* \cdot C)^* \cdot B \cdot D^* \\ (D + C \cdot A^* \cdot B)^* \cdot C \cdot A^* & (D + C \cdot A^* \cdot B)^* \end{bmatrix}.$$

There is a commutative identity for each $m \geq n \geq 1$, surjective function $\rho : m \to n$, family of functions $\rho_i : m \to m \mid i \in m$ that respect $\rho$, i.e. satisfy

$$\forall i, j \in m \,.\, \rho(j) = \rho(\rho_i(j)),$$

and $n \times m$ matrix $A$ of expressions $A_{pi} \mid p \in n, i \in m$. We first define an $m \times m$ matrix $B$ by

$$B_{ij} \stackrel{\text{def}}{=} \sum_{k \in \rho_i^{-1}(j)} A_{\rho(i),k}$$

and an $n \times n$ matrix $C$ by

$$C_{pq} \stackrel{\text{def}}{=} \sum_{k \in \rho^{-1}(q)} A_{pk}.$$

Viewing $\rho$ as an $m \times n$ matrix with

$$\begin{aligned} \rho_{ip} &= 1 \text{ , if } \rho(i) = p \\ &= 0 \text{ , otherwise,} \end{aligned}$$

the commutative identity is the $n \times m$ equations of

$$B^* \cdot \rho \;=\; \rho \cdot C^*.$$

Bloom and Ésik showed that the axiom

$$1^* \;=\; 1$$

together with the two axioms above, the commutative identities, their duals (in a sense we leave undefined) and the semiring axioms $C1$–$10$ are complete.

The former contains the *group identities* proposed in [Con71]. For each finite group $\langle G, \circ, \_^{-1}, I\rangle$ there is an identity over variables $X_g \mid g \in G$. Suppose $G = \{0, \ldots, m-1\}$ and $I = 0$. We define an $m \times m$ matrix $M$ by

$$M_{ij} \;\stackrel{\text{def}}{=}\; X_{i^{-1} \circ j}.$$

The identity is then

$$[\,1 \;\; 0 \;\; \cdots \;\; 0\,] \cdot M^* \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \;=\; \left(\sum_{g \in G} X_g\right)^*,$$

which is the first equation of the commutative identity for $n = 1$, $\rho(i) = 0$, $\rho_i(j) = i \circ j$, $A_{1i} = X_i$.

Krob showed that these identities for all finite groups $G$ together with the axioms $C1$–$13$ and

$$\begin{aligned} (X \cdot Y)^* &= 1 + X \cdot (Y \cdot X)^* \cdot Y \\ (X + Y)^* &= (X^* \cdot Y)^* \cdot X^* \end{aligned}$$

are complete.

They are not all sound for bisimulation — taking $G$ to be the two element group $\{1, -1\}$ with multiplication gives the identity

$$(a + b \cdot a^* \cdot b)^* + a^* \cdot b \cdot (a + b \cdot a^* \cdot b)^* \;=\; (a + b)^*,$$

the left hand side of which can evolve via $\xrightarrow{b}$ to a state which is not successfully terminated, which the right hand side cannot.

A finite equational system has been given by Yanov [Con71, p.108] that is complete for $*$-expressions $E$ such that $E = E + 1$. Returning to bisimulation for a moment, a finite equational system has been given by Fokkink and Zantema [FZ94] for $*$-expressions of the form

$$E ::= a \mid 1 \mid E + E \mid E \cdot E \mid E^{\star}E.$$

It will be discussed in Chapter 4. It is not clear whether there is any precise connection between these results.

Two finite impure horn clause systems have been given by Salomaa in [Sal66], one of which has already been discussed. Finally, finite pure horn clause systems have been given by Arkhangelskii and Gorshkov [AG87], Boffa and Krob [Bof90, Kro91] and Kozen [Koz94]. Sample rules from these are

$$\begin{array}{c}(E_1 + E_2)^* \cdot E_3 = (F_1 + F_2)^* \cdot F_3 \\ \rightarrow \\ (E_1 + E_2)^* \cdot E_3 = (F_2 + F_1 \cdot E_1^* \cdot E_2)^* \cdot (F_3 + F_1 \cdot E_1^* \cdot E_3)\end{array}$$

$$E \cdot E = E \quad \rightarrow \quad E^* = 1 + E$$

$$E \cdot F + F = F \quad \rightarrow \quad E^* \cdot F + F = F$$

respectively.

### 2.3.5 Infinitary rules and Denotational models

There is a large body of other work involving axiomatisations over process calculi, some of which is mentioned here for completeness. Firstly there is work (e.g. [BBK87,BW90]) on ACP using the Approximation Induction Principle. A family of unary operators $\pi_n \mid n \in \omega$ is introduced with transitions

$$\frac{E \xrightarrow{a} F}{\pi_{n+1}(E) \xrightarrow{a} \pi_n(F).}$$

The AIP is then the implication

$$(\forall n \in \omega \,.\, \pi_n(E) = \pi_n(F)) \rightarrow E = F.$$

It is sound for bisimulation over finitely branching labelled transition systems and can be used to give finite complete axiomatisations for a large class of process calculi, e.g. those that are GSOS definable [ABV92]. Being infinitary it does not fit in any of the classes of axiomatisation considered in §2.2. In [BK84] there is a variant of Milner's system from [Mil84] (reproduced in §2.3.2) for a syntax close to our sequences of formal equations.

A number of axiomatisations for calculi of infinite state processes have been presented in a sequent style, for example in [CHM94,CHM93,Chr93,HS91,Hüt91]. It is not clear whether corresponding horn clause axiomatisations can be given, even allowing additional predicates such as 'guardedness' conditions.

In this thesis we are only concerned with axiomatising *equivalences*. Partial orders over processes have been introduced for several reasons — to match a notion of effective testing [Mil81], to allow for divergent [Wal88] and under-specified [CS90] processes and to allow a CPO-based denotational semantics [HP80,Hen81, Hen88,AH88,Abr91]. Walker gives in [Wal88] a complete system using a variant of the unique fixed point rule of [Mil84]. The CPO models admit infinitary rules such as $\omega$-*induction*. We recall a definition from [Hen88] for term models which requires a definition of the syntactic approximants to a recursive term:

$$\begin{aligned}
E^0 &\stackrel{\text{def}}{=} \bot \\
X^{n+1} &\stackrel{\text{def}}{=} X \\
(aE)^{n+1} &\stackrel{\text{def}}{=} aE^{n+1} \\
(E+F)^{n+1} &\stackrel{\text{def}}{=} E^{n+1} + F^{n+1} \\
(\mu X E)^{n+1} &\stackrel{\text{def}}{=} E^{n+1}[(\mu X E)^n / X]
\end{aligned}$$

It is then

$$(\forall n \in \omega \,.\, E^n \leq F) \rightarrow E \leq F$$

which again does not fit into any of the classes of axiomatisation considered in §2.2.

I expect that the following is known in the folklore but is perhaps worth pointing out in print. Bisimulation preorders such as those appearing in the above references work well for strong bisimulation but less so when one abstracts from certain actions. In particular any CPO equipped with continuous functions for $a, b, c, \tau$ and $+$ that satisfies axioms $A1$–4 and $\tau\tau X = \tau X$ will identify $\mu$-expressions that are not even weak-completed-trace equivalent, e.g. writing $\tau^\omega$ for $\mu X\ \tau X$ and the standard semantic function as $[\![\_]\!]$

$$[\![a\tau^\omega + a\tau b\tau c\tau^\omega]\!] = [\![a\tau^\omega + a\tau b\tau^\omega + a\tau b\tau c\tau^\omega]\!],$$

so when using such a model divergence can only be regarded as catastrophic.

# Chapter 3

# Axiomatisation over $\mu$-expressions

In this chapter we formulate and prove our main non-axiomatisability result, that there is no finite equational axiomatisation for bisimulation over $\mu$-expressions.

To state the result a precise definition of the admissible equational axiomatisations is required, preferably as large as possible. For a syntax with variable binding, such as the $\mu$-expressions, there does not seem to be a canonical definition. Consider for example the following three rules from [Mil84].

$$\begin{array}{llcl} A4 & E + E & = & E \\ B1 & \mu X E & = & E[\mu X E / X] \\ D & E = F[E/X] \wedge X \text{ guarded in } F & \rightarrow & E = \mu X F \end{array}$$

The last involves an implication and an additional predicate symbol — for either reason we would not call it an equational axiom. The first is formally an equation *scheme*, standing for the infinite set of equations obtained by replacing $E$ by each $\mu$-expression. Its schematic nature is harmless, however, as it could be replaced by the equation

$$X + X \quad = \quad X$$

with a standard inference rule of substitution. The second is again schematic but is not equivalent to any equation within the syntax of $\mu$-expressions. Indeed, to equationally express anything of interest about fixed points, such as the simple properties below, some notation for substitution is required.

$$\begin{array}{llcl}
 & \mu X E & = & \mu X E[E/X] \\
B1 & \mu X E & = & E[\mu X E\,/X] \\
E2 & \mu X E[F/X] & = & E[\mu X F[E/X]\,/X] \\
C_{12} & \mu X\ E[X,X,X,X] & = & \mu X\ E[X,X,\mu Y\ E[X,Y,X,Y],\mu Y\ E[X,Y,X,Y]]
\end{array}$$

To express these directly we could allow axioms containing substitutions and metavariables over terms ($E, F$) and variables ($X, Y$), together with a rather complicated rule for their instantiation. It would be awkward to characterise the sound axioms of this form, however, so instead we will embed the $\mu$-expressions in a simply typed lambda calculus and work up to $\beta\eta$ equality. Axioms such as the above can be written as equations containing variables of higher type rather than as equation schemes, with substitution appearing only in the rules defining $\beta\eta$ equality. This simplifies the technical development and also gives added significance to some of the intermediate results as the terms of higher type can be viewed as a fragment of a higher order process calculus.

The main theorem, stated in §3.2, asserts the nonexistence of finite axiomatisations containing at most first order variables. These axiomatisations may contain (the translations of) schemes such as those above. Generalising the result to axiomatisations containing variables of arbitrary type is discussed in Chapter 6.

## 3.1 Outline of chapter

The proof rests on the fact that finite axiomatisations only provide bisimulations of certain 'widths', which we illustrate for the first axiom above. Writing $a^n$ for $\overbrace{a\ldots a}^{n\ \text{times}}$, repeated use of $\mu X E = \mu X E[E/X]$ can change the length of a loop only by factors of 2, i.e. it can derive the 'internal' unfolding

$$\mu X a^n X = \mu X a^{2^k n} X$$

for any $k \geq 0$ but not

$$\mu X a^n X = \mu X a^{pn} X$$

for any prime $p > 2$. We show that for any finite set of axioms there is some bound corresponding to this '2'.

The details are somewhat lengthy (occupying the remainder of this chapter) but may be of some independent interest. In particular, notions of higher order transition system and bisimulation are given which might be interesting when considering richer higher order calculi. We show a finite-state property and hence decidability of higher order bisimulation at arbitrary type.

The intermediate results can be applied to give an easy proof of the nonexistence of finite axiomatisations of bisimulation over calculi of $*$-expressions containing a zero process. This is done in Chapter 4, where it is related to the positive result of [FZ94] for BPA$^*$ (which has no zero).

In the next section we define the lambda calculus used and state the nonexistence result precisely. We then define an extensional equivalence $\sim_{ext}$ over open terms that contains all sound equations and give a useful alternative characterisation of it. In §3.3 higher order transition systems and higher order bisimulation are defined and in §3.4 we show that all processes are 'finite state' in the appropriate sense and hence that higher order bisimulation is decidable. An inductive characterisation of the transitions of a substituted term is given in §3.5. In §3.6 higher order bisimulation and $\sim_{ext}$ are shown to coincide over the base type terms containing first order variables (the $T_1^P$ defined in the next section) — hence all sound axioms lie within finite higher order bisimulations. Finally in §3.7 we construct from these bisimulations an equivalence over terms which is preserved by all proofs. A pair of bisimilar but inequivalent terms is then picked out.

## 3.2 Basic definitions

From now on we shall be considering terms of a simply typed lambda calculus with a single base type $P$ of processes and the following constants:

$$\begin{aligned}
&0 : P \\
&a : P \to P \ \text{ for each } a \in Act \\
&+ : P \to P \to P \\
&\mathit{fix} : (P \to P) \to P
\end{aligned}$$

We will usually work up to $\beta\eta$ equality, using abstraction to allow parameterised equations. This is in contrast to taking $\beta$-reduction to be of comparable computational interest to the labelled transitions, e.g. in the work of Nielson [Nie89].

Notation and definitions will be taken from [Mit90]. In particular we write typing judgements and typed equations as $\Gamma \rhd E : \sigma$ and $\Gamma \rhd E = F : \sigma$, where $E, F$ are terms, $\sigma$ is a type, $\Gamma$ an assignment of types to a finite set of variables. The inference rules for typing and $\beta\eta$ equality are reproduced in Figure 3–1. It is presumed wherever $\Gamma, x : \alpha$ appears that $x$ does not occur in $\Gamma$ and further that the terms appearing in a typed equation $\Gamma \rhd E = F : \sigma$ are both typable with type $\sigma$ in $\Gamma$. There is an unfortunate clash of notation between the $\rhd$ of typing judgments and the visibility predicate $\rhd$. As both are standard we make only a small typographical distinction — it should be clear from context which is intended. If $\mathcal{E}$ a set of typed equations we write $\mathcal{E} \vdash \Gamma \rhd E = F : \sigma$ to mean that $\Gamma \rhd E = F : \sigma$ is derivable from $\mathcal{E}$ using the rules in Figure 3–1 — formally, that it is derivable by those rules augmented by

$$\frac{}{\Gamma \rhd E = F : \sigma} \; (\Gamma \rhd E = F : \sigma) \in \mathcal{E}.$$

We sometimes elide existentially quantified $\Gamma$.

The order of a type is as usual:

$$\begin{array}{rcl} \mathrm{order}(P) & = & 0 \\ \mathrm{order}(\sigma \to \tau) & = & \max\{1 + \mathrm{order}(\sigma), \mathrm{order}(\tau)\}. \end{array}$$

We take some type assignment $\mathcal{K}$ with a countable infinity of variables at each type. For $k \in \omega \cup \{\omega\}$ we write $T_k^\sigma$ for the set of terms $E$ for which there is some finite $\Gamma \subseteq \mathcal{K}$, containing only variables of order $\leq k$, such that $\Gamma \rhd E : \sigma$. In particular $T_0^\sigma$ contains terms of type $\sigma$ whose only free variables are of type $P$.

There is an obvious bijection between the $\mu$-expressions and the long $\beta\eta$ normal forms in $T_0^P$ with for example

$$\mu X aY + X \leftrightarrow \mathit{fix}\ \lambda x : P.\ +\ (ay)\ (x).$$

For any equivalence over $\mu$-expressions this induces an equivalence over $T_0^P$, closing under $\beta\eta$ equality.

**Definition** An *axiomatisation* for an equivalence $\sim$ over $\mu$-expressions is a set $\mathcal{E}$ of typed equalities that is sound, i.e.

$$\forall E, F \in T_0^P\ .\ \mathcal{E} \vdash E = F : P \Rightarrow E \sim F$$

$$\frac{}{x:\sigma \triangleright x:\sigma}\ \mathit{var} \qquad \frac{}{\{\} \triangleright c:\sigma}\ \mathit{cst}$$

$$\frac{\Gamma \triangleright M:\sigma}{\Gamma, x:\tau \triangleright M:\sigma}\ \mathit{addhyp}$$

$$\frac{\Gamma, x:\sigma \triangleright M:\tau}{\Gamma \triangleright (\lambda x:\sigma.\ M):\sigma \to \tau} \to \mathit{Intro} \qquad \frac{\Gamma \triangleright M:\sigma \to \tau \quad \Gamma \triangleright N:\sigma}{\Gamma \triangleright MN:\tau} \to \mathit{Elim}$$

$$\frac{\Gamma \triangleright M = N:\sigma}{\Gamma, x:\tau \triangleright M = N:\sigma}\ \mathit{addhyp} \qquad \frac{}{\Gamma \triangleright M = M:\sigma}\ \mathit{ref}$$

$$\frac{\Gamma \triangleright M = N:\sigma}{\Gamma \triangleright N = M:\sigma}\ \mathit{sym} \qquad \frac{\Gamma \triangleright L = M:\sigma \quad \Gamma \triangleright M = N:\sigma}{\Gamma \triangleright L = N:\sigma}\ \mathit{tran}$$

$$\frac{\Gamma, x:\sigma \triangleright M = N:\tau}{\Gamma \triangleright (\lambda x:\sigma.\ M) = (\lambda x:\sigma.\ N):\sigma \to \tau}\ \xi \qquad \frac{\Gamma \triangleright M = M':\sigma \to \tau \quad \Gamma \triangleright N = N':\sigma}{\Gamma \triangleright MN = M'N':\tau}\ \mu$$

$$\frac{y \notin FV(M)}{\Gamma \triangleright \lambda x:\sigma.\ M = \lambda y:\sigma.\ M[y/x]:\sigma \to \tau}\ \alpha$$

$$\frac{}{\Gamma \triangleright (\lambda x:\sigma.\ M)N = M[N/x]:\tau}\ \beta \qquad \frac{x \notin FV(M)}{\Gamma \triangleright \lambda x:\sigma.\ (Mx) = M:\sigma \to \tau}\ \eta$$

**Figure 3–1:** Lambda calculus

and complete, i.e.

$$\forall E, F \in T_0^P \,.\, \mathcal{E} \vdash E = F : P \Leftarrow E \sim F.$$

We assume w.l.g. that $\mathcal{E}$ contains only equalities at type $P$ (otherwise they can be applied to new variables of appropriate types).

Note that if $\sim$ is not both a congruence for $a, +, \mu X$ and substitutive (i.e. $E \sim F \Rightarrow E[G/x] \sim F[G/x]$) then there can be no axiomatisation in this sense. Further, bisimulation is a substitutive congruence.

Some candidate axioms (corresponding to the axiom schemes given earlier) are given below, taking $\Gamma = \{e : P \to P,\ f : P \to P,\ z : P \to P \to P \to P \to P\}$.

$$\Gamma \rhd \mathit{fix}\ e = \mathit{fix}\ \lambda x : P.\ e(ex) : P$$

$B1$ $\quad \Gamma \rhd \mathit{fix}\ e = e(\mathit{fix}\ e) : P$

$E2$ $\quad \Gamma \rhd \mathit{fix}\ \lambda x : P.\ e(fx) = e(\mathit{fix}\ \lambda x : P.\ f(e(x))) : P$

$C_{12}$ $\quad \Gamma \rhd \mathit{fix}\ \lambda x : P.\ zxxxx = \mathit{fix}\ \lambda x : P.\ zxx(\mathit{fix}\ \lambda y : P.\ zxyxy)(\mathit{fix}\ \lambda y : P.\ zxyxy) : P$

They can all be shown sound for bisimulation by using Theorem 8 below. Clearly any of the axioms of the infinite system given in §2.3.1 could be written in this form.

We define the order of a finite set of typed equalities to be the maximum order of any variable (free or bound) therein. The main theorem can now be stated.

**Theorem 6** *If $Act$ is non-empty there is no finite axiomatisation of order $\leq 1$ for any substitutive congruence $\sim$ finer than (or identical to) bisimulation that for all $n \geq 1$ satisfies*

$$\mu X aX \sim \mu X a^n X.$$

The restriction to axioms of order $\leq 1$ simplifies the problem considerably, as we will discuss later. We have yet to prove the result without it but note that the candidate axioms above are all of order $1$, supporting the view that it admits an interestingly rich class of axioms.

In the rest of this section we characterise the sound axioms.

**Definition** If $\rho$ is a substitution with a finite domain that includes the free variables of $E_1, E_2, \ldots$ and for all $x : \sigma$ in its domain $\rho(x) \in T_0^\sigma$ then we say that $\rho$ is a substitution for $E_1, E_2, \ldots$.

**Definition** If $\sim$ is a substitutive congruence over $\mu$-expressions the relation $\sim_{ext}$ over $T^P_\omega$ is given by

$$E \sim_{ext} F \quad \textit{iff} \quad \text{for all substitutions } \rho \text{ for } E, F \text{ we have } E\rho \sim F\rho.$$

Sound axioms clearly lie within $\sim_{ext}$. For the converse we give another characterisation.

**Definition** If $\mathcal{E} = \{\, \Gamma_i \triangleright M_i = N_i : P \mid i \in I \,\}$ is a set of typed equations then the relation $=_\mathcal{E}$ over $T^P_0$ is the smallest equivalence such that:

- If $\rho$ is a substitution for $M_i, N_i$ then $M_i\rho =_\mathcal{E} N_i\rho$.
- $=_\mathcal{E}$ is a congruence for contexts of the form

$$C ::= \_ \mid x \mid 0 \mid aC \mid C + C \mid \mathit{fix}\ \lambda x : P.\ C.$$

- $=_\mathcal{E}$ is closed under $\beta\eta$ equivalence.

**Proposition 30** *If $M, N \in T^P_0$ then $\mathcal{E} \vdash M = N : P$ iff $M =_\mathcal{E} N$.*

**Corollary 31** *If $\sim$ is a substitutive congruence over $\mu$-expressions then a set of typed equations $\mathcal{E}$ is sound for $\sim$ iff $\mathcal{E} \subseteq \sim_{ext}$.*

Proof (of Proposition 30)

We need two facts about derivations of typing and typed equality judgements:

**Lemma 32** *If $\mathcal{E} \vdash \Gamma \triangleright M = N : \sigma$ then there is a derivation in which all instances of $\mu$ have an instance of $\mathit{ref}$ as one premise and all rules below an instance of $\mathit{tran}$ are themselves instances of $\mathit{tran}$.*

Proof We write derivations of typed equalities in a linear syntax, using the rule names, not in tree form. The proof is via two rewrite systems over derivations of typed equalities. The first, generated by the rule

$$\mu(d, e) \quad \mapsto \quad \mathit{tran}(\mu(d, \mathit{ref}()), \mu(\mathit{ref}(), e)), \text{ if } d \neq \mathit{ref}() \neq e$$

clearly terminates as it reduces the number of redexes and the normal forms satisfy the first condition. The second, generated by the rules

$$\begin{array}{rcl}
\mu(tran(d,e), ref()) & \mapsto & tran(\mu(d, ref()), \mu(e, ref())) \\
\mu(ref(), tran(d,e)) & \mapsto & tran(\mu(ref(), d), \mu(ref(), e)) \\
add_hyp(tran(d,e)) & \mapsto & tran(add_hyp(d), add_hyp(e)) \\
sym(tran(d,e)) & \mapsto & tran(sym(d), sym(e)) \\
\xi(tran(d,e)) & \mapsto & tran(\xi(d), \xi(e)),
\end{array}$$

terminates as it reduces the total number of rules below each instance of $tran$. It introduces no redexes of the first so the normal forms satisfy both conditions of the lemma. □

**Lemma 33** *If $\Gamma, z:\sigma \triangleright G:\tau$, $\{\} \triangleright M:\sigma$ and $\{\} \triangleright N:\sigma$ then there exists some $n$ and $G_j \mid j \in 1..n$ such that:*

- $\Gamma, z:\sigma \triangleright G_j:\tau$.
- *$G_j$ is in long $\beta\eta$ normal form and contains at most one free occurrence of $z$.*
- *$G[M/z] =_{\beta\eta} G_1[M/z]$, $G_n[N/z] =_{\beta\eta} G[N/z]$ and $\forall j \in 1..n-1$ . $G_j[N/z] =_{\beta\eta} G_{j+1}[M/z]$.*

PROOF Let $H_1$ be the normal form of $G$ and define $H_{k+1}$ by substituting $N$ for the leftmost occurrence of $z$ in $H_k$ and taking the normal form. Let $n$ be the least such that $H_n$ has at most one occurrence of $z$ ($n$ must exist as otherwise $G[N/z]$ is not strongly normalising) and for $i \in 1..n$ substitute $M$ for all except the leftmost occurrence of $z$ in $H_i$ and let $G_i$ be the normal form thereof. □

Now we can prove Proposition 30. The right-to-left implication is trivial. For the left-to-right we suppose that $d$ is a derivation of $\mathcal{E} \vdash \Gamma \triangleright M = N : P$ in the form of Lemma 32 and show $M =_{\mathcal{E}} N$ by induction on $d$. Either the bottom rule in $d$ is an instance of $tran$ and the induction hypothesis can be used or $d$ contains no instances of $tran$. The unique top rule in $d$ is then either one of $ref, \alpha, \beta, \eta$ in which case $M =_{\beta\eta} N$ and so $M =_{\mathcal{E}} N$, or an axiom $\Gamma_i \triangleright M_i = N_i : P$. Suppose $\Gamma_i = x_1:\sigma_1, \ldots, x_n:\sigma_n$. There is a term $G$ and new variable $z$ such that

$$\Gamma, z:\sigma_1 \to \cdots \sigma_n \to P \triangleright G:P$$

$$\begin{aligned} G[\lambda x_1 : \sigma_1. \ \cdots \lambda x_n : \sigma_n. \ M_i \, / z] &=_{\beta\eta} M \\ G[\lambda x_1 : \sigma_1. \ \cdots \lambda x_n : \sigma_n. \ N_i \, / z] &=_{\beta\eta} N. \end{aligned}$$

It now suffices to use Lemma 33 and note that $G_j[\lambda x_1 : \sigma_1. \ \cdots \lambda x_n : \sigma_n. \ M_i \, / z] =_{\mathcal{E}} G_j[\lambda x_1 : \sigma_1. \ \cdots \lambda x_n : \sigma_n. \ N_i \, / z]$. □

We note that, taking $\sim$ to be bisimulation, $\sim_{ext}$ is a congruence in the following sense.

**Proposition 34** *If $E \sim_{ext} F$ and $E' \sim_{ext} F'$ then $aE \sim_{ext} aF$, $+EE' \sim_{ext} +FF'$ and fix $\lambda x : P.\ E \sim_{ext}$ fix $\lambda x : P.\ F$.*

PROOF Straightforward, using the analogous properties of $\sim$ from Proposition 7. □

## 3.3 Higher order bisimulation

To show a non-axiomatisability result we need to capture some limitation of any finite set of sound axioms. To do this a more intensional characterisation of $\sim_{ext}$, without its quantification over all substitutions, is required. A suitable equivalence is defined below in two steps. Firstly a notion of higher order transition system is given and then higher order bisimulation (written $\sim_{ho}$) is defined over it. We generalise the definition of bisimulation over open $\mu$-expressions in [Mil84], reproduced in §2.1.1. There (where variables are all of 'type $P$') an extended transition system is defined over $\mu$-expressions consisting of the usual labelled transitions together with predicates $E \triangleright X$ (pronounced '$E$ sees $X$' or '$X$ is visible in $E$'). An extended bisimulation then requires matching of visibilities at each state, as well as transitions, and can be shown equal to the relevant special case of $\sim_{ext}$. The essential fact about these visibility predicates is that the transitions of a $\mu$-expression $E[F/X]$ can be calculated from the visibilities of $E$ and the transitions of $E$ and $F$. In this work variables may be of higher type and so applied to arguments (which may themselves be of higher type), so we need a more sophisticated visibility predicate to get an analogous result. The restriction to low order variables is not yet needed so this is all still over terms with free variables of arbitrary type, i.e. the terms $T^P_\omega$.

**Notation** From now on we take all terms mentioned to be in normal form unless stated otherwise and work up to alpha equivalence when convenient. We write $+$ infix except when emphasizing the distinction between lambda calculus terms and others. Bisimulation over $T_0^P$ will be written $\sim$.

**Definition** Take the relations $\xrightarrow{a} \mid a \in Act$ and $\rhd$ to be the least over normal forms in $T_\omega^P$ such that

$$\frac{}{aE \xrightarrow{a} E} \qquad\qquad \frac{}{E \rhd E}$$

$$\frac{E \xrightarrow{a} E'}{E + F \xrightarrow{a} E'} \text{ and sym.} \qquad\qquad \frac{E \rhd E'}{E + F \rhd E'} \text{ and sym.}$$

$$\frac{E \xrightarrow{a} E'}{\mathit{fix}\ \lambda y : P.\ E \xrightarrow{a} E'[\mathit{fix}\ \lambda y : P.\ E\,/y]} \quad \frac{E \rhd E'}{\mathit{fix}\ \lambda y : P.\ E \rhd E'[\mathit{fix}\ \lambda y : P.\ E\,/y]}$$

We will usually only be concerned with visibility judgments of the form $E \rhd x\tilde{F}$. For example consider the term

$$E \stackrel{\text{def}}{=} \mathit{fix}\ \lambda y : P.\ by + x(zx + y)$$

with $\{x : P \to P,\ z : (P \to P) \to P\} \rhd E : P$. We have

$$\begin{array}{rl} & E \xrightarrow{b} E \\ \text{and} & E \rhd x(zx + E). \end{array}$$

There are a number of simple properties of the transition and visibility relations that will be used without comment.

**Lemma 35** *For all $E, F, G \in T_\omega^P$ and substitutions $\rho$:*

- *If $E \xrightarrow{a} F$ then $E\rho \xrightarrow{a} F\rho$.*
- *If $E \rhd F$ then $E\rho \rhd F\rho$.*
- *If $E \rhd F \xrightarrow{a} G$ then $E \xrightarrow{a} G$.*
- *If $E \rhd F \rhd G$ then $E \rhd G$.*

- *If $E \xrightarrow{a} F$ then $FV(F) \subseteq FV(E)$.*

- *If $E \rhd F$ then $FV(F) \subseteq FV(E)$.*

PROOF Straightforward inductions on the derivations of the judgments. $\square$

Higher order bisimulation is defined using this transition system.

**Definition** If $R$ is a relation on $T_\omega^P$ then the typed relations $\hat{R}^\sigma$ over $T_\omega^\sigma$ are given by

- $\hat{R}^P = R$

- $E\ \hat{R}^{\sigma \to \tau}\ F$ iff for all $x : \sigma \in \mathcal{K}$, if $x$ is not free in $E, F$ then $Ex\ \hat{R}^\tau\ Fx$.

Note that we are writing $Ex$ for its normal form, according to the convention above.

**Notation** We abbreviate the application $((xE_1)\ldots)E_n$ by $x\tilde{E}$, in which $x$ is typically assumed to be of a type $\sigma_1 \to \ldots \sigma_n \to P$.

**Definition** A relation $R$ over $T_\omega^P$ is a higher order bisimulation if $E\ R\ F$ implies

- If $E \xrightarrow{a} E'$ then $\exists F'\ .\ F \xrightarrow{a} F' \wedge E'\ R\ F'$.

- If $E \rhd x\tilde{E}$ then $\exists \tilde{F}\ .\ F \rhd x\tilde{F} \wedge \forall i\ .\ E_i\ \hat{R}^{\sigma_i}\ F_i$.

and symmetrically. We write $\sim_{ho}$ for the union of all higher order bisimulations and $\sim_{ho}^\sigma$ for $\hat{\sim_{ho}}^\sigma$.

**Proposition 36** *$\sim_{ho}$ is a higher order bisimulation and is an equivalence relation.*

PROOF Straightforward except for transitivity where Proposition 38 is needed. $\square$

For example, if

$$
\begin{aligned}
E &\stackrel{\text{def}}{=} \mathit{fix}\ \lambda x : P.\ y(yx) \\
F &\stackrel{\text{def}}{=} \mathit{fix}\ \lambda x : P.\ yx,
\end{aligned}
$$

then $E \sim_{ho} F$ is shown by taking a relation $\{\langle E, F\rangle, \langle yE, F\rangle\}$, the only relevant visibilities or transitions being

$$\begin{array}{ll} E \rhd y(y(E)) & \qquad F \rhd yF \\ yE \rhd yE & \qquad F \rhd yF. \end{array}$$

The terms in the $T^\sigma_\omega$ can be viewed as a small fragment of the higher order $\pi$ calculus of [San93], taking a single object sort $s \mapsto ()$ and the agents without parallel composition, restriction, infinitary sum, matching, variables of sorts containing $s$ or infinitely many/higher order defined constants. This is discussed further in §4.2.

## 3.4 Finite state and decidability properties

As we are working with a mild generalization of finite state processes it is to be expected that all higher order bisimulations between them are in some sense finitely generated. In the sequel we need only the following result for terms in $T^P_1$.

**Definition** The *derivatives* of a term $E \in T^P_1$ are $\mathrm{der}(E) \overset{\text{def}}{=} \{ F \in T^P_1 \mid E {\rightsquigarrow}^* F \}$, where $\rightsquigarrow$ is the least relation over $T^P_1$ such that

$$\begin{array}{rcl} E \longrightarrow E' & \Rightarrow & E \rightsquigarrow E' \\ E \rhd x\tilde{E} & \Rightarrow & \forall i \;.\; E \rightsquigarrow E_i. \end{array}$$

This generalises the standard definition for transition systems.

**Lemma 37** *If $E \in T^P_1$ then $\mathrm{der}(E)$ is finite. Further if $E \sim_{ho} F$ then there is a higher order bisimulation contained in $\mathrm{der}(E) \times \mathrm{der}(F)$ relating them.*

PROOF This is a special case of Corollary 46 and Proposition 47 below. □

In the rest of this section we show a generalization of this result for terms in $T^P_\omega$ containing free variables of arbitrary types and hence the decidability of higher order bisimulation for arbitrary terms. This is slightly subtle as any higher order bisimulation between non-trivial terms in $T^P_\omega$ must be infinite. We will need to know that higher order bisimulation preserves sets of free variables.

**Proposition 38** *If $E \sim_{ho} F$ then $FV(E) = FV(F)$.*

PROOF We show that free variables are accessible via the transition system in a sense that is preserved by higher order bisimulation.

**Notation** $\mathcal{V}$ ranges over sets of variables such that each $\mathcal{V}^\sigma$ and $\mathcal{K}^\sigma - \mathcal{V}^\sigma$ is infinite.

**Definition** For a set $\mathcal{V}$ of variables $\rightsquigarrow_\mathcal{V}$ is the least relation over normal forms in $T^P_\omega$ such that

$$\begin{aligned} E \longrightarrow E' \quad &\Rightarrow \quad E \rightsquigarrow_\mathcal{V} E' \\ E \rhd x\tilde{E} \wedge \tilde{z} \cap (FV(E_i) \cup \mathcal{V}) = \{\} \quad &\Rightarrow \quad E \rightsquigarrow_\mathcal{V} E_i\tilde{z}. \end{aligned}$$

**Lemma 39** *If $E \rightsquigarrow_\mathcal{V} F$ then for all terms $G$, variables $w \in \mathcal{V}^P$ and substitutions $\rho$ with $\mathrm{dom}(\rho) \cup \mathrm{ran}(\rho) \subseteq \mathcal{K} - \mathcal{V}$ there exists $\rho'$ with $\mathrm{dom}(\rho') \cup \mathrm{ran}(\rho') \subseteq \mathcal{K} - \mathcal{V}$ such that $E\rho[G/w] \rightsquigarrow_\mathcal{V} F\rho'[G/w]$.*

PROOF If $E \xrightarrow{a} F$ we can take $\rho' = \rho$, otherwise $E \rhd y\tilde{E}$ and $F = E_i\tilde{z}$ with $\tilde{z} \cap (FV(E_i) \cup \mathcal{V}) = \{\}$. Take some $\tilde{z}'$ such that $\tilde{z}' \cap (FV(E_i) \cup \mathcal{V} \cup FV(E_i\rho[G/w]) \cup \mathrm{dom}(\rho)) = \{\}$ and put $\rho' = [\tilde{z}'/\tilde{z}] \circ \rho$. □

**Corollary 40** *If $E \rightsquigarrow^n_\mathcal{V} \rhd x\tilde{E}$, $x \in \mathcal{V}$ and $\rho$ is a substitution with $\mathrm{dom}(\rho) \cup \mathrm{ran}(\rho) \subseteq \mathcal{K} - \mathcal{V}$ then there exists $\tilde{E}'$ such that $E\rho \rightsquigarrow^n_\mathcal{V} \rhd x\tilde{E}'$.*

**Corollary 41** *If $E \rightsquigarrow^*_\mathcal{V} \rhd x\tilde{E}$ and $\{x, y\} \subseteq \mathcal{V}$ then there exists $\tilde{E}'$ such that $\mathit{fix}\ \lambda y : P.\ E \rightsquigarrow^*_\mathcal{V} \rhd x\tilde{E}'$.*

PROOF Both are simple inductions along $\rightsquigarrow^n_\mathcal{V}$. □

**Lemma 42** $x \in FV(E)$ *iff* $\forall \mathcal{V} \,.\, x \in \mathcal{V} \Rightarrow \exists \tilde{E} \,.\, E \rightsquigarrow^*_\mathcal{V} \rhd x\tilde{E}$.

PROOF $\Rightarrow$: By induction on $E$, using Corollary 41 for the $\mathit{fix}$ case. $\Leftarrow$: If $E \rightsquigarrow^*_\mathcal{V} F$ then the free variables of $F$ are either present in $E$ or not in $\mathcal{V}$. □

Suppose $x \in FV(E)$. Take any $\mathcal{V}$ containing $FV(E) \cup FV(F)$. By Lemma 42 $E \rightsquigarrow^n_\mathcal{V} \rhd x\tilde{E}$. By induction on $n$ we can show $F \rightsquigarrow^n_\mathcal{V} \rhd x\tilde{F}$ (using Corollary 40) then by Lemma 42 $x \in FV(F)$. □

We give an alternative definition of higher order bisimulation that is *not* necessarily infinite.

**Definition** If $R$ is a relation on $T_\omega^P$ then the typed relations $\overline{R}^\sigma$ over $T_\omega^\sigma$ are given by

- $\overline{R}^P = R$
- $E\ \overline{R}^{\sigma\to\tau}\ F$ iff there exists $x:\sigma \in \mathcal{K}$ that is not free in $E, F$ such that $Ex\ \overline{R}^\tau\ Fx$.

**Definition** Such an $R$ is a *loose* higher order bisimulation if $E\ R\ F$ implies

- If $E \xrightarrow{a} E'$ then $\exists F'\ .\ F \xrightarrow{a} F' \wedge E'\ R\ F'$.
- If $E \rhd x\tilde{E}$ then $\exists \tilde{F}\ .\ F \rhd x\tilde{F} \wedge \forall i\ .\ E_i\ \overline{R}^{\sigma_i}\ F_i$.

and symmetrically.

A loose higher order bisimulation generates a higher order bisimulation as follows.

**Definition** If $R$ is a relation on $T_\omega^P$ then $E\ Cl(R)\ F$ iff there exists an injective type-respecting substitution $\rho : FV(E,F) \to \mathcal{K}$ such that $E\rho\ R\ F\rho$.

**Proposition 43** *If $R$ is a loose higher order bisimulation then $Cl(R)$ is a higher order bisimulation.*

PROOF

> **Lemma 44** *If $E\ \overline{R}^\sigma\ F$ then for all injective $\rho : FV(E,F) \to \mathcal{K}$ we have $E\rho\ \widehat{Cl(R)}^\sigma\ F\rho$.*
>
> PROOF By induction on the type $\sigma$. □

We now check that $Cl(R)$ is a higher order bisimulation. Suppose $E\ Cl(R)\ F$, i.e. for some $\rho$ that $E\rho\ R\ F\rho$. If $E \xrightarrow{a} E'$ then $E\rho \xrightarrow{a} E'\rho$ so (as $R$ a loose higher order bisimulation) there is some $B$ such that $F\rho \xrightarrow{a} B\ R^{-1}\ E'\rho$. Now $\rho$ is invertible so $F = F\rho\rho^{-1} \xrightarrow{a} B\rho^{-1}$ and moreover $E'\ Cl(R)\ B\rho^{-1}$.

Suppose $E \rhd x\tilde{E}$, then $E\rho \rhd \rho(x)(\tilde{E}\rho)$ so $F\rho \rhd \rho(x)\tilde{B}$ with $\forall i \,.\, E_i\rho \; \overline{R}^{\sigma_i} \; B_i$. Now $F \rhd x(\tilde{B}\rho^{-1})$ and by Lemma 44 $\forall i \,.\, E_i \; \widehat{Cl(R)}^{\sigma_i} \; B_i\rho^{-1}$. □

Given terms $E, F$ and a higher order bisimulation $R$ with $E \; R \; F$ we shall now show that it is possible to make particular choices of their derivatives such that these are finite and $R$ restricted to them is a loose higher order bisimulation.

We suppose some total order, isomorphic to the naturals, on each $\mathcal{K}^\sigma$ and that $\mathcal{V} \subseteq \mathcal{K}$ is a set of variables with each $\mathcal{V}^\sigma$ and $\mathcal{K}^\sigma - \mathcal{V}^\sigma$ infinite. We let $f$ range over computable type-respecting partial functions $\mathcal{K} \to \mathcal{P}_{\text{fin}}\mathcal{K}$ with finite and computable domain. They are extended to total functions $f : \mathcal{P}_{\text{fin}}\mathcal{K} \to \mathcal{P}_{\text{fin}}\mathcal{K}$ by

$$f(A) \stackrel{\text{def}}{=} (A - \text{dom}(f)) \cup \bigcup \{\, f(a) \mid a \in A \cap \text{dom}(f) \,\}.$$

The function with empty domain will be written $\{\}$. These functions will be used to calculate the free variables of subterms in their original contexts.

**Definition** The full application of a term is given by

$$\begin{aligned} E *_f \;&\stackrel{\text{def}}{=}\; E \text{ for } E : P \\ E *_f \;&\stackrel{\text{def}}{=}\; (Ex) *_f \text{ for } E : \sigma \to \tau, \end{aligned}$$

where $x$ is the first variable in $(\mathcal{K}^\sigma - \mathcal{V}^\sigma) - f(FV(E))$.

**Definition** The $f$-derivatives of a term $E \in T^P_\omega$ are $\text{der}_f(E) \stackrel{\text{def}}{=} \{\, F \in T^P_\omega \mid E {\rightsquigarrow}^*_f F \,\}$ where $\rightsquigarrow_f$ is the least relation over $T^P_\omega$ such that

$$\begin{aligned} E \longrightarrow E' \;&\Rightarrow\; E \rightsquigarrow_f E' \\ E \rhd x\tilde{E} \;&\Rightarrow\; \forall i \,.\, E \rightsquigarrow_f E_i *_f . \end{aligned}$$

The proper $f$-derivatives of $E$ are $\text{der}^+_f(E) \stackrel{\text{def}}{=} \{\, F \in T^P_\omega \mid E {\rightsquigarrow}^+_f F \,\}$.

**Proposition 45** *If $R$ is a higher order bisimulation with $E \; R \; F$ then $Q \stackrel{\text{def}}{=} R \cap (\text{der}_{\{\}} E \times \text{der}_{\{\}} F)$ is a loose higher order bisimulation.*

Proof We can show $E \; \hat{R}^\sigma \; F \Rightarrow E *_{\{\}} \; R \; F *_{\{\}}$ and $E *_{\{\}} \; Q \; F *_{\{\}} \Rightarrow E \; \overline{Q}^\sigma \; F$ by induction on types using Proposition 38. The result is then straightforward. □

**Corollary 46** *$E \sim_{ho} F$ iff there is a loose higher order bisimulation contained in $\text{der}_{\{\}} E \times \text{der}_{\{\}} F$.*

PROOF Immediate from Propositions 43 and 45. $\square$

**Proposition 47** *The set of $f$-derivatives $\mathrm{der}_f(E)$ of a term is finite and computable.*

PROOF

**Lemma 48** *If $x \in \mathcal{V}$ then $\forall F \in T^\sigma_\omega \,.\, F *_{f \oplus x \mapsto f(FV(E)-\{x\})} [\mathit{fix}\ \lambda x : P.\ E\,/x] = F[\mathit{fix}\ \lambda x : P.\ E\,/x] *_f$.*

PROOF Induction on $\sigma$. $\square$

**Lemma 49** *If $x \in \mathcal{V}$ then $\forall n \,.\, E[\mathit{fix}\ \lambda x : P.\ E\,/x] \leadsto^{\leq n}_f F \iff \exists E' \,.\, E \leadsto^{\leq n}_{f \oplus x \mapsto f(FV(E)-\{x\})} E' \wedge E'[\mathit{fix}\ \lambda x : P.\ E\,/x] = F$.*

PROOF Induction on $n$. $\square$

It is now straightforward to show by induction on $E$ that $\forall f \,.\, \mathrm{der}^+_f(E)$ is finite and computable. $\square$

**Corollary 50** *$\sim_{ho}$ is decidable.*

PROOF By Corollary 46, to decide $E \sim_{ho} F$ we need only consider the set of relations contained in $\mathrm{der}_{\{\}}\, E \times \mathrm{der}_{\{\}}\, F$. This set is finite and computable and moreover it is clearly decidable whether any element of it is a loose higher order bisimulation. $\square$

We note that $\sim_{ho}$ is also a congruence in a limited sense, although this fact will not be used until Chapter 6.

**Proposition 51**

1. *If $E \sim_{ho} F$ and $E' \sim_{ho} F'$ then $aE \sim_{ho} aF$, $+EE' \sim_{ho} +FF'$ and $\mathit{fix}\ \lambda x : P.\ E \sim_{ho} \mathit{fix}\ \lambda x : P.\ F$.*

2. *If $E \sim^\tau_{ho} F$ and $x : \sigma$ then $\lambda x : \sigma.\ E \sim^{\sigma \to \tau}_{ho} \lambda x : \sigma.\ F$.*

3. *If $E \sim^\sigma_{ho} F$ and $x : \sigma \to \tau$ then $xE \sim^\tau_{ho} xF$.*

4. *If $E \sim^{\sigma \to \tau}_{ho} F$ and $x : \sigma \notin FV(E) \cup FV(F)$ then $Ex \sim^\tau_{ho} Fx$.*

PROOF The only interesting part is the *fix* case of 1, for which it suffices to check that

$$\{\, E[\mathit{fix}\,\lambda x : P.\ G\,/x], F[\mathit{fix}\,\lambda x : P.\ G\,/x] \mid E, F, G \in T^P_\omega \text{ and } E \sim_{ho} F \,\}$$

is a loose higher order bisimulation. $\square$

## 3.5 The transition system of a substituted term

In this section we consider the transitions of a substituted term such as $E\rho$. The transition and visibility predicates are related by the following.

**Lemma 52** *If $E \in T^P_\omega$, $\rho$ is a substitution for $E$ and $E\rho \xrightarrow{a} A$ then either $E \xrightarrow{a} E'$ and $E'\rho = A$ or $E \rhd x\tilde{E}$ and $\rho(x)(\tilde{E}\rho) \xrightarrow{a} A$.*

PROOF Induction on the derivation of $E\rho \xrightarrow{a} A$. $\square$

In general there will be a complex pattern of $\beta$ reduction involved in reducing the $\rho(x)(\tilde{E}\rho)$ term appearing above to normal form. If $E \in T^P_1$ and $E$ and $\mathrm{ran}(\rho)$ are all in normal form, however, it is simple, allowing a direct inductive characterisation of the transitions of $E\rho$. The rest of this section is devoted to giving that characterisation.

For the remainder of the chapter we consider only terms in $T^P_1$, i.e. terms of type $P$ containing at most first order variables. We discuss whether this restriction can be removed in §6.1.

We consider a substitution $\rho$ with a finite domain containing at most first order variables and a range with variables only of type $P$. We suppose that $\mathrm{dom}(\rho) \cap FV(\mathrm{ran}(\rho)) = \{\}$. We further suppose w.l.g. that there is a set $\mathcal{Z} \subseteq \mathcal{K}$ of variables disjoint from $\mathrm{dom}(\rho)$ and $FV(\mathrm{ran}(\rho))$ and for all $y : \underbrace{P \to \cdots \to P}_{n \text{ times}} \to P \in \mathrm{dom}(\rho)$ that $\rho(y)$ is of the form

$$\lambda z_1 : P.\ \cdots \lambda z_n : P.\ H_y$$

for a term $H_y$, with each $z_i \in \mathcal{Z}$.

In the sequel $y$ ranges over $\mathrm{dom}(\rho)$, $z$ over $\mathcal{Z}^P$ and $x$ over $\mathcal{K}^P - \mathrm{dom}(\rho) - \mathcal{Z}$.

**Definition** The relation $\rhd_\rho \subseteq T_1^P \times (\mathcal{K}^P - \mathrm{dom}(\rho) - \mathcal{Z})$ is the least such that

1. $E \rhd x \Rightarrow E \rhd_\rho x$
2. $E \rhd y\tilde{E} \wedge H_y \rhd x \Rightarrow E \rhd_\rho x$
3. $E \rhd y\tilde{E} \wedge H_y \rhd z_j \wedge E_j \rhd_\rho x \Rightarrow E \rhd_\rho x.$

For $a \in Act$ let $\xrightarrow{a}_\rho \subseteq T_1^P \times T_1^P$ be the least relation such that

1. $E \xrightarrow{a} F \Rightarrow E \xrightarrow{a}_\rho F$
2. $E \rhd y\tilde{E} \wedge H_y \xrightarrow{a} H' \Rightarrow E \xrightarrow{a}_\rho H'[\tilde{E}/\tilde{z}]$
3. $E \rhd y\tilde{E} \wedge H_y \rhd z_j \wedge E_j \xrightarrow{a}_\rho B \Rightarrow E \xrightarrow{a}_\rho B.$

**Proposition 53** *These relations agree with the transition system, i.e.*

$$\begin{array}{rcl} E\rho \rhd x & \iff & E \rhd_\rho x \\ E\rho \xrightarrow{a} B & \iff & \exists F \,.\, E \xrightarrow{a}_\rho F \wedge F\rho = B. \end{array}$$

PROOF The right-to-left implications are straightforward inductions. To show the others we first pick out a controlled sub-relation of the $\beta$ reduction of $E\rho$.

> **Definition** Let the relation $\longrightarrow_\beta \subseteq T_1^P \times T_1^P$ be the least relation such that
>
> 1. $y\tilde{E} \longrightarrow_\beta H_y[\tilde{E}/\tilde{z}]$
> 2. $E \longrightarrow_\beta F \Rightarrow \mathit{fix}\ \lambda w : P.\ E \longrightarrow_\beta \mathit{fix}\ \lambda w : P.\ F$
> 3. $(E_j \longrightarrow_\beta E'_j \wedge \forall i \neq j \,.\, E_i = E'_i) \Rightarrow c\tilde{E} \longrightarrow_\beta c\tilde{E}'$ for any constant or variable $c$.
>
> This is related to $\beta\eta$ equality by the following.
>
> **Lemma 54** *For all $E$ there is some $F$ such that $E \longrightarrow_\beta{}^* F$ and $FV(F) \cap \mathrm{dom}(\rho) = \{\}$.*
>
> PROOF One can show that otherwise $E\rho$ has an infinite sequence of $\beta$ reductions. □

**Lemma 55** *If $E \longrightarrow_\beta F$ then $E\rho = F\rho$.*

PROOF Induction on $E \longrightarrow_\beta F$. □

It is related to the transition relation by

**Lemma 56** *If $E \longrightarrow_\beta F \xrightarrow{a}_\rho H$ then for some $G$ $E \xrightarrow{a}_\rho G \longrightarrow_\beta{}^* H$. Further, if $E \longrightarrow_\beta F \rhd_\rho x$ then $E \rhd_\rho x$.*

PROOF By somewhat tedious inductions on derivations of $\longrightarrow_\beta$. □

Now suppose $E\rho \xrightarrow{a} B$. By Lemmas 54 and 55 there is an $F$ such that $E \longrightarrow_\beta{}^* F$ and $E\rho = F \xrightarrow{a} B$. By the definition of $\xrightarrow{a}_\rho$ we have $F \xrightarrow{a}_\rho B$ so using Lemma 56 we have $E \xrightarrow{a}_\rho A \longrightarrow_\beta{}^* B$ for some $A$. Finally by Lemma 55 we have $A\rho = B$.

Suppose $E\rho \rhd x$. As before there is an $F$ such that $E \longrightarrow_\beta{}^* F$ and $E\rho = F \rhd x$. By the definition of $\rhd_\rho$ we have $F \rhd_\rho x$ so using Lemma 56 we have $E \rhd_\rho x$. □

## 3.6 The coincidence of $\sim_{ext}$ and $\sim_{ho}$

We now show that over $T_1^P$ the equivalences $\sim_{ext}$ and $\sim_{ho}$ coincide.

**Theorem 7** *If $Act$ is nonempty and $E \sim_{ext} F$ then $E \sim_{ho} F$.*

PROOF Suppose there is some action $a \in Act$. By Lemma 37 there is a largest $N$ such that some derivative of $E$ or $F$ is higher order bisimilar to $a^N 0$ (take $N = 0$ if there are none such). Using this we construct a substitution $\rho$ for $E, F$. For $y : \underbrace{P \to \cdots \to P}_{m \text{ times}} \to P$ put

$$\begin{array}{rcl} \rho(y) & \stackrel{\text{def}}{=} & \lambda z_1 : P. \ \cdots \lambda z_m : P. \ aA_y \\ A_y & \stackrel{\text{def}}{=} & a0 + a^{N+3+y}0 + \sum_{i \in 1..m} aB_i \\ B_i & \stackrel{\text{def}}{=} & a^{2i}0 + a^{2i+1}z_i \end{array}$$

(eliding some injective function from variables to the naturals). The following three lemmas show that $\rho$ is sufficiently discriminating.

**Lemma 57** *If $E' \in \mathrm{der}(E) \cup \mathrm{der}(F)$ then $\forall \tilde{G} \ . \ E'\rho \not\sim_{ho} A_y[\tilde{G}\rho/\tilde{z}]$.*

PROOF Suppose not, then $E'\rho \xrightarrow{a} \sim_{ho} P \stackrel{\text{def}}{=} a^{N+2+y}0$. By Lemma 52 either $E' \xrightarrow{a} E'' \wedge E''\rho \sim_{ho} P$ or $E' \rhd y'\tilde{E} \wedge A_{y'}[\tilde{E}\rho/\tilde{z}] \sim_{ho} P$. Both contradict the definition of $\rho$. □

**Lemma 58** *If $y \neq y'$ then $\forall \tilde{E}, \tilde{F} \,.\, A_y[\tilde{E}/\tilde{z}] \not\sim_{ho} A_{y'}[\tilde{F}/\tilde{z}]$.*

PROOF The transition to $a^{N+2+y}0$ of the left hand side cannot be matched by the right. □

**Lemma 59** *If $A_y[\tilde{E}/\tilde{z}] \sim_{ho} A_y[\tilde{F}/\tilde{z}]$ then $\forall i \,.\, E_i \sim_{ho} F_i$.*

PROOF Straightforward, by consideration of the ways in which a transition $A_y[\tilde{E}/\tilde{z}] \xrightarrow{a} B_i[\tilde{E}/\tilde{z}]$ can be matched. □

We can now check that

$$R \stackrel{\text{def}}{=} \{\, E', F' \mid E' \in \text{der}(E), F' \in \text{der}(F) \text{ and } E'\rho \sim_{ho} F'\rho \,\}$$

is a higher order bisimulation. Consider $E' \; R \; F'$ and suppose $E' \xrightarrow{b} E''$. By Lemma 35 $E'\rho \xrightarrow{b} E''\rho$ so $F'\rho \xrightarrow{b} \sim E''\rho$. By Lemma 52 either $F' \xrightarrow{b} F'' \wedge F''\rho \sim E''\rho$ — then $E'' \; R \; F''$ — or $F' \rhd y\tilde{F} \wedge b = a \wedge A_y[\tilde{F}\rho/\tilde{z}] \sim E''\rho$ — contradicting Lemma 57. Now suppose $E' \rhd y\tilde{E}$. We have $E'\rho \xrightarrow{a} A_y[\tilde{E}\rho/\tilde{z}]$ so $F'\rho \xrightarrow{a} \sim A_y[\tilde{E}\rho/\tilde{z}]$. By Lemma 52 either $F' \xrightarrow{a} F'' \wedge F''\rho \sim A_y[\tilde{E}\rho/\tilde{z}]$ — contradicting Lemma 57 — or $F' \rhd y'\tilde{F} \wedge A_{y'}[\tilde{F}\rho/\tilde{z}] \sim A_y[\tilde{E}\rho/\tilde{z}]$ — in which case by Lemma 58 $y = y'$ and by Lemma 59 $\forall i \,.\, E_i\rho \sim F_i\rho$ so $\forall i \,.\, E_i \; R \; F_i$. □

**Remark** The presence of $+$ is essential for this theorem. Consider a variable $y : P \to P \to P$. The $+$-free normal forms in $T_0^{P \to P \to P}$ are of the form $\lambda x_1 : P.\ \lambda x_2 : P.\ E$ for $E$ of

$$E ::= 0 \mid aE \mid x \mid \mathit{fix}\ \lambda x : P.\ E$$

and so must ignore one or both of their arguments. The appropriate definition of $\sim_{ext}$ is $\sim_{ext}^{+\mathit{free}}$ where $A \sim_{ext}^{+\mathit{free}} B$ iff for all substitutions for $A, B$ with $+$-free range $A\rho \sim B\rho$. Taking the terms $A \stackrel{\text{def}}{=} y\ 0\ (y\ a0\ 0)$ and $B \stackrel{\text{def}}{=} y\ 0\ (y\ aa0\ 0)$ we have $A \sim_{ext}^{+\mathit{free}} B$ but $A \not\sim_{ho} B$. This prevents a cheap proof of non-finite-axiomatisability over the $+$-free fragment. For the even simpler $+$-free fragment with a single action we conjecture that there *is* a finite equational axiomatisation using the axiom

$$\{f : P \to P,\ g : P \to P\} \rhd \mathit{fix}\ \lambda x : P.\ f(gx) = \mathit{fix}\ \lambda x : P.\ f(g(g(x)) : P,$$

which is sound only because this fragment is so inexpressive.

**Theorem 8** *If* $E \sim_{ho} F$ *then* $E \sim_{ext} F$.

PROOF Given $E \sim_{ho} F$ we must show for all substitutions $\rho$ for $E, F$ that $E\rho \sim_{ho} F\rho$. This is done indirectly. We construct below transition systems $E{\star}\rho, F{\star}\rho$ and show in the next two lemmas that

$$E\rho \sim_{ho} E{\star}\rho \sim_{ho} F{\star}\rho \sim_{ho} E\rho.$$

□

The (straightforward) direct proof is not given as in the next section information relating the 'loop structure' of $E\rho$ and $F\rho$ is extracted from these bisimulations.

The transition system $E{\star}\rho$ differs from $E\rho$ in that states that might be identified by the non-injectivity of $\rho$ are split apart. For example if

$$\begin{aligned} E &\stackrel{\text{def}}{=} y + aE' \\ E' &\stackrel{\text{def}}{=} \mathit{fix}\ \lambda x : P.\ a(z + ax) \\ \rho(y) \stackrel{\text{def}}{=} \rho(z) &\stackrel{\text{def}}{=} 0 \end{aligned}$$

then

$$E\rho \underset{a}{\overset{a}{\rightleftarrows}} E'\rho$$

whereas $E{\star}\rho$ is

$$\langle E\rangle \xrightarrow{a} \langle E'\rangle \underset{a}{\overset{a}{\rightleftarrows}} \langle z + aE'\rangle$$

To define $E{\star}\rho$ the inference system of §3.3 is extended to one for inferring transitions labelled by nonempty finite sequences of actions, with the rules

$$\frac{E \xrightarrow{a} F}{E \xrightarrow{a\epsilon} F}\ \mathit{single} \qquad\qquad \frac{E \xrightarrow{a} F \quad F \xrightarrow{l} G}{E \xrightarrow{al} G}\ \mathit{cons}.$$

If $d$ is an inference tree of this system with conclusion $E \xrightarrow{l} F$ we write $d : E \xrightarrow{l} F$.

**Definition** Given a substitution $\rho$ for $E_0$ with $\rho$ and notation as in §3.5 the transition system $E_0{\star}\rho$ has states

$$\begin{aligned} S \stackrel{\text{def}}{=}\ & \{\, \langle E\rangle \mid E \in \mathrm{der}(E_0)\,\} \\ \uplus\ & \{\langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d\rangle \mid\ d : H_y \xrightarrow{l} H' \\ & \qquad\qquad \text{and } \exists E \in \mathrm{der}(E_0)\ .\ E \rhd y\tilde{E}\} \end{aligned}$$

with root $\langle E_0 \rangle$. The transition relation $\overset{a}{\leadsto}$ of $E_0 \star \rho$ is the least relation such that

1. $E \xrightarrow{a} F \Rightarrow \langle E \rangle \overset{a}{\leadsto} \langle F \rangle$

2. $E \rhd y\tilde{E} \wedge d : H_y \xrightarrow{a} H' \Rightarrow \langle E \rangle \overset{a}{\leadsto} \langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, \mathit{single}(d) \rangle$

3. $E \rhd y\tilde{E} \wedge H_y \rhd z_i \wedge \langle E_i \rangle \overset{a}{\leadsto} s \Rightarrow \langle E \rangle \overset{a}{\leadsto} s$

4. $(\exists E \in \mathrm{der}(E_0)\ .\ E \rhd y\tilde{E}) \wedge d : H_y \xrightarrow{l} H' \wedge d' : H' \xrightarrow{a} H'' \Rightarrow$
   $\langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d \rangle \overset{a}{\leadsto} \langle H''[\tilde{E}/\tilde{z}], y\tilde{E}, \mathit{cons}(d, d') \rangle$

5. $(\exists E \in \mathrm{der}(E_0)\ .\ E \rhd y\tilde{E}) \wedge d : H_y \xrightarrow{l} H' \wedge H' \rhd z_i \wedge \langle E \rangle \overset{a}{\leadsto} s \Rightarrow$
   $\langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d \rangle \overset{a}{\leadsto} s$.

The visibilities $\rhd'$ of $E_0 \star \rho$ are given by

1. (There is no corresponding case.)

2. $E \rhd y\tilde{E} \wedge H_y \rhd x \Rightarrow \langle E \rangle \rhd' x$

3. $E \rhd y\tilde{E} \wedge H_y \rhd z_i \wedge \langle E_i \rangle \rhd' x \Rightarrow \langle E \rangle \rhd' x$

4. $(\exists E \in \mathrm{der}(E_0)\ .\ E \rhd y\tilde{E}) \wedge d : H_y \xrightarrow{l} H' \wedge H' \rhd x \Rightarrow$
   $\langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d \rangle \rhd' x$

5. $(\exists E \in \mathrm{der}(E_0)\ .\ E \rhd y\tilde{E}) \wedge d : H_y \xrightarrow{l} H' \wedge H' \rhd z_i \wedge \langle E \rangle \rhd' x \Rightarrow$
   $\langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d \rangle \rhd' x$.

**Lemma 60** *If $\rho$ is a substitution for $E$ then $E\rho \sim_{ho} E \star \rho$.*

PROOF One can check that

$$\begin{array}{rl} & \{\, A\rho, \langle A \rangle \mid \langle A \rangle \text{ is a state of } E \star \rho \,\} \\ \cup & \{\, A\rho, \langle A, F, d \rangle \mid \langle A, F, d \rangle \text{ is a state of } E \star \rho \,\} \end{array}$$

is a higher order bisimulation, using induction on the transition derivations. □

**Lemma 61** *If $E \sim_{ho} F$ and $\rho$ is a substitution for $E, F$ then $E \star \rho \sim_{ho} F \star \rho$.*

PROOF By Lemma 37 there is a finite higher order bisimulation $R$ with $E\ R\ F$. Let the relation $Q$ between the states of $E{\star}\rho$ and $F{\star}\rho$ be

$$\begin{array}{rll} Q & \stackrel{\text{def}}{=} & \{\, \langle E' \rangle, \langle F' \rangle \mid E'\ R\ F' \,\} \\ & \cup & \{\, \langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d \rangle, \langle H'[\tilde{F}/\tilde{z}], y\tilde{F}, d \rangle \mid \begin{array}[t]{l} \text{there exist } E' \in \mathrm{der}(E) \text{ and} \\ F' \in \mathrm{der}(F) \text{ such that} \\ E' \rhd y\tilde{E}, F' \rhd y\tilde{F}, \\ \forall j\ .\ E_j\ R\ F_j \text{ and } d : H_y \stackrel{l}{\longrightarrow} H' \,\}. \end{array} \end{array}$$

One can check that $Q$ is a higher order bisimulation betwixt $E{\star}\rho$ and $F{\star}\rho$, using induction on the transition derivations. □

## 3.7 Loop properties

The instantiations $E\rho, F\rho$ of a higher order bisimilar $E, F$ are uniform in a sense captured by the following definition and theorem.

**Definition** For $u \geq 1$ take the predicate $L_u$ and equivalence relation $\equiv_u$ over states in a transition system to be:

- $L_u s$ iff $s$ has a loop with no prime factor $\geq u$, i.e. $s \longrightarrow^* s' \longrightarrow^n s'$ for some $s'$ and $n$ with no prime factor $\geq u$.

- $s \equiv_u s'$ iff $\forall v \geq u\ .\ L_v s \Leftrightarrow L_v s'$.

**Theorem 9** *If $E \sim_{ho} F$ then there is some $u \geq 1$ such that for all substitutions $\rho$ for $E, F$ $E\rho \equiv_u F\rho$.*

PROOF This follows from the following two lemmas. □

**Lemma 62** *If $\rho$ is a substitution for $E$ then $\forall u \geq 1\ .\ E\rho \equiv_u E{\star}\rho$.*

PROOF See Appendix B. □

**Definition** For a finite relation $U \subseteq A \times B$ say the *width* of $U$ is

$$\mathrm{width}(U) \stackrel{\text{def}}{=} \max\{\max_{a\in A} \#\{\, b \mid a\ U\ b \,\},\ \max_{b\in B} \#\{\, a \mid a\ U\ b \,\}\}.$$

**Lemma 63** *If $E \sim_{ho} F$ then there is some $u \geq 1$ such that for all substitutions $\rho$ for $E, F$ $E\star\rho \equiv_u F\star\rho$.*

PROOF Consider the $R$ and $Q$ in the proof of Lemma 61. We first give a $u \geq 1$, dependent on $R$ but not on $\rho$, strictly greater than the width of $Q$. Let $R' \subseteq T_1^P \times T_1^P$ be

$$y\tilde{E}\ R'\ y\tilde{F} \quad \Longleftrightarrow \quad \exists E' \in \mathrm{der}(E),\ F' \in \mathrm{der}(F)\ .\ E' \rhd y\tilde{E} \wedge F' \rhd y\tilde{F} \wedge \forall i\ .\ E_i\ R\ F_i$$

and put $u \stackrel{\text{def}}{=} 1 + \max\{\mathrm{width}(R),\ \mathrm{width}(R')\}$. To see that $E\star\rho \equiv_u F\star\rho$ suppose that $s\ Q\ t, v \geq u$ and $L_v s$, i.e. for some $s'$ and some $n \geq 1$ with no prime factors $\geq v$ that $s \longrightarrow^* s' \longrightarrow^n s'$. As $Q$ is a higher order bisimulation there exist $t_i \mid i \geq 0$ such that

$$\begin{array}{ccccccccc} & & s' & & s' & & s' & \cdots \\ & & Q & & Q & & Q & \cdots \\ t & \longrightarrow^* & t_0 & \longrightarrow^n & t_1 & \longrightarrow^n & t_2 & \cdots \end{array}$$

but $\#\{\, t_i \mid i \geq 0 \,\} < u$ so for some $k \in 1..u-1$ we have $t \longrightarrow^* t' \longrightarrow^{kn} t'$. Further, $kn$ has no prime factors $\geq v$. □

The equivalences $\equiv_u$ have the following congruence property.

**Lemma 64** *For $M, N \in T_0^P$ and $C[\_]$ a context from*

$$C ::= \_ \mid x \mid 0 \mid aC \mid C + C \mid \mathit{fix}\ \lambda x : P.\ C$$

*(where $x : P \in \mathcal{K}$), if $M \sim_{ho} N$ and $M \equiv_u N$ then $C[M] \equiv_u C[N]$.*

PROOF By induction on contexts using the following lemma. □

**Definition** $\mathrm{loops}\, E \stackrel{\text{def}}{=} \{\, k \mid \exists E'\ .\ E \longrightarrow^* E' \longrightarrow^k E' \,\}$

**Lemma 65**

$$\begin{array}{rcl} \mathrm{loops}\, aE & = & \mathrm{loops}\, E \\ \mathrm{loops}\, E + F & = & \mathrm{loops}\, E \cup \mathrm{loops}\, F \\ \mathrm{loops}\, \mathit{fix}\ \lambda x : P.\ E & = & \mathrm{loops}\, E \cup \left\{ \sum_{i \in I} k_i \;\middle|\; \forall i \in I\ .\ E \longrightarrow^{k_i} \rhd x \right\} \end{array}$$

*where $I$ ranges over finite sets.*

PROOF Straightforward. □

The following is included for later use.

**Lemma 66** *If $E \longrightarrow^* F \longrightarrow^l F$ then there exist a term $\mathit{fix}\ \lambda x : P.\ G$, a finite set $I$ and $k_i \geq 1 \mid i \in I$ such that $E \longrightarrow^* \rhd \mathit{fix}\ \lambda x : P.\ G$, $\sum_I k_i = l$ and $\forall i\ .\ G \longrightarrow^{k_i} \rhd x$.*

PROOF Induction on $E$ using Lemma 65. $\square$

Finally we can put together the results of this chapter.

**Lemma 67** *If $\mathcal{E}$ is a finite set of typed equalities that is sound for bisimulation and of order $\leq 1$ then there is some $u \geq 1$ such that for all $M, N \in T_0^P$ if $\mathcal{E} \vdash M = N : P$ then $M \equiv_u N$.*

PROOF We assume w.l.g. that $\mathcal{E}$ contains only equalities at type $P$. By Corollary 31 each equation lies within $\sim_{ext}$ and so by Theorem 7 within $\sim_{ho}$. Take $u$ to be the largest of those given by Theorem 9 applied to each equation. By Proposition 30 $M =_{\mathcal{E}} N$. An induction on a derivation of this then suffices, using Lemma 64 in the inductive steps and Theorem 9 at the uses of axioms. $\square$

PROOF (of Theorem 6) The main theorem follows by noting that if $\mathcal{E}$ is sound for an equivalence finer than bisimulation then it is sound for bisimulation and that, if $q$ is the smallest prime strictly greater than the $u$ given by the previous lemma, then

$$\mathit{fix}\ \lambda x : P.\ ax \not\equiv_u \mathit{fix}\ \lambda x : P.\ a^q x.$$

$\square$

# Chapter 4

# A Miscellany

In this chapter we include a number of applications and developments of the theory of higher order bisimulation. We consider the axiomatisability of bisimulation over $*$-expressions, relate higher order bisimulation to the higher order $\pi$ calculus and cast our results into the language of non-well-founded set theory. We also give a definition of weak higher order congruence and show that weak congruence is finitely axiomatisable relative to bisimulation. The sections may be read in any order.

## 4.1 Axiomatisation over $*$-expressions

Bisimulation equivalence of $*$-expressions has recently received attention in [BBP94], where the addition of operators that suffice to express the regular processes is considered. The results of the previous chapter can be used to give an easy proof of a non-axiomatisability result for the basic calculus.

### 4.1.1 Axiomatisation with and without $0$

To discuss axioms over $*$-expressions we find it convenient to introduce a set of variables disjoint from $Act$. An axiomatisation will be a set of pairs of open terms together with rules for equivalence, congruence and instantiation of variables. This differs slightly from the usual setup in which instantiation of actions is permitted.

Fokkink and Zantema have given (in [FZ94]) a complete axiomatisation for the $*$-expressions that do not contain $0$, $1$ or the unary $*$.

**Theorem 10 (Fokkink and Zantema)** *The axioms below are sound and complete for bisimulation over BPA*, i.e. over expressions of* $\{a, +, \cdot, \_^{\star}\_\}$.

$$
\begin{array}{rcl}
x + y &=& y + x \\
(x + y) + z &=& x + (y + z) \\
x + x &=& x \\
(x + y) \cdot z &=& x \cdot z + y \cdot z \\
(x \cdot y) \cdot z &=& x \cdot (y \cdot z) \\
x \cdot (x^{\star} y) + y &=& x^{\star} y \\
x^{\star}(y \cdot z) &=& (x^{\star} y) \cdot z \\
x^{\star}(y \cdot ((x + y)^{\star} z) + z) &=& (x + y)^{\star} z
\end{array}
$$

PROOF This is immediate from the result of [FZ94] and Proposition 17. □

If the zero process is added, however, there can be no finite axiomatisation.

**Theorem 11** *There is no finite axiomatisation for bisimulation over any set of $*$-expressions closed under* $\{0, a, +, \cdot\}$ *and one of* $\{\_^{*}, \_^{\star}\_\}$.

This can be shown using the results of Chapter 3. We first note that the $*$-expressions (and axioms over them) can be faithfully encoded into our lambda calculus, encoding sequential composition using function composition at type $P \to P$.

**Definition** Take the map $[\![\_]\!]$ from $*$-expressions to lambda calculus terms of type $P \to P$ to be

$$
\begin{array}{rcl}
[\![x]\!] &\stackrel{\text{def}}{=}& x \ \text{ where we suppose } x : P \to P \in \mathcal{K} \\
[\![a]\!] &\stackrel{\text{def}}{=}& a \\
[\![0]\!] &\stackrel{\text{def}}{=}& \lambda y : P.\ 0 \\
[\![1]\!] &\stackrel{\text{def}}{=}& \lambda y : P.\ y \\
[\![E + F]\!] &\stackrel{\text{def}}{=}& \lambda y : P.\ +([\![E]\!] y)([\![F]\!] y) \\
[\![E \cdot F]\!] &\stackrel{\text{def}}{=}& \lambda y : P.\ [\![E]\!]([\![F]\!] y) \\
[\![E^{*}]\!] &\stackrel{\text{def}}{=}& \lambda y : P.\ \mathit{fix}\ \lambda z : P.\ +(y)([\![E]\!] z)
\end{array}
$$

$$[\![E^\star F]\!] \stackrel{\text{def}}{=} \lambda y : P.\ \mathit{fix}\ \lambda z : P.\ +([\![F]\!]y)([\![E]\!]z)$$

**Lemma 68** *If $y : P \in \mathcal{K}$ then*

$$E \sim F \iff [\![E]\!]y \sim_{ho} [\![F]\!]y.$$

PROOF It is straightforward to check that $E\surd \iff [\![E]\!]y \rhd y$, $E \xrightarrow{a} E' \Rightarrow [\![E]\!]y \xrightarrow{a} [\![E']\!]y$ and $[\![E]\!]y \xrightarrow{a} A \Rightarrow \exists E'\ .\ E \xrightarrow{a} E' \wedge [\![E']\!]y = A$. □

PROOF (of Theorem 11) To prove the theorem we consider a sound finite set $\mathcal{A}$ of axioms. The encodings of these must be higher order bisimilar (as the terms constructed in the proof of Theorem 7 are all expressible) so $[\![\mathcal{A}]\!]$ is a sound set in our earlier sense. By Lemma 67 there is some $u$ such that if $M, N \in T_0^P$ and $[\![\mathcal{A}]\!] \vdash M = N : P$ then $M \equiv_u N$. Now consider the relevant pair of terms below, where $q$ is the smallest prime strictly greater than $u$.

$$\begin{array}{lllllll} E & \stackrel{\text{def}}{=} & a^* \cdot 0 & \quad & E & \stackrel{\text{def}}{=} & a^\star 0 \\ F & \stackrel{\text{def}}{=} & (\overbrace{a \cdot \ldots \cdot a}^{q \text{ times}})^* \cdot 0 & & F & \stackrel{\text{def}}{=} & (\overbrace{a \cdot \ldots \cdot a}^{q \text{ times}})^\star 0 \end{array}$$

In each case $[\![E]\!]y \not\equiv_u [\![F]\!]y$ so $[\![\mathcal{A}]\!] \not\vdash [\![E]\!]y = [\![F]\!]y : P$ so $\mathcal{A} \not\vdash E = F$ but $E \sim F$. □

The addition of other operators does not affect the result, so long as their semantics are expressible by an encoding into lambda terms as above.

In [Fok94] Fokkink introduced a subcalculus of the $*$-expressions that contains a zero process but for which bisimulation is nonetheless finitely equationally axiomatisable. Denoted MPA$^*_\delta$ (for Minimal Process Algebra) it is obtained by requiring the first operand of sequential compositions and of binary iterations to be single actions, i.e. in our notation the terms of

$$E ::= 0 \mid a \cdot E \mid E + E \mid a^\star E.$$

Aceto and Ingólfsdóttir have recently given (in [AI95]) an extension of that axiomatisation that is complete for observation congruence. These positive axiomatisability results, for BPA$^*$ and MPA$^*_\delta$, can be seen as depending on the inexpressiveness of the calculi, in that the processes required in the proof of Theorem 11 cannot be expressed.

### 4.1.2 Axiomatisation with $1$

In the light of the previous subsection it is natural to question whether bisimulation over expressions of the signature $\{a, 1, +, \cdot, \_^*\}$ is axiomatisable. Unfortunately a naive application of the same idea does not show non-axiomatisability. For a proof analogous to that of Theorem 11 we would need a term $G$ such that $G^*$ and $(G^q)^*$ are bisimilar but not provably equal. By the following lemma any such $G$ must be successfully terminated.

**Lemma 69** *For $q \geq 2$, if $G^* \sim (G^q)^*$ then either $G\surd$ or $\nexists H\,.\,G {\longrightarrow}^* H\surd$.*

PROOF Suppose for a contradiction that $\neg G\surd$ and $G{\longrightarrow}^+\surd$. Writing $n(E)$ for the smallest non-zero natural such that $E{\longrightarrow}^n\surd$ we have $n(G^*) = n(G) \geq 1$ and $n((G^q)^*) = q\;n(G^*) > n(G^*)$, hence $G^* \not\sim (G^q)^*$. $\square$

Taking the simplest nontrivial example $G \overset{\text{def}}{=} a + 1$ it is easy to see that for any $n$ there is a bisimulation of width 2 relating $((a+1)^n)^*$ and $((a+1)^{n+1})^*$, namely

$$\begin{aligned} R &\overset{\text{def}}{=} \{((a+1)^n)^*,\ ((a+1)^{n+1})^*\} \\ &\cup\ \{\, C^n_m,\ C^{n+1}_{m+1} \mid 0 \leq m \leq n-1\,\} \\ &\cup\ \{C^n_{n-1},\ C^{n+1}_0\} \end{aligned}$$

where

$$\begin{aligned} C^n_0 &\overset{\text{def}}{=} 1\cdot((a+1)^n)^* \\ C^n_m &\overset{\text{def}}{=} (1\cdot(a+1)^m)\cdot((a+1)^n)^*\ , \text{for } m \geq 1. \end{aligned}$$

The technique of Chapter 3 therefore cannot be used to show that $(a+1)^*$ and $((a+1)^q)^*$ are not provably equal. The question of axiomatisability over $\{a, 1, +, \cdot, \_^*\}$ remains open.

## 4.2 Higher order process calculi

The lambda calculus of §3.2 was introduced to allow a clean definition of interesting classes of equational axioms over the $\mu$-expressions, not involving ad hoc metanotation. It is, however, of additional interest when viewed as a simple fragment of a higher order process calculus such as the higher order $\pi$ calculus of Sangiorgi [San93,San94]. In particular the theory of strong higher order bisimulation is not well understood even for this fragment (the work of Sangiorgi is concerned only with the weak case). We will give a precise connection between higher order bisimulation (as defined in Chapter 3) and *normal bisimulation* as defined in [San94] for the strictly higher order fragment of the higher order $\pi$ calculus. This confirms the intuition that, in both, the equivalence of abstractions is checked by applying them to simple 'new' terms (new variables and triggers respectively).

### 4.2.1 Syntax and transitions

The basic version of the higher order $\pi$ calculus of [San94] has guarded summation, parallel composition, restriction and the application of a variable or constant:

$$P ::= \sum_{i \in I} \alpha_i.P_i \mid P_1 \,|\, P_2 \mid \nu a P \mid Y\langle\tilde{A}\rangle \mid D\langle\tilde{A}\rangle$$

where prefixes $\alpha$ are inputs or outputs:

$$\alpha ::= a(\tilde{X}) \mid \overline{a}\langle\tilde{A}\rangle$$

(with $a \neq \tau$) and agents $A$ are abstractions:

$$A ::= (\tilde{X})P.$$

Here $\tilde{X}$ and $\tilde{A}$ range over (possibly empty) tuples. It is assumed that all terms obey a suitable sorting discipline. There is an evident bijection between the lambda terms in $T^P_\omega$ (in normal form, containing only guarded summation and modulo axioms $A1$–$4$) and the higher order $\pi$ terms in which there is no parallel composition, restriction or infinitary summation, the only defined constant is

$$\mathrm{FIX} \stackrel{\text{def}}{=} (X)X\langle\mathrm{FIX}\langle X\rangle\rangle$$

$$\frac{}{\overline{a}\langle\tilde{A}\rangle.P \xrightarrow{\overline{a}\langle\tilde{A}\rangle} P}\ \mathit{out} \qquad \frac{}{a(\tilde{X}).P \xrightarrow{a\langle\tilde{A}\rangle} P[\tilde{A}/\tilde{X}]}\ \mathit{inp}$$

$$\frac{P \xrightarrow{\mu} P'}{P \mid Q \xrightarrow{\mu} P' \mid Q}\ \mathit{par} \qquad \frac{P \xrightarrow{\overline{a}\langle\tilde{A}\rangle} P' \quad Q \xrightarrow{a\langle\tilde{A}\rangle} Q'}{P \mid Q \xrightarrow{\tau} P' \mid Q'}\ \mathit{com}$$

$$\frac{P \xrightarrow{\mu} P'}{P + Q \xrightarrow{\mu} P'}\ \mathit{sum} \qquad \frac{P \mid !P \xrightarrow{\mu} P'}{!P \xrightarrow{\mu} P'}\ \mathit{rep}$$

**Figure 4–1:** Higher order $\pi$ transition system

and all names appear only as inputs with a nullary sort — i.e. the terms of

$$P ::= \sum_{i \in I} a_i\langle\rangle.P_i \mid Y\langle\tilde{A}\rangle \mid \mathrm{FIX}\langle A\rangle.$$

We define higher order bisimulation ($\sim_{ho}$) over these terms via the bijection. Unfortunately the definition of normal bisimulation and the results of [San94] are in terms of a slightly different calculus — with replication in place of constants and with unary variables and names, i.e. variables only of types from

$$\sigma ::= P \mid \sigma \to P$$

and names that carry single values of these types. We will therefore only obtain a result for lambda terms that do not contain *fix* , however the definition will be generalised to allow arbitrary types. We take a fragment of the higher order $\pi$ calculus that is just large enough to contain the terms of

$$P ::= \sum_{i \in I} a_i\langle\rangle.P_i \mid Y\langle\tilde{A}\rangle \qquad \dagger$$

and to admit a definition of normal bisimulation. The terms are those of

$$P ::= \sum_{i \in I} \alpha_i.P_i \mid P_1 \mid P_2 \mid Y\langle\tilde{A}\rangle \mid !P. \qquad \ddagger$$

The labelled transition system semantics for this is given in Figure 4–1, in which $\mu$ ranges over all action labels and the symmetric counterparts of $\mathit{par}$, $\mathit{com}$ and $\mathit{sum}$ are omitted. For the sub-fragment with only unary variables and names the transition system coincides with that of [San94]. Our omission of restriction is therefore harmless in that normal bisimulation (as defined in the next section) over this sub-fragment coincides with normal bisimulation over the whole calculus of [San94]. A similar coincidence presumably holds between normal bisimulation over the latter and normal bisimulation over the 'reduced higher order $\pi$ calculus' of [San93], which allows name passing in addition to process passing.

### 4.2.2 Normal bisimulation

Three equivalences over the higher order $\pi$ calculus are defined and shown equal in [San93] — context bisimulation, barbed congruence and normal bisimulation. The idea behind the latter is that to check the input case of the equivalence of two processes it suffices to consider inputs of a particular form, the *triggers*. A trigger is an agent of the form

$$\mathrm{Tr}_{m_i} \stackrel{\mathrm{def}}{=} (\tilde{X})\overline{m}_i\langle\tilde{X}\rangle.0$$

for a name $m_i$ of suitable sort. Similarly to check the equivalence of processes with free variables it suffices to substitute triggers on new, distinct names for the variables and then check equivalence.

**Definition** A symmetric relation $R$ over closed terms of ‡ is a *normal bisimulation* if for any $P\ R\ Q$ and distinct $\tilde{m}$ not in the free names of $P, Q$:

- If $P \stackrel{a\langle\mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle}{\longrightarrow} P'$ then $\exists Q'\ .\ Q \stackrel{a\langle\mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle}{\Longrightarrow} Q'$ and $P'\ R\ Q'$.
- If $P \stackrel{\overline{a}\langle A_1,\ldots,A_n\rangle}{\longrightarrow} P'$ then $\exists Q', \tilde{B}\ .\ Q \stackrel{\overline{a}\langle B_1,\ldots,B_n\rangle}{\Longrightarrow} Q'$ and

$$\begin{array}{c} P'\,|!m_1(\tilde{X}_1).A_1\langle\tilde{X}_1\rangle\,|\cdots|!m_n(\tilde{X}_n).A_n\langle\tilde{X}_n\rangle \\ R \\ Q'\,|!m_1(\tilde{X}_1).B_1\langle\tilde{X}_1\rangle\,|\cdots|!m_n(\tilde{X}_n).B_n\langle\tilde{X}_n\rangle. \end{array}$$

- If $P \stackrel{\tau}{\longrightarrow} P'$ then $\exists Q'\ .\ Q \Longrightarrow Q'$ and $P'\ R\ Q'$.

(Here $\stackrel{\alpha}{\Longrightarrow} \stackrel{\mathrm{def}}{=} \stackrel{\tau}{\longrightarrow}^* \stackrel{\alpha}{\longrightarrow} \stackrel{\tau}{\longrightarrow}^*$ and $\Longrightarrow \stackrel{\mathrm{def}}{=} \stackrel{\tau}{\longrightarrow}^*$, as usual.) The union of all normal bisimulations is written $\sim_{Nr}$. It is lifted to open terms $P, Q$ with free variables $X, Y, \ldots$ by

$$P \sim_{Nr} Q \quad \textit{iff} \quad P\rho \sim_{Nr} Q\rho$$

where $\rho$ is a substitution $[\mathrm{Tr}_{m_X}/X,\ \ \mathrm{Tr}_{m_Y}/Y, \ldots]$ for some new distinct names $m_X, m_Y, \ldots$.

This is the natural generalisation of the definition of [San94] to processes with free variables of arbitrary type, but without restriction.

### 4.2.3 The coincidence of $\sim_{ho}$ and $\sim_{Nr}$

**Proposition 70** *If $P, Q$ are terms of †, i.e. of*

$$P ::= \sum_{i \in I} a_i\langle\rangle.P_i \mid Y\langle\tilde{A}\rangle$$

*then $P \sim_{ho} Q$ iff $P \sim_{Nr} Q$.*

PROOF In the following we assume that the set of names is the disjoint union of $Act$ (ranged over by $a$) and a set $M$ (ranged over by $m$). $E$ and $F$ range over abstractions $(\tilde{X})P$ of the above processes. We work up to the associativity, commutativity and $0$-absorbtion of $|$ and the law $!P = P\,|!P$. A *trigger substitution* $\rho$ is a substitution of distinct triggers for all free variables in terms to which it is applied. For the right-to-left direction we let

$$R \stackrel{\text{def}}{=} \{\, E, F \mid E\rho \sim_{Nr} F\rho \text{ and } \rho \text{ is a trigger substitution} \,\}$$

and check that this is a higher order bisimulation.

Suppose $E \;\; R \;\; F$ and $E \rhd y\tilde{E}$. Clearly $E\rho \xrightarrow{\overline{m}_y\langle\tilde{E}\rho\rangle} 0$ so there is some $\tilde{B}$ such that $F\rho \stackrel{\overline{m}_y\langle\tilde{B}\rangle}{\Longrightarrow} 0$ and for any new $\tilde{m}$ and $\tilde{X}_i \mid i \in 1..n$

$$\begin{array}{c} !m_1(\tilde{X}_1).E_1\rho\langle\tilde{X}_1\rangle \mid \cdots \mid !m_n(\tilde{X}_n).E_n\rho\langle\tilde{X}_n\rangle \\ \sim_{Nr} \\ !m_1(\tilde{X}_1).B_1\langle\tilde{X}_1\rangle \mid \cdots \mid !m_n(\tilde{X}_n).B_n\langle\tilde{X}_n\rangle. \end{array}$$

Hence there is some $\tilde{F}$ such that $F \rhd y\tilde{F}$ and $\tilde{F}\rho = \tilde{B}$. Further, by the definition of normal bisimulation, for any $i$ and new $\tilde{m}_i$

$$(E_i\langle\tilde{X}_i\rangle)\rho \oplus [\mathrm{Tr}_{m_{i1}}/X_{i1}, \ldots] \quad \sim_{Nr} \quad (F_i\langle\tilde{X}_i\rangle)\rho \oplus [\mathrm{Tr}_{m_{i1}}/X_{i1}, \ldots]$$

so $(E_i\langle\tilde{X}_i\rangle) \; R \; (F_i\langle\tilde{X}_i\rangle)$ and $E_i \; \hat{R}^{\sigma_i} \; F_i$.

The reasoning for $E \xrightarrow{a} E'$ is straightforward.

For the left-to-right direction we take

$$\begin{array}{rll} R_0 & \stackrel{\text{def}}{=} & \{\, E\rho, F\rho \mid E \sim^{P}_{ho} F \text{ and } \rho \text{ is a trigger substitution} \,\} \\ & \cup & \{\, !m(\tilde{X}).E\rho\langle\tilde{X}\rangle, !m(\tilde{X}).F\rho\langle\tilde{X}\rangle \mid \;\; E \sim^{\sigma_1\cdots\sigma_n \to P}_{ho} F \\ & & \qquad \rho \text{ is a trigger substitution} \\ & & \qquad \tilde{X} \text{ are not free in } E, F \\ & & \qquad m \text{ does not occur in } \mathrm{ran}(\rho) \,\} \end{array}$$

and close up under certain parallel compositions. Let $R$ be the smallest relation containing $R_0$ such that if $P \ R \ Q,\ P' \ R \ Q'$ and no name occurs both in input and output positions in $P \mid P'$ or $Q \mid Q'$ then $P \mid P' \ R \ Q \mid Q'$. We check that $R$ is a normal bisimulation. No communication can occur in any process in $\mathrm{dom}(R) \cup \mathrm{ran}(R)$ so the only transitions that need be considered are the following, in which $P \ R \ Q,\ E \sim^{\sigma}_{ho} F$ and $\tilde{m},\ \tilde{X}$ are new:

- $P \mid E\rho \xrightarrow{a\langle\rangle} P \mid E'\rho$ as $E \xrightarrow{a} E'$
  $R$
  $Q \mid F\rho$

- $P \mid !m(\tilde{X}).E\rho\langle\tilde{X}\rangle \xrightarrow{m\langle \mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle} P \mid !m(\tilde{X}).E\rho\langle\tilde{X}\rangle \mid E\rho\langle \mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle$
  $R$
  $Q \mid !m(\tilde{X}).F\rho\langle\tilde{X}\rangle$

- $P \mid E\rho \xrightarrow{\overline{m}_y\langle\tilde{E}\rho\rangle} P \mid 0$ as $E \triangleright y\tilde{E}$
  $R$
  $Q \mid F\rho$

For the first, $F \xrightarrow{a} F' \sim_{ho} E'$ so $(Q \mid F\rho) \xrightarrow{a\langle\rangle} (Q \mid F'\rho) \ R \ (P \mid E'\rho)$.

For the second

$$Q \mid !m(\tilde{X}).F\rho\langle\tilde{X}\rangle \quad \xrightarrow{m\langle \mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle} \quad Q \mid !m(\tilde{X}).F\rho\langle\tilde{X}\rangle \mid F\rho\langle \mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle.$$

We have $E\langle\tilde{X}\rangle \sim^{P}_{ho} F\langle\tilde{X}\rangle$ so $E\langle\tilde{X}\rangle\rho \oplus [\mathrm{Tr}_{m_1}/X_1,\ldots] \ R_0 \ F\langle\tilde{X}\rangle\rho \oplus [\mathrm{Tr}_{m_1}/X_1,\ldots]$, hence $E\rho\langle \mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle \ R_0 \ F\rho\langle \mathrm{Tr}_{m_1},\ldots,\mathrm{Tr}_{m_n}\rangle$ and the targets of the transitions are related by $R$.

For the third there exists some $\tilde{F}$ such that $F \triangleright y\tilde{F}$ and $\forall i \,.\, E_i \sim^{\sigma_i}_{ho} F_i$. It follows that

$$Q \mid F\rho \quad \xrightarrow{\overline{m}_y\langle\tilde{F}\rho\rangle} \quad Q \mid 0.$$

Further, for all $i$ and new $\tilde{X}_i$

$$!m_i(\tilde{X}_i).E_i\rho\langle\tilde{X}_i\rangle \;\; R_0 \; !m_i(\tilde{X}_i).F_i\rho\langle\tilde{X}_i\rangle$$

so

$$(P \,|!m_1(\tilde{X}_1).E_1\rho\langle\tilde{X}_1\rangle \,|\cdots) \;\; R \;\; (Q \,|!m_1(\tilde{X}_1).F_1\rho\langle\tilde{X}_1\rangle \,|\cdots).$$

□

We expect that this proposition could be generalised to all lambda terms, i.e. to

$$P ::= \sum_{i\in I} a_i\langle\rangle.P_i \mid Y\langle\tilde{A}\rangle \mid \mathrm{FIX}\langle A\rangle,$$

without difficulty, although the statement would require a higher order $\pi$ calculus with both replication and constants.

The definitions of barbed congruence and context bisimulation both involve quantification over higher order $\pi$ contexts and so cannot be easily related to $\sim_{ho}$ or $\sim_{ext}$ (except via the above proposition, of course). This prevents us making use of Sangiorgi's results to obtain a cheap proof of the equality of $\sim_{ho}$ and $\sim_{ext}$.

## 4.3 Non-well-founded set theory

In the usual conception of set theory the membership relation is required to be well-founded — there is, for example, no set which is a member of itself. The Zermelo-Fraenkel axiomatisation expresses this by the axiom of Foundation:

FA: Any non-empty set $X$ has an element $Y$ such that $X \cap Y = \{\}$.

In [Acz88] Aczel discusses a number of alternative axioms that admit non-well-founded sets. He focuses on the axioms of ZFC with an axiom AFA (of Anti-Foundation) replacing FA. We recall its statement.

**Definition** A *graph* is a set of nodes together with a set of edges, i.e. ordered pairs of the nodes. A *decoration* of a graph is an assignment of a set to each of its nodes such that the elements of the set assigned to a node are the sets assigned to the children of that node.

AFA: Any graph has a unique decoration.

Pointed graphs can therefore be taken to denote unique sets of ZFC$-$FA+AFA. To say when they denote the same set we define bisimulation between pointed graphs in the obvious way.

**Definition** A *bisimulation* between pointed graphs $\langle N, \longrightarrow, n_0\rangle$ and $\langle N', \longrightarrow', n_0'\rangle$ is a relation $R \subseteq N \times N'$ such that $n_0 \; R \; n_0'$ and if $n \; R \; n'$ then:

- If $n \longrightarrow n_1$ then $\exists n_1' \,.\, n' \longrightarrow' n_1' \wedge n_1 \; R \; n_1'$.
- If $n' \longrightarrow' n_1'$ then $\exists n_1 \,.\, n \longrightarrow n_1 \wedge n_1 \; R \; n_1'$.

**Proposition 71 (Aczel)** *Two pointed graphs denote the same set iff they are bisimilar.*

These definitions of pointed graphs and bisimulation are almost identical to those of §2.1.1 of charts, taking a single action and no variables, and bisimulation thereof. (They differ only in that the sets involved are non-well-founded and standard, respectively, which we gloss over.) We can therefore take the closed $\mu$-expressions over a single action to denote non-well-founded sets. For more intuitive notation one could write $0$ as $\{\}$, the single action as an outfix $\{\_\}$, $+$ as $\cup$ and $\{A_1, \ldots, A_n\}$ as an abbreviation for $\{A_1\} \cup \cdots \cup \{A_n\}$, with for example

$$\begin{aligned} \mu X \; \{X\} &= \mu X \; \{\{X\}\} \\ \mu X \; \{\{\}, X\} &= \mu X \; \{\{\}, \{\{\}, X\}\}. \end{aligned}$$

The collection of non-well-founded sets denotable by $\mu$-expressions is reasonably large. It coincides with the *hereditarily finite* sets of [Acz88, Ch.1]. In [Acz88] particular non-well-founded sets are described using either pictures of pointed graphs or pointed sets of guarded equations. We expect that the simpler syntax of $\mu$-expressions is sometimes more convenient, just as it is for process algebra.

We trivially have the following corollary of Theorem 6.

**Corollary 72** *There is no finite axiomatisation of order $\leq 1$ for non-well-founded set equality of $\mu$-expressions over a single action.*

## 4.4 Weak congruence: relative axiomatisability and at higher order

In this section we show that weak congruence of $\mu$-expressions is finitely equationally axiomatisable relative to bisimulation. We also give a definition of weak higher order congruence. This enables us to prove the soundness of the new axioms simply by exhibiting some small weak higher order bisimulations.

### 4.4.1 Weak higher order congruence

The adaption of the definition of higher order bisimulation to the weak case is straightforward. The only subtlety involved is that we must require the arguments of a variable to be weak (higher order) congruent, not merely weak (higher order) bisimilar. For example, taking $x : P \to P$, the terms $x\ (a0)$ and $x\ (\tau a0)$ cannot be identified as substituting the identity for $x$ would give non-congruent terms.

**Definition** Let $R, S$ be relations over $T^P_\omega$. The relation $\mathcal{R}(R, S)$ over $T^P_\omega$ is defined by $E\ \mathcal{R}(R, S)\ F$ iff

- If $E \xrightarrow{a} E'$ then $\exists F'\ .\ F \overset{a}{\Longrightarrow} F' \wedge E'\ S\ F'$.
- If $E \rhd x\tilde{E}$ then $\exists \tilde{F}\ .\ F \xrightarrow{\tau}{}^{*} \rhd x\tilde{F} \wedge \forall i\ .\ E_i\ \hat{R}^{\sigma_i}\ F_i$.

and symmetrically. The relation $\mathcal{S}(R, S)$ over $T^P_\omega$ is defined by $E\ \mathcal{S}(R, S)\ F$ iff

- If $E \xrightarrow{a} E'$ then $\exists F'\ .\ F \overset{\hat{a}}{\Longrightarrow} F' \wedge E'\ S\ F'$.
- If $E \rhd x\tilde{E}$ then $\exists \tilde{F}\ .\ F \xrightarrow{\tau}{}^{*} \rhd x\tilde{F} \wedge \forall i\ .\ E_i\ \hat{R}^{\sigma_i}\ F_i$.

and symmetrically. For $E, F \in T^P_\omega$ we say $E$ is *weak higher order congruent* to $F$, written $E \approx^c_{ho} F$, if there exist $R, S$ such that $R \subseteq \mathcal{R}(R, S)$, $S \subseteq \mathcal{S}(R, S)$ and $E\ R\ F$.

The sound axioms can be characterised as before.

**Definition** The relation $\approx^c_{ext}$ over $T^P_\omega$ is given by

$$E \approx^c_{ext} F \quad \textit{iff} \quad \text{for all substitutions } \rho \text{ for } E, F \text{ we have } E\rho \approx^c F\rho.$$

By Corollary 31 an axiom $E = F$ is sound for $\approx^c$ precisely if $E \approx^c_{ext} F$.

In the rest of this section we again consider only terms in $T^P_1$, i.e. terms of type $P$ containing at most first order variables. We show an analogue of Theorem 8, that any weak higher order congruent pairs of terms are sound.

**Theorem 12** *If $E \approx^c_{ho} F$ then $E \approx^c_{ext} F$.*

PROOF Given $E \approx^c_{ho} F$ and a substitution $\rho$ for $E, F$ we must show that $E\rho \approx^c F\rho$. We are not going to show a non-axiomatisability result for $\approx^c$ (although one might well) and so do not need to factor this through the derived transition systems $E\star\rho$, $F\star\rho$. We adopt conditions on variables as in §3.5, in particular taking $\rho(y)$ to be of the form $\lambda z_1 : P. \cdots \lambda z_n : P.\ H_y$. Suppose that $R \subseteq \mathcal{R}(R, S)$, $S \subseteq \mathcal{S}(R, S)$, $E\ R\ F$ and that $R$ and $S$ relate only terms contained in $\mathrm{der}(E), \mathrm{der}(F)$. We let

$$\begin{array}{rlll} Q & \stackrel{\text{def}}{=} & \{\, E'\rho, F'\rho \mid E'\ (R \cup S)\ F' \,\} & \\ & \cup & \{\, H'[\tilde{E}\rho/\tilde{z}], H'[\tilde{F}\rho/\tilde{z}] \mid & \text{there exist } E' \in \mathrm{der}(E), F' \in \mathrm{der}(F) \text{ such that} \\ & & & E' \rhd y\tilde{E}, F' \rhd y\tilde{F}, \\ & & & \forall j\ .\ E_j\ R\ F_j \text{ and } H_y \longrightarrow^+ H' \,\}. \end{array}$$

It is straightforward to check that $Q$ is a weak bisimulation, using Proposition 53 and induction on the definitions of $\xrightarrow{a}_\rho$, $\rhd_\rho$. Moreover, if $E'\ R\ F'$ then $E'\rho\ \mathcal{W}(Q)\ F'\rho$. $\square$

We have no immediate application for the converse, that $E \approx^c_{ext} F$ implies $E \approx^c_{ho} F$, and so leave it and other investigation of the weak case to future work.

### 4.4.2 A finite axiomatisation of weak congruence relative to bisimulation

The new axioms are $F1$–$3$, which presented as schemas are:

$$
\begin{array}{lrcl}
F1 & \mu\left\langle \begin{array}{l} X = E + aY \\ Y = F + \tau G \end{array} \right\rangle & = & \mu\left\langle \begin{array}{l} X = E + aY + aG \\ Y = F + \tau G \end{array} \right\rangle \\
F2 & \mu\left\langle \begin{array}{l} X = E + \tau Y \\ Y = F + G \end{array} \right\rangle & = & \mu\left\langle \begin{array}{l} X = E + \tau Y + G \\ Y = F + G \end{array} \right\rangle \\
F3 & a\mu X E & = & a\mu X E + \tau X.
\end{array}
$$

We show below that these suffice to saturate (the fixed point of) any standard equation sequence. Over saturated charts weak congruence and bisimulation coincide so the relative axiomatisability result follows immediately, i.e. $F1$–$3$ together with a rule

$$F4 \qquad E \sim F \;\;\rightarrow\;\; E = F$$

are complete for $\approx^c$.

Axiom schemas $F1$–$3$ can be written as a set of pure equations between lambda terms. Unfortunately $F1$ and $F3$ involve variables ($a$) ranging over $Act$ so strictly speaking an infinite set of such equations is required. Taking $\Gamma = \{e : P \to P \to P,$ $f : P \to P \to P,\; g : P \to P \to P,\; h : P \to P\}$ and an arbitrary constant $a : P \to P$:

$$
\begin{array}{ll}
F1'_a & \Gamma \rhd \mathit{fix}\ \lambda x : P.\ (exA_1) + (aA_1) = (exA_1) + (aA_1) + (a(gxA_1)) : P \\
& \text{where } A_1 \stackrel{\text{def}}{=} \mathit{fix}\ \lambda y : P.\ (fxy) + (\tau(gxy)) \\
F2' & \Gamma \rhd \mathit{fix}\ \lambda x : P.\ (exA_2) + (\tau A_2) = (exA_2) + (\tau A_2) + (gxA_2) : P \\
& \text{where } A_2 \stackrel{\text{def}}{=} \mathit{fix}\ \lambda y : P.\ (fxy) + (gxy) \\
F3'_a & \Gamma \rhd a\,\mathit{fix}\ \lambda x : P.\ hx = a\,\mathit{fix}\ \lambda x : P.\ (hx) + (\tau x) : P
\end{array}
$$

all instances of $F1'_a$ and $F3'_a$ for $a \in Act$ are required. One could take a modified signature which has base types $P$ and $A$ and constants

$$
\begin{array}{l}
0 : P \\
a : A \;\; \text{for each } a \in Act \text{ (including } \tau) \\
. : A \to P \to P \\
+ : P \to P \to P \\
\mathit{fix}\ : (P \to P) \to P.
\end{array}
$$

The equations $F3'_a \mid a \in Act$ could then be expressed by a single equation

$$F3'' \quad \{\alpha : A,\, h : P \to P\} \triangleright \alpha.\mathit{fix}\ \lambda x : P.\ hx = \alpha.\mathit{fix}\ \lambda x : P.\ (hx) + (\tau.x) : P,$$

and similarly for $F1'_a \mid a \in Act$. One could then obtain a precise finite axiomatisability result. We do not do this because the definitions (e.g. of weak higher order congruence) would become somewhat more complex and no more enlightening. The completeness proof will be presented in terms of the schematic forms $F1$–$3$. The soundness proof shows that all instances of $F1'_a$, $F2'$ and $F3'_a$ are sound. It would presumably be straightforward to carry out both for the modified signature.

**Lemma 73** *For any $a \in Act$ the axioms $F1'_a$, $F2'$ and $F3'_a$ are sound.*

PROOF By Theorem 12 and Corollary 31 we need only show that the terms in each equation are weak higher order congruent. For $F1'_a$ let $B, C$ be the left and right hand terms of the equation respectively and let

$$R \stackrel{\text{def}}{=} \{\langle B, C\rangle,\, \langle A_1[B/x], A_1[C/x]\rangle,\, \langle gB(A_1[B/x]), gC(A_1[C/x])\rangle\}.$$

It is routine to check that $R \subseteq \mathcal{R}(R, R)$ (and hence $R \subseteq \mathcal{S}(R, R)$) so $B \approx^c_{ho} C$. The reasoning for the other axioms is similar. □

**Definition** A chart $S = \langle S, \longrightarrow, \triangleright, s_0\rangle$ is *root-unwound* if $\neg\exists s\ .\ s \xrightarrow{a} s_0$. It is *saturated* if it is root-unwound, accessible and:

1. $s \stackrel{a}{\Longrightarrow} s' \Rightarrow s \xrightarrow{a} s'$
2. $s \xrightarrow{\tau}{}^{*} \triangleright X \Rightarrow s \triangleright X$
3. $s \neq s_0 \Rightarrow s \xrightarrow{\tau} s$.

Note that any standard sequence is, when viewed as a chart, root-unwound.

**Lemma 74** *For saturated charts $S, T$, if $S \approx^c T$ then $S \sim T$.*

PROOF It is routine to check that $\approx$, restricted to $(S - s_0) \times (T - t_0)$, is a strong bisimulation and then that $s_0 \sim t_0$. □

**Lemma 75** *If for all substitutions $\rho$ with $X, X' \notin \mathrm{dom}(\rho)$ we have $\vdash A = B$, where*

$$A \stackrel{\text{def}}{=} \mu \left\langle \begin{array}{c} X = E\rho \\ X' = E'\rho \end{array} \right\rangle \qquad \textit{and} \qquad B \stackrel{\text{def}}{=} \mu \left\langle \begin{array}{c} X = E''\rho \\ X' = E'\rho \end{array} \right\rangle,$$

*then for all (possibly empty) sequences $\langle \tilde{Y} = \tilde{F} \rangle$, $\langle \tilde{Z} = \tilde{G} \rangle$ we have $\vdash L = R$, where*

$$L \stackrel{\text{def}}{=} \mu \left\langle \begin{array}{c} \tilde{Y} = \tilde{F} \\ X = E \\ X' = E' \\ \tilde{Z} = \tilde{G} \end{array} \right\rangle \qquad \textit{and} \qquad R \stackrel{\text{def}}{=} \mu \left\langle \begin{array}{c} \tilde{Y} = \tilde{F} \\ X = E'' \\ X' = E' \\ \tilde{Z} = \tilde{G} \end{array} \right\rangle.$$

PROOF By the definition of the fixed point of a sequence there is a substitution $\rho$ with domain $\tilde{Z}$ such that

$$\begin{array}{rcl} L & = & \mu \langle \tilde{Y} = \tilde{F}\rho[\mu X' E'\rho / X'][A/X] \rangle \\ R & = & \mu \langle \tilde{Y} = \tilde{F}\rho[\mu X' E'\rho / X'][B/X] \rangle. \end{array}$$

By the premise and congruence rules, for all $i$

$$\vdash F_i\rho[\mu X' E'\rho / X'][A/X] = F_i\rho[\mu X' E'\rho / X'][B/X]$$

so by Lemma 18 $\vdash L = R$. □

**Lemma 76** *For any standard sequence $\langle X_i = E_i \mid i \in m \rangle$ there is a saturated standard sequence $\langle X_i = F_i \mid i \in m \rangle$ such that $A1$–$5$,$B3$,$F1$–$3 \vdash \mu \langle X_i = E_i \mid i \in m \rangle = \mu \langle X_i = F_i \mid i \in m \rangle$.*

PROOF For a standard sequence $\langle Z_i = G_i \mid i \in m \rangle$, considered as a chart, let the measure $f\,\langle Z_i = G_i \mid i \in m \rangle$ be the pair

$$\langle \#\{\, i, a, j \mid i \stackrel{a}{\Longrightarrow} j \wedge \neg i \stackrel{a}{\longrightarrow} j \,\},\ \#\{\, i, W \mid i \stackrel{\tau}{\longrightarrow}^{*} \rhd W \wedge \neg i \rhd W \,\} \rangle.$$

We show by induction on $f\,\langle \tilde{X} = \tilde{E} \rangle$ (with the product ordering) that there exists a provably equal $\langle \tilde{X} = \tilde{F} \rangle$ satisfying conditions 1 and 2 of the definition of a saturated chart. The base case, $f\,\langle \tilde{X} = \tilde{E} \rangle = \langle 0, 0 \rangle$, is trivial. Suppose $f\,\langle \tilde{X} = \tilde{E} \rangle = \langle p+1, q \rangle$.

> **Definition** For $n \geq 0$ and states $s, s'$ of a chart say $s \stackrel{a}{\Longrightarrow}_n s'$ if there exist $n_1, n_2 \geq 0$ such that $s \stackrel{\tau}{\longrightarrow}^{n_1} \stackrel{a}{\longrightarrow} \stackrel{\tau}{\longrightarrow}^{n_2} s'$, $n_1 + n_2 = n$ and further that $n_1 + n_2$ is minimal amongst the $n'_1 + n'_2$ such that $s \stackrel{\tau}{\longrightarrow}^{n'_1} \stackrel{a}{\longrightarrow} \stackrel{\tau}{\longrightarrow}^{n'_2} s'$.

Clearly there exist $i, a, j, n \geq 1$ such that $i \overset{a}{\Longrightarrow}_n j$. In fact there exist $i, a, j$ such that $i \overset{a}{\Longrightarrow}_1 j$, as otherwise we can pick a non-minimal subsequence of the transitions in $i \overset{a}{\Longrightarrow}_n j$. It follows that $\neg(i \overset{a}{\longrightarrow} j)$ and that there is some $k$ such that either (1) $i \overset{\tau}{\longrightarrow} k \overset{a}{\longrightarrow} j$ or (2) $i \overset{a}{\longrightarrow} k \overset{\tau}{\longrightarrow} j$. Suppose (1). Then

$$\begin{aligned} E_i &= E_i + \tau X_k \\ E_k &= E_k + a X_j. \end{aligned}$$

By $\langle \tilde{X} = \tilde{E} \rangle$ standard we know that $k \neq 0 \neq j$ and by Lemma 19 w.l.g. $k = i + 1$, so $\langle \tilde{X} = \tilde{E} \rangle$ is

$$\left\langle \begin{array}{l} \tilde{Y} = \tilde{F} \\ X_i = E_i + \tau X_k \\ X_k = E_k + a X_j \\ \tilde{Z} = \tilde{G} \end{array} \right\rangle$$

for some (possibly empty) $\tilde{Y} = \tilde{F}$ and $\tilde{Z} = \tilde{G}$. By Lemma 75 this is provably equal, using $F2$, to

$$\left\langle \begin{array}{l} \tilde{Y} = \tilde{F} \\ X_i = E_i + \tau X_k + a X_j \\ X_k = E_k + a X_j \\ \tilde{Z} = \tilde{G} \end{array} \right\rangle,$$

which has measure $\langle p, q \rangle$. The case for (2) is similar but using $F1$.

Suppose $f\,\langle \tilde{X} = \tilde{E} \rangle = \langle p, q+1 \rangle$. This is similar to (1) above but instantiating the '$G$' of $F2$ by a variable rather than by a term $aX_j$.

This completes the induction. It remains only to add $\tau$ loops to ensure that condition 3 holds. This is straightforward using $F3$. □

**Theorem 13** *The axioms $F1$–$4$ are sound and complete for weak congruence ($\approx^c$).*

PROOF Soundness is immediate from Lemma 73 and the fact that $\sim \subseteq \approx^c$. For completeness suppose that $E \approx^c F$. By Lemma 20 there are standard accessible sequences $\langle \tilde{X} = \tilde{E} \rangle$, $\langle \tilde{Y} = \tilde{F} \rangle$ such that

$$\begin{aligned} AB &\vdash E = \mu \langle \tilde{X} = \tilde{E} \rangle \\ AB &\vdash F = \mu \langle \tilde{Y} = \tilde{F} \rangle \end{aligned}$$

(here $AB$ denotes the axioms $A1$–$5$,$B1$–$3$ introduced in §2.3). By Lemma 76 there are saturated sequences $\langle \tilde{X} = \tilde{E}' \rangle$, $\langle \tilde{Y} = \tilde{F}' \rangle$ such that

$$\begin{aligned} ABF1\text{–}3 &\vdash \mu\langle \tilde{X} = \tilde{E} \rangle = \mu\langle \tilde{X} = \tilde{E}' \rangle \\ ABF1\text{–}3 &\vdash \mu\langle \tilde{Y} = \tilde{F} \rangle = \mu\langle \tilde{Y} = \tilde{F}' \rangle. \end{aligned}$$

By soundness and Lemma 21 $\langle \tilde{X} = \tilde{E}' \rangle \approx^c \langle \tilde{Y} = \tilde{F}' \rangle$ so by Lemma 74 $\langle \tilde{X} = \tilde{E}' \rangle \sim \langle \tilde{Y} = \tilde{F}' \rangle$. By Lemma 21 again $\mu\langle \tilde{X} = \tilde{E}' \rangle \sim \mu\langle \tilde{Y} = \tilde{F}' \rangle$ so using $F4$ $\vdash \mu\langle \tilde{X} = \tilde{E}' \rangle = \mu\langle \tilde{Y} = \tilde{F}' \rangle$. Finally, all uses of axioms from $A$ and $B$ can, as these are sound for bisimulation, be replaced by uses of $F4$. $\square$

It follows that, if the pure implication $E3$ (from §2.3.3) could be shown sound for $\approx^c$, the axioms $A1$–$5$, $B1$–$3$, $E1$–$3$, $F1$–$3$ would be a finite pure horn clause axiomatisation for $\approx^c$. This would be as good a positive result for $\approx^c$ as could be expected, as we expect that the proof of Chapter 3 could, with some added complexity, be adapted to show that there is no finite equational axiomatisation for $\approx^c$.

# Chapter 5

# Implicational theory

In this chapter we investigate the implicational theory of finite state processes up to bisimulation. We consider sequents

$$E_1 = F_1 \wedge \cdots \wedge E_m = F_m \vdash G_1 = H_1 \vee \cdots \vee G_n = H_n$$

over $\mu$-expressions (for $m, n \geq 0$). The first two sections are devoted to showing that unification problems over $\mu$-expressions have finite computable complete sets of unifiers. The decidability of $\vdash$ is an easy corollary. These results carry over to the case of finite processes. By taking a singleton set of actions we then have finite complete sets of unifiers for certain set and non-well-founded set problems, generalising the result of [BB88]. We do not, however, have a good characterisation of the minimal complete sets of unifiers — some coarse cardinality bounds thereof are given.

These sequents do not appear to provide enough power to finitely axiomatise bisimulation (we do not give a proof of this). It was shown by Bloom and Ésik that sequents over expressions containing first order variables suffice (see §2.3.3). The decidability of such sequents is left for future work.

For a richer signature, sequents such as the above may be regarded as interesting implicit process specifications, with unification results providing automated implementation. The solving of equations between CCS expressions involving parallel composition has been studied in some depth. We refer the reader to [Liu93] in which Liu reviews previous work and shows that rather severe restrictions on the forms of equations are required for decidability (of the existence of solutions).

We note that closely related work has been done (independently) by Drost, who in [Dro94] gives a rule based unification algorithm, a proof of its correctness and a complexity analysis. We have not investigated how the sets of unifiers produced are related to those described here.

## 5.1 The consequence relation and unification

**Definition** $E_1 = F_1 \wedge \cdots \wedge E_m = F_m \vdash G_1 = H_1 \vee \cdots \vee G_n = H_n$ holds iff for all substitutions $\rho$ if $\forall i \in 1..m \,.\, E_i\rho \sim F_i\rho$ then $\exists j \in 1..n \,.\, G_j\rho \sim H_j\rho$.

The following is immediate from the definition.

**Proposition 77** $\vdash$ *is a regular, ordinary consequence relation as defined in [Avr91], i.e. the rules below for reflexivity, cut and weakening are sound and multiplicity can be neglected. Letting* $A$ *and* $\Gamma, \Delta, \Theta$ *range over equations and finite sets of equations respectively:*

$$\frac{}{A \vdash A}\ \mathit{refl}$$

$$\frac{\Gamma_1 \vdash \Delta_1, A \quad A, \Gamma_2 \vdash \Delta_2}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\ cut$$

$$\frac{\Gamma \vdash \Delta}{\Theta, \Gamma \vdash \Delta}\ w-l \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \Theta}\ w-r$$

We recall some standard definitions from [BS93], instantiating them to the setting of bisimulation over $\mu$-expressions.

**Definition** A *unification problem* is a finite set $\{E_1 = F_1, \ldots, E_n = F_n\}$ of equations between $\mu$-expressions. A *unifier* for this is a substitution $\rho$ satisfying all the equations up to bisimulation, i.e. such that $\forall i \in 1..n \,.\, E_i\rho \sim F_i\rho$. The set $U$ of all unifiers can be ordered by $\sigma < \theta$ iff there exists a substitution $\lambda$ such that for all variables $x$ free in the problem $\sigma(x)\lambda \sim \theta(x)$. (This is clearly a preorder.) A *complete* set $C \subseteq U$ of unifiers is one such that $\forall \theta \in U \,.\, \exists \sigma \in C \,.\, \sigma < \theta$. A *minimal complete* set is a complete set for which $<$ is a discrete order.

**Lemma 78** *If for any unification problem there is a finite computable complete set of unifiers then the consequence relation is decidable.*

PROOF Consider the sequent

$$E_1 = F_1 \wedge \cdots \wedge E_m = F_m \vdash G_1 = H_1 \vee \cdots \vee G_n = H_n$$

and let $C$ be a finite complete set of unifiers for the problem $\{E_1 = F_1, \ldots, E_m = F_m\}$. We show that the sequent holds iff $\forall \sigma \in C \,.\, \exists j \in 1..n \,.\, G_j\sigma \sim H_j\sigma$. The left to right direction is immediate from the definitions. For the other, consider an arbitrary substitution $\rho$ such that $\forall i \in 1..m \,.\, E_i\rho \sim F_i\rho$. By definition $\rho$ is a unifier so there is some $\sigma \in C \,.\, \sigma < \rho$, i.e. $\exists \lambda \,.\, \forall x \,.\, \rho(x) \sim \sigma(x)\lambda$. By the premise $\exists j \in 1..n \,.\, G_j\sigma \sim H_j\sigma$ but bisimulation is substitutive so $G_j\sigma\lambda \sim H_j\sigma\lambda$. □

## 5.2 Unification for charts

It is convenient to work with transition systems rather than directly with $\mu$-expressions. Accordingly we define charts and chart substitution. These are essentially as in §2.1.1, differing only in that for this chapter charts have no start state.

**Definition** A *chart* is a tuple $\langle S, \longrightarrow, \triangleright \rangle$ where $S$ is a set of states, $\longrightarrow \subseteq S \times \mathit{Act} \times S$ is a transition relation and $\triangleright \subseteq S \times \mathit{Var}$ is a visibility predicate.

Except where otherwise stated we consider only finite state charts, i.e. those for which all three components are finite. Substitution over charts is as expected — we give an explicit definition for concreteness.

**Definition** A *chart substitution* $\tau$ is a chart (say $T$) together with a function $\tau$ from a finite subset $A$ of $\mathit{Var}$ to the state set of $T$. The result of a substitution $S\tau$ is a chart with states $S \uplus T$ and the least transition and visibility relations satisfying the following clauses.

$$\frac{s \xrightarrow{a} s'}{\mathrm{inl}(s) \xrightarrow{a} \mathrm{inl}(s')} \qquad \frac{t \xrightarrow{a} t'}{\mathrm{inr}(t) \xrightarrow{a} \mathrm{inr}(t')} \qquad \frac{s \triangleright y \quad y \in A \quad \tau(y) \xrightarrow{a} t'}{\mathrm{inl}(s) \xrightarrow{a} \mathrm{inr}(t')}$$

$$\frac{s \triangleright x \notin A}{\mathrm{inl}(s) \triangleright x} \qquad \frac{t \triangleright x}{\mathrm{inr}(t) \triangleright x} \qquad \frac{s \triangleright y \in A \quad \tau(y) \triangleright x}{\mathrm{inl}(s) \triangleright x}$$

**Notation** We will often omit the injections into disjoint unions. If $s$ is a state of a chart $S$ then $s\sigma$ is the corresponding state in $S\sigma$.

**Definition** If $S$ is a chart $\langle S, \longrightarrow, \rhd \rangle$ then $Var(S)$ is the range of $\rhd$, i.e. the set of variables visible to any state within it and $Act(S)$ is the set of actions occurring in it. If $s$ is a state of a chart then $\mathrm{vis}(s) \stackrel{\text{def}}{=} \{ x \mid s \rhd x \}$.

**Definition** A *unification problem* is a finite set $\{ s_i = s_i' \mid i \in I \}$ of pairs of states within a chart. A *unifier* for such is a substitution $\theta$ with domain $Var(S)$ such that $\forall i \in I \,.\, s_i\theta \sim s_i'\theta$.

**Notation** From now on we fix a unification problem as above. We let $R$ range over pairs of a symmetric relation $R \subseteq S \times S$ and a relation $\longrightarrow_R \subseteq Var(S) \times Act \times S$.

**Definition** For any $R$ we define a substitution $\sigma_R$ as follows. First, we introduce new variables $V_A$ indexed by nonempty sets $A$ of the variables occurring in $S$ and put

$$\mathcal{V}_R \stackrel{\text{def}}{=} \{ V_A \mid \forall s, s' \in S \,.\, s \mathrel{R} s' \Rightarrow ((A \cap \mathrm{vis}(s) \neq \{\}) \iff (A \cap \mathrm{vis}(s') \neq \{\})) \}.$$

We define a chart $S_2 \stackrel{\text{def}}{=} \langle S \uplus Var(S), \longrightarrow \cup \rhd \longrightarrow_R \cup \longrightarrow_R, \rhd_R \rangle$ where

$$\begin{array}{rcl} s \rhd_R V_A & \iff & V_A \in \mathcal{V}_R \wedge \exists x \,.\, s \rhd x \in A \\ x \rhd_R V_A & \iff & V_A \in \mathcal{V}_R \wedge x \in A. \end{array}$$

The substitution $\sigma_R$ takes a variable $x \in Var(S)$ to the state $x$ of $S_2$.

**Definition** $R$ is a *solution bisimulation* if:

1. $R$ is a bisimulation on the chart $\langle S, \longrightarrow \cup \rhd \longrightarrow_R, \{\} \rangle$.

2. $\forall i \in I \,.\, s_i \mathrel{R} s_i'$.

**Lemma 79** *The set of solution bisimulations is finite and computable.*

Proof There are a bounded number of relations $R, \longrightarrow_R$ and it is decidable whether each is a solution bisimulation. □

**Lemma 80** *For any solution bisimulation $R$, the substitution $\sigma_R$ is a unifier.*

PROOF The chart $S\sigma_R$ has states $S \uplus (S \uplus \mathit{Var}(S))$. We let $\mathrm{inl}(s), \mathrm{inr}(s), x$ range over these components respectively. Letting $\alpha, \beta$ range over $\{l, r\}$ it is straightforward to show that it has the transitions generated by

$$\begin{aligned} s \xrightarrow{a} s' &\Rightarrow \mathrm{in}\alpha(s) \xrightarrow{a} \mathrm{in}\alpha(s') \\ s \rhd \xrightarrow{a}_R s' &\Rightarrow \mathrm{in}\alpha(s) \xrightarrow{a} \mathrm{inr}(s') \\ x \xrightarrow{a}_R s' &\Rightarrow x \xrightarrow{a} \mathrm{inr}(s') \end{aligned}$$

and the visibilities generated by

$$\begin{aligned} s \rhd_R V_A &\Rightarrow \mathrm{in}\alpha(s) \rhd V_A \\ x \rhd_R V_A &\Rightarrow x \rhd V_A. \end{aligned}$$

It is then straightforward to check that

$$Q \stackrel{\text{def}}{=} \{ \mathrm{in}\alpha(s), \mathrm{in}\beta(s') \mid s \; R \; s' \}$$

is a bisimulation with $\forall i \in I \; . \; \mathrm{inl}(s_i) \; Q \; \mathrm{inl}(s_i')$. □

This gives us a finite computable set of unifiers:

$$\{ \sigma_R \mid R \text{ is a solution bisimulation} \}.$$

It remains to show that this is complete.

**Definition** A substitution $\theta$ with domain $\mathit{Var}(S)$ *satisfies* an $R$ (which is not necessarily a solution bisimulation) if:

3. $\forall s, t \in S \; . \; s \; R \; t \Rightarrow s\theta \sim t\theta$.

4. $\forall x \in \mathit{Var}(S), s \in S \; . \; x \xrightarrow{a}_R s \iff \theta(x) \xrightarrow{a} \sim s\theta$.

**Lemma 81** *If $\theta$ is a substitution with domain $\mathit{Var}(S)$ that satisfies $R$ then $\theta$ factors through the substitution $\sigma_R$.*

PROOF

> **Definition** Given a finite set $P$ of processes let their intersection $\bigcap P$ be some minimal member of the bisimulation class of
>
> $$\sum \{ a.q \mid \forall p \in P \; . \; p \xrightarrow{a} \sim q \} + \sum \{ x \mid \forall p \in P \; . \; p \rhd x \}.$$
>
> If all elements of $P$ are finite state then $\bigcap P$ is also finite state.

**Definition** The substitution $\lambda$ has range $\mathcal{V}_R$ as defined above and for $V_A \in \mathcal{V}_R$

$$\lambda(V_A) \stackrel{\text{def}}{=} \bigcap\{\, \theta(x) \mid x \in A \,\}.$$

We will show for all $x \in \mathit{Var}(S)$ that $\theta(x) \sim \sigma_R(x)\lambda$. First we give two lemmas exploiting the structure of the set $\mathcal{V}_R$ of variables.

**Lemma 82** *If $\theta(x) \xrightarrow{a} p$ and $\neg\exists s \in S \,.\, s\theta \sim p$ then $\exists V_A \in \mathcal{V}_R \,.\, x \in A \wedge \forall y \in A \,.\, \theta(y) \xrightarrow{a} \sim p$.*

**Lemma 83** *If $\theta(x) \rhd z$ then $\exists V_A \in \mathcal{V}_R \,.\, x \in A \wedge \forall y \in A \,.\, \theta(y) \rhd z$.*

PROOF Only the first is given here — the other is similar. We show that there exists a strictly increasing sequence of sets $A_k \subseteq \mathit{Var}(S)$ with $A_0 \stackrel{\text{def}}{=} \{x\}$ all of which satisfy

$$\forall y \in A_k \,.\, \theta(y) \xrightarrow{a} \sim p.$$

Suppose $V_{A_k} \notin \mathcal{V}_R$, otherwise we are done. There are then some $s, s' \in S$ and $y$ such that

$$\begin{aligned} &s \; R \; s' \\ &y \in A_k \cap \mathrm{vis}(s) \\ &\{\} = A_k \cap \mathrm{vis}(s'). \end{aligned}$$

By induction $\theta(y) \xrightarrow{a} \sim p$ so $s\theta \xrightarrow{a} \sim p$. By clause 3 $s'\theta \xrightarrow{a} \sim p$ so either

- $s' \xrightarrow{a} s'' \wedge s''\theta \sim p$, contradicting the premise, or
- $\exists y' \in \mathrm{vis}(s') - A_k \,.\, \theta(y') \xrightarrow{a} \sim p$, in which case we put $A_{k+1} \stackrel{\text{def}}{=} A_k \cup \{y'\}$.

$\square$

It is useful to use a simple 'bisimulation up to' technique.

**Definition** If $\langle S, \longrightarrow_S, \rhd_S \rangle, \langle T, \longrightarrow_T, \rhd_T \rangle$ are charts then $Q \subseteq S \times T$ is a bisimulation up to $\sim$ if $s \; Q \; t$ implies

- $s \xrightarrow{a} s' \Rightarrow \exists t' \,.\, t \xrightarrow{a} t' \wedge s' \; (\sim \cup \sim Q \sim) \; t'$
- $s \rhd_S x \Rightarrow t \rhd_T x$

and symmetrically.

**Proposition 84** *If $Q$ is a bisimulation up to $\sim$ then $\sim \cup \sim Q \sim$ is a bisimulation.*

PROOF Straightforward. □

Now let

$$\begin{aligned} S_1 &\stackrel{\text{def}}{=} \langle S \cup \mathit{Var}(S), \longrightarrow, \rhd \cup \mathrm{id}_{\mathit{Var}(S)} \rangle \\ S_2 &\stackrel{\text{def}}{=} \langle S \cup \mathit{Var}(S), \longrightarrow \cup \rhd \longrightarrow_R \cup \longrightarrow_R, \rhd_R \rangle \\ Q &\stackrel{\text{def}}{=} \mathrm{id}_{S \cup \mathit{Var}(S)}. \end{aligned}$$

We check below that $Q$ is a bisimulation up to $\sim$ between $S_1\theta$ and $S_2\lambda$. The result follows by noting that the state $x$ of $S_1$ is bisimilar to the one state chart '$x$' and that the state $x$ of $S_2$ is $\sigma_R(x)$ by definition.

We give only the transition part for pairs $\langle s, s \rangle \in Q$ — the rest is similar. There are two cases for transitions of $s\theta$:

- $s \xrightarrow{a} s' \wedge s\theta \xrightarrow{a} s'\theta$. Then $s\lambda \xrightarrow{a} s'\lambda$.
- $s \rhd x \wedge \theta(x) \xrightarrow{a} p \wedge s\theta \xrightarrow{a} p$. Suppose $\exists s' \in S. s'\theta \sim p$. Then by 4$\Leftarrow$ $x \xrightarrow{a}_R s'$ so $s \xrightarrow{a}_{S_2} s'$ and $s\lambda \xrightarrow{a} s'\lambda \; Q^{-1} \; s'\theta \sim p$. Now suppose not. Then by Lemma 82 $\exists V_A \in \mathcal{V} \,.\, x \in A \wedge \lambda(V_A) \xrightarrow{a} \sim p$ so $s \rhd_R V_A$ and $s\lambda \xrightarrow{a} \sim p$.

For transitions of $s\lambda$ there are three cases:

- $s \xrightarrow{a} s' \wedge s\lambda \xrightarrow{a} s'\lambda$. Then $s\theta \xrightarrow{a} s'\theta$.
- $s \rhd x \xrightarrow{a}_R s' \wedge s\lambda \xrightarrow{a} s'\lambda$. Then by 4 $\Rightarrow$ $\theta(x) \xrightarrow{a} \sim s'\theta$ so $s\theta \xrightarrow{a} \sim s'\theta \; Q \; s'\lambda$.
- $s \rhd_R V_A \wedge \lambda(V_A) \xrightarrow{a} q \wedge s\lambda \xrightarrow{a} q$. Then by the definitions of $\rhd_R$ and $\lambda$ $s\theta \xrightarrow{a} \sim q$.

□

**Remark** It is not the case that for any solution bisimulation $\sigma_R$ satisfies $R$ — the right-to-left implication of 4 may fail, essentially because $\longrightarrow_R$ is too small. For example take $S \stackrel{\text{def}}{=} \langle \{1,2,3\}, \{\}, \{\langle 1, x \rangle\} \rangle$, the problem $\{2 = 3\}$, $R \stackrel{\text{def}}{=} \{\langle 2,3 \rangle, \langle 3,2 \rangle\}$

and $\longrightarrow_R \stackrel{\text{def}}{=} \{x \xrightarrow{a} 2\}$. Here $\mathcal{V}_R = \{V_{\{x\}}\}$ and $\sigma_R(x) = V_{\{x\}} + a0$. We have $\sigma_R(x) \xrightarrow{a} \sim (3\sigma_R)$ but $\neg(x \xrightarrow{a}_R 3)$.

**Lemma 85** *If $\theta$ is a unifier then there is a solution bisimulation $R(\theta)$ that is satisfied by it.*

PROOF Define $R, \longrightarrow_R$ by:

3. $\forall s, t \in S \,.\, s \; R \; t \iff s\theta \sim t\theta$.

4. $\forall x \in \mathit{Var}(S), s \in S \,.\, x \xrightarrow{a}_R s \iff \theta(x) \xrightarrow{a} \sim s\theta$.

Trivially, $R$ is symmetric (and reflexive), is satisfied by $\theta$ and $\forall i \in I \,.\, s_i \; R \; s_i'$. It remains to check that $R$ is a bisimulation on $\langle S, \longrightarrow \cup \rhd \longrightarrow_R, \{\}\rangle$. Suppose $s\theta \sim t\theta$. There are two possible transitions:

- $s \xrightarrow{a} s'$, in which case $s\theta \xrightarrow{a} s'\theta$.
- $s \rhd y \xrightarrow{a}_R s'$, in which case $\theta(y) \xrightarrow{a} \sim s'\theta$ so $s\theta \xrightarrow{a} \sim s'\theta$.

In either case $t\theta \xrightarrow{a} \sim s'\theta$ and one of the following holds.

- $t \xrightarrow{a} t' \wedge t'\theta \sim s'\theta$. Then $t \; (\xrightarrow{a} \cup \rhd \xrightarrow{a}_R) \; t' \; R \; s'$.
- $t \rhd y \wedge \theta(y) \xrightarrow{a} \sim s'\theta$. Then $y \xrightarrow{a}_R s'$ so $t \; (\xrightarrow{a} \cup \rhd \xrightarrow{a}_R) \; s' \; R \; s'$.

□

**Theorem 14** *The set $\{\, \sigma_R \mid R$ is a solution bisimulation$\,\}$ is a finite computable complete set of unifiers.*

PROOF It is finite and computable by Lemma 79 and contains only unifiers by Lemma 80. For completeness suppose $\theta$ is an arbitrary unifier. By Lemma 85 there is a solution bisimulation $R$ satisfied by it, then by Lemma 81 $\theta$ factors through $\sigma_R$. □

**Corollary 86** *The consequence relation $\vdash$ over sets of equations between $\mu$-expressions is decidable.*

PROOF By Lemma 78. □

**Remark** The conjunction can be internalised, i.e.

$$E_1 = F_1 \wedge \cdots \wedge E_m = F_m \vdash G_1 = H_1 \vee \cdots \vee G_n = H_n$$

holds iff

$$\left(\sum_{i \in 1..m} a_i E_i = \sum_{i \in 1..m} a_i F_i\right) \vdash G_1 = H_1 \vee \cdots \vee G_n = H_n$$

does.

## 5.3 Minimal sets of unifiers

The complete sets of the previous section can be reduced somewhat, as follows.

**Corollary 87** *The set of all $\sigma_{R,\longrightarrow_R}$ where*

- $R, \longrightarrow_R$ *is a solution bisimulation.*
- $R$ *is an equivalence relation containing bisimulation over $S$.*
- $R, \longrightarrow_R$ *is minimal with respect to set inclusion among the $R', \longrightarrow_R$ satisfying the above.*

*is a complete set of unifiers.*

PROOF From Theorem 14 and the observations that for $\langle R, \longrightarrow_R\rangle, \langle R', \longrightarrow_R\rangle$ if $R \subseteq R'$ then $\sigma_R < \sigma_{R'}$ and if $R'$ is the reflexive symmetric transitive closure of $R$ then $\sigma_R = \sigma_{R'}$. □

These may still not be minimal. For example the set for the problem $a0+x = a0+x$ contains both $x \mapsto V_{\{x\}}$ and $x \mapsto a0+V_{\{x\}}$. We do not have a nice characterisation of the minimal sets (in contrast to that of [BB88] for certain set equations) but coarse bounds on their size can be easily given. An upper bound from counting the number of relations $R, \longrightarrow_R$ is

$$2^{|S| \times |S|} \times 2^{|\mathit{Var}(S)| \times |\mathit{Act}(S)| \times |S|}.$$

For an exponential lower bound, the problem

$$a0 \quad = \quad a0 + \sum_{i \in 1..n} x_i$$

has $2^n$ incomparable unifiers.

## 5.4 Unification for finite processes

Theorem 14 further implies the existence of finite complete sets of unifiers for the finite process expressions, i.e. those given by

$$E ::= 0 \mid X \mid aE \mid E + E.$$

The reasoning is general and may well be known. We suppose $\sim$ is a substitutive congruence and consider a subset $T$ of terms satisfying the following (allowing $E, F$ to range over all $\mu$-expressions).

- $Var \subseteq T$
- $E, F \in T \Rightarrow E[F/x] \in T$
- $E[F/x] \sim\in T \Rightarrow \exists E' \in T \,.\, E \sim E'$
- $E[F/x] \sim\in T \wedge x \in FV(E) \Rightarrow \exists F' \in T \,.\, F \sim F' \wedge FV(F') \subseteq FV(F)$

The finite process expressions and bisimulation clearly satisfy these conditions.

**Definition** A $T$-*unification problem* is a finite set $\{E_1 = F_1, \ldots, E_n = F_n\}$ of equations between members of $T$. A $T$-*unifier* for this is a substitution $\rho$ with range in $T$ satisfying all the equations up to bisimulation, i.e. such that $\forall i \in 1..n \,.\, E_i\rho \sim F_i\rho$. The set $U_T$ of all $T$-unifiers can be ordered by $\sigma <_T \theta$ iff there exists a substitution $\lambda$ with range in $T$ such that for all variables $x$ free in the problem $\sigma(x)\lambda \sim \theta(x)$. (This is clearly a preorder.) A $T$-*complete* set $C \subseteq U_T$ of $T$-unifiers is one such that $\forall \theta \in U_T \,.\, \exists \sigma \in C \,.\, \sigma_T <_T \theta$.

**Proposition 88** *If $C$ is complete for a $T$-problem and satisfies*

$$\forall \sigma \in C \,.\, \forall E \in \mathrm{ran}(\sigma) \,.\, E \sim\in T \Rightarrow E \in T$$

*then $C \cap U_T$ is $T$-complete for it.*

PROOF The last two conditions above imply their vector forms, i.e.

- $E[\tilde{F}/\tilde{x}] \sim\in T \Rightarrow \exists E' \in T \,.\, E \sim E'$

- $\tilde{E}[\tilde{F}/\tilde{x}] \sim\in T^m \wedge \tilde{x} \subseteq FV(\tilde{E}) \Rightarrow \exists \tilde{F}' \in T^n \,.\, \forall i \,.\, F_i \sim F_i' \wedge FV(F_i') \subseteq FV(F_i)$

by induction on the length of $\tilde{x}$. Now suppose $\theta$ is a $T$-unifier. By $C$ complete there is a substitution $\lambda$ such that $\forall x.\sigma(x)\lambda \sim \theta(x)$. By the second and first vector forms we can find $\lambda'$ and $\sigma'$ with ranges in $T$. The premise then implies that the range of $\sigma$ is also in $T$ and we have $\sigma <_T \theta$. □

**Corollary 89** *There are finite computable complete sets of unifiers for problems over the finite process expressions.*

PROOF It remains only to note that if $\exists E' \in T \,.\, E' \sim E$ then such an $E'$ can be constructed. □

**Remark** The consequence relation $\vdash_f$ obtained by restricting to finite process expressions is strictly larger than $\vdash$. The inclusion is trivial and for example $ax = x \vdash_f a0 = 0$ but $ax = x \not\vdash a0 = 0$.

**Remark** Lemmas 81 and 85 hold for infinite state unifiers $\theta$ (of finite state problems) so the set of unifiers given in Theorem 14 is also complete for such and the consequence relation $\vdash_i$ obtained by allowing arbitrary charts coincides with $\vdash$ over the finite state problems.

# Chapter 6

# Conclusion

The main contributions of the thesis are, briefly, as follows.

In Chapter 2 a conceptually simple (albeit notationally complex) infinite set of equation schemes over $\mu$-expressions was given and shown to be complete for bisimulation. This gave rise to conjectures about the completeness of various subsets, one of which led to the result of Chapter 3.

In Chapter 3 a simply typed lambda calculus was introduced in which a large class of equation schemes over $\mu$-expressions can be expressed as equations. A higher order bisimulation was defined over the lambda terms, shown to be decidable and (when restricted to terms with first order variables) to coincide with a more extensional equality. Examination of the coincidence showed that equational axioms preserve certain loop properties and hence that no finite, equational, first order axiomatisation for bisimulation over the $\mu$-expressions can exist.

In Chapter 4 it was shown that no finite equational axiomatisation for bisimulation over sets of $*$-expressions containing a zero process can exist, in sharp contrast to the extant positive result for BPA$^*$. The lambda terms were viewed as a fragment of the higher order $\pi$ calculus and higher order bisimulation was shown to coincide with the normal bisimulation of Sangiorgi. The work was related to the theory of non-well-founded sets. Finally, a finite equational axiomatisation of weak congruence relative to bisimulation was given. The soundness proof made use of a definition of weak higher order congruence.

In Chapter 5 it was shown that unification problems over $\mu$-expressions have finite computable complete sets of unifiers and hence that certain sequents over $\mu$-expressions are decidable.

In the rest of this chapter we discuss possible future work, including the several problems that have been left open.

## 6.1 Higher order bisimulation

In Chapter 3 we were able to show the coincidence of $\sim_{ho}$ and $\sim_{ext}$ only for terms containing variables of at most first order types. We conjecture that the coincidence holds for all terms but have been unable to prove either inclusion. Our proof for the first order case depended on an inductive characterisation of the transitions of a substituted term $E\rho$ (Proposition 53) which in turn depended on a tight relationship between $\beta$-reduction and transitions (Lemma 56). These cannot be simply generalized to the general case, for which we expect that a more sophisticated induction (corresponding to a normalisation proof for general $\beta$ reduction) is required. It is natural to consider the lifting of $\sim_{ho}$ and $\sim_{ext}$ to logical relations over terms of all types (containing variables of arbitrary types). This enables a number of related conjectures to be stated.

We recall from [Mit90] the definition of logical relation over the typed applicative structure of terms in the $T^{\sigma}_{\omega}$.

**Definition** If $\sim$ is a relation over $T^{P}_{\omega}$ then

$$\begin{aligned} \overset{\text{LR}^P}{\sim} &\overset{\text{def}}{=} \sim \\ E \overset{\text{LR}^{\sigma\to\tau}}{\sim} F &\overset{\text{def}}{=} \forall E', F' \in T^{\sigma}_{\omega} \,.\, E' \overset{\text{LR}^{\sigma}}{\sim} F' \Rightarrow EE' \overset{\text{LR}^{\tau}}{\sim} FF'. \end{aligned}$$

The relations $\overset{\text{LR}^{\sigma}}{\sim}$ are a *logical relation* if for each constant $c:\sigma$ (here with $c$ one of $0, a, +, \mathit{fix}$ ) we have $c \overset{\text{LR}^{\sigma}}{\sim} c$.

It is immediate from Propositions 34 and 51 that $\overset{\text{LR}^{\sigma}}{\sim}_{ext}$ and $\overset{\text{LR}^{\sigma}}{\sim}_{ho}$ are logical partial equivalence relations. The conjectures can now be stated, in each of which the types $\sigma$ and $\tau$ are universally quantified.

**Conjecture 90**

1. $\overset{\text{LR}^{\sigma}}{\sim}_{ext}$ *is reflexive.*

2. $\overset{\text{LR}^\sigma}{\sim}_{ho}$ *is reflexive.*

3. *If* $E \sim^\sigma_{ho} F$ *and* $G : \tau$ *then* $E[G/x] \sim^\sigma_{ho} F[G/x]$.

4. $\sim^\sigma_{ho} = \overset{\text{LR}^\sigma}{\sim}_{ho}$.

5. $\sim_{ho} \subseteq \sim_{ext}$.

6. $\sim_{ho} = \sim_{ext}$.

7. $\sim^\sigma_{ho} = \overset{\text{LR}^\sigma}{\sim}_{ext}$.

There are a number of implications between these, mostly trivial. We note that using Proposition 51 clause 3 can be shown to be equivalent to the congruence of $\sim^\sigma_{ho}$ for application on the left or right. Further, clause 3 implies clause 5 and is implied by clause 6. We note also that clause 4 is equivalent to the conjunction of clauses 2 and 3, using straightforward inductions on types.

A possible starting point for the inclusion $\sim_{ho} \subseteq \sim_{ext}$ is therefore the proof by Sangiorgi that context bisimulation is a congruence for application [San94, Appendix B]. The work in [Sta82], in which $\beta\eta$ equality is reduced to $\beta\eta$ equality at a particular type, may also be relevant.

For the inclusion $\sim_{ext} \subseteq \sim_{ho}$ a simple proof would need, for any $E \sim_{ext} F$, a substitution $\rho$ for $E, F$ such that a higher order bisimulation relating $E$ and $F$ can be extracted from a bisimulation relating $E\rho$ and $F\rho$. A sample problematic case is the pair of terms

$$\begin{array}{rcl} E & \overset{\text{def}}{=} & x\ (\lambda w : P.\ x\ (\lambda u : P.\ cw + du)) \\ F & \overset{\text{def}}{=} & x\ (\lambda w : P.\ x\ (\lambda u : P.\ cu + dw)), \end{array}$$

where $x : (P \to P) \to P$, which are not higher order bisimilar. This rules out substitutions of the form

$$\rho(x) \overset{\text{def}}{=} \lambda y : P \to P.\ a_x(y\ G),$$

with $y$ not free in $G$, as for these $E\rho \sim F\rho$.

There are three other obvious developments of higher order bisimulation.

In §3.4 higher order bisimulation was shown to be decidable by the exponential technique of checking whether any relation over a given set is a loose bisimulation. One could look for a tighter complexity result, based on the partition refinement algorithm of Paige and Tarjan [PT87].

For higher order process calculi one is interested in equivalences that abstract from internal actions. In §4.4 we defined a weak higher order congruence, $\approx^c_{ho}$, that does this. One could develop the theory of $\approx^c_{ho}$, in particular by showing a converse to Theorem 12, i.e. showing that $\approx^c_{ext} \subseteq \approx^c_{ho}$, and exhibiting a tight connection between $\approx^c_{ho}$ and the equivalences over the higher order $\pi$ calculus. The literature contains many different equivalences over transition systems that abstract from internal actions. There is at least as much scope for different higher order equivalences, and even less understanding of the criteria for choosing between them.

Throughout this thesis we have considered only sequential nondeterministic processes from very simple signatures. For these it is possible to characterise the extensional equality $\sim_{ext}$ over open terms by a bisimulation involving visibility judgments, resting on results such as

$$E[F/X] \xrightarrow{a} A \quad \textit{iff} \quad (E \xrightarrow{a} E' \wedge E'[F/X] = A) \vee (E \rhd X \wedge F \xrightarrow{a} A)$$

for the $\mu$-expressions and Proposition 53 for lambda terms with first order variables. There does not appear to be an analogous result for terms involving parallel composition (although formulating a precise result to that effect is nontrivial). One could investigate the class of GSOS-definable (say) operators for which there are analogues of the visibility judgment and higher order bisimulation — perhaps something like those for which the target of any transition depends only on the 'near future' behaviour of all except one argument.

## 6.2 Axiomatisability

The space of possible finite axiomatisability results is very large. They can be classified along three main dimensions.

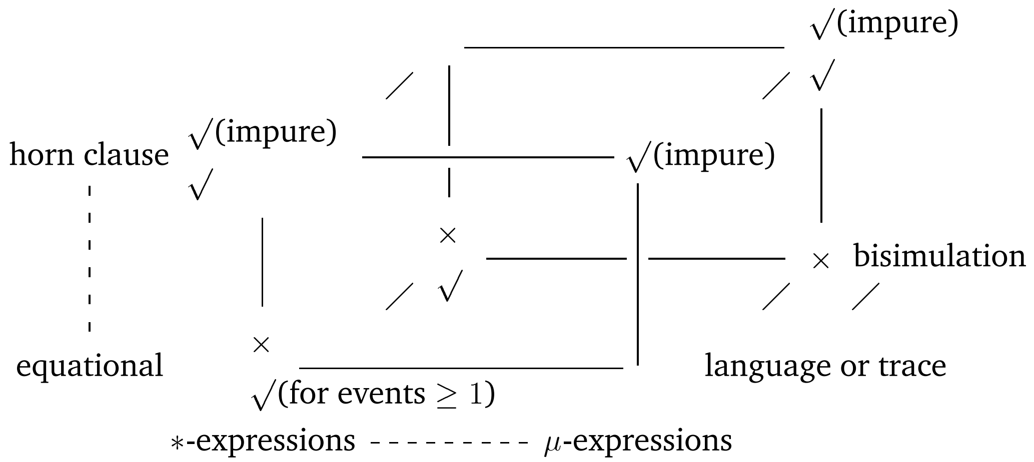
Syntax: we have considered $\mu$-expressions and several classes of $*$-expressions.

Equivalence: this can be loosely subdivided into a linear-branching time dimension and a dimension of abstraction from internal actions. We have considered infinite term equality, (strong) bisimulation, weak congruence and (strong) language/trace congruence.

Logic: we have considered equational and horn clause (both pure and impure) axiomatisations.

For axiomatisations over $\mu$-expressions there are further dimensions, namely the orders of free variables that may occur in axioms and whether variables over actions are admitted. We have generally considered only the first order case.

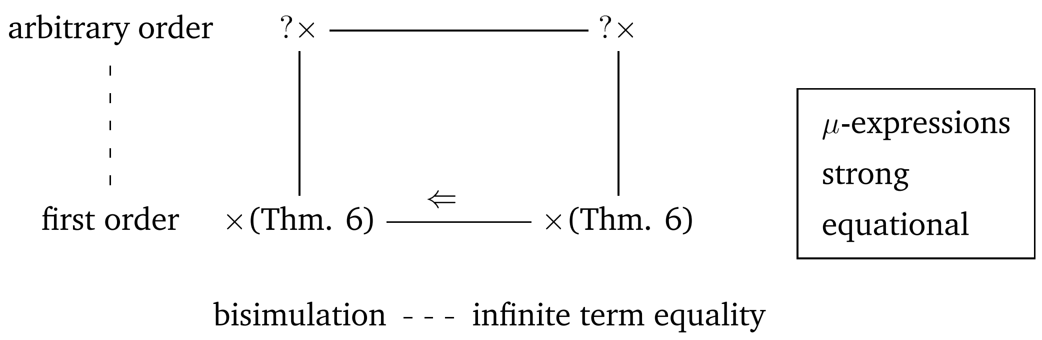
Some of the known results were summarised in Figure 2–2, page 21. It is reproduced below, omitting references. The results shown are for strong equivalences, that do not abstract from internal actions, and the results for $\mu$-expressions are for axioms with at most first order variables and without variables over actions.




There are a great many conjectures suggested by this and by the multi-dimensional diagram of which it is a slice. A few of these seem to have some compelling interest, albeit largely technical. They are discussed below. It should be noted that bisimulation of $\mu$-expressions is very close to infinite term equality of terms from an arbitrary first order signature with recursion. Some of these questions could be posed in that more general setting.

In the diagrams below conjectures are indicated by '?', sometimes annotated by $\surd$ or $\times$ if we expect a positive or negative result. Relative axiomatisability results are indicated by double arrows $\Rightarrow$.

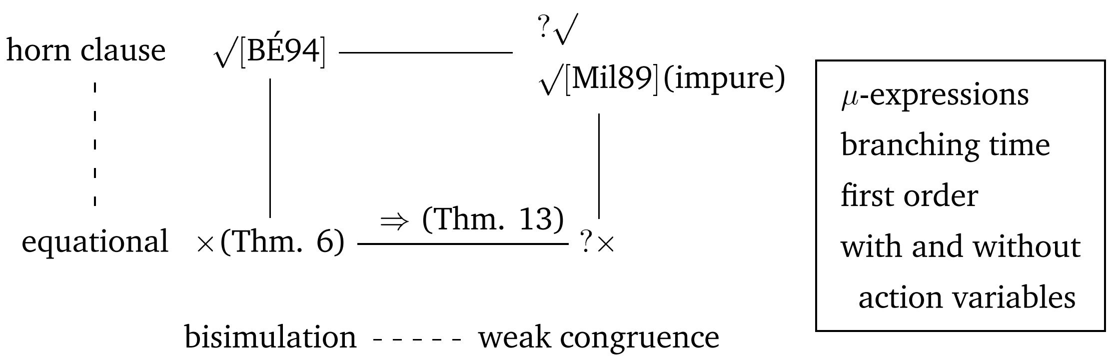
**With variables of arbitrary order**




We would like to generalise Theorem 6 to allow variables of arbitrary order, not simply those of first order. A proof would presumably depend on a proof of the coincidence of $\sim_{ho}$ and $\sim_{ext}$ for all terms, as discussed in the previous section. For simplicity one could first consider infinite term equality. In the first order case this might enable the (somewhat complex) proof of Theorem 6 to be simplified and in the arbitrary order case be a useful stepping stone towards the result for bisimulation.

A further question is the existence of good (possibly infinitary) $\omega$-complete axiomatisations for bisimulation, i.e. axiomatisations complete for $\sim_{ext}$ over $T_\omega^P$ or $T_1^P$, not just over $T_0^P$.

**Abstracting from $\tau$**




The left edge is the back right edge of the cube. As discussed in §4.4.2, to express interesting axioms about weak congruence as pure typed equations we need to modify the signature slightly, adding a base type $A$ of actions and taking constants

$$\begin{aligned}
&0 : P\\
&a : A \;\; \text{for each } a \in Act \text{ (including } \tau)\\
&. : A \to P \to P\\
&+ : P \to P \to P\\
&\mathit{fix} : (P \to P) \to P.
\end{aligned}$$

We conjecture that the proof of Theorem 6 could be adapted to this signature without difficulty and to weak congruence at the cost of some uninteresting complications. A more interesting result would be to show that axiom $E3$, used in the pure horn clause axiomatisation of bisimulation in [BÉ94] and reproduced in §2.3.3, is sound for weak congruence. Together with the result of [BÉ94] and Theorem 13 (the relative axiomatisation of weak congruence over bisimulation) this would give a pure horn clause axiomatisation for weak congruence.

### Linear time

$$\begin{array}{lccc}
\text{horn clause} & \begin{array}{c}?\surd\\ \surd[\text{Rab93}]\,(\text{impure})\end{array} & \text{———————} & \surd[\text{BÉ94}] \\
\vdots & | & & | \\
\text{equational} & ?\times & \overset{?\Leftarrow}{\text{———————}} & \times\,(\text{Thm. 6}) \\
 & \text{trace congruence} & \text{- - - - - -} & \text{bisimulation}
\end{array}$$

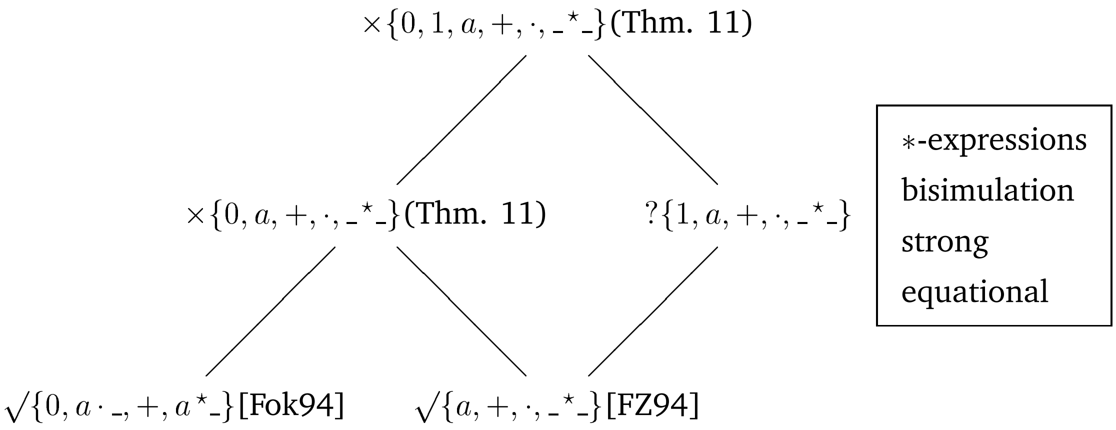
$\mu$-expressions
strong
first order
with and without
action variables

This is the right hand face of the cube. The non-axiomatisability proof of Theorem 6 for bisimulation of $\mu$-expressions and that reproduced in §2.3.4 from [Con71] for language equivalence of $*$-expressions are based on similar intuitions, that a finite set of axioms cannot introduce an arbitrary prime factor into the length of a recursion or iteration. We would like to make this precise, e.g. by giving a non-axiomatisability proof for all equivalences, finer than or equal to trace congruence, over $\mu$-expressions and hence over $*$-expressions.

Another interesting question about trace congruence is whether it is equationally axiomatisable relative to bisimulation. A negative result would immediately give non-axiomatisability. A positive result could be combined with the result of

[BÉ94], if $E3$ is sound for trace congruence, to give a pure horn clause axiomatisation.

**Bisimulation of $*$-expressions**




This expands the lower back left corner of the cube. The lines indicate signature inclusions. As discussed in §4.1.2 we would like to know whether bisimulation over the expressions of $\{1, a, +, \cdot, \_^\star\_\}$ is equationally axiomatisable and cannot use a simple adaption of the proof of Theorem 11, depending only on the widths of axioms, to show that it is not. One could also investigate whether the axiomatisation of weak congruence over $\{0, a \cdot \_, +, a^\star\_\}$ given in [AI95] can be adapted to $\{a, +, \cdot, \_^\star\_\}$, possibly using ideas from the proof of Theorem 13 that over $\mu$-expressions weak congruence is axiomatisable relative to bisimulation.

**Incompleteness of axiomatisations with unbounded width**

The proof of Theorem 6 is somewhat crude in that it ignores all information about putative axiomatisations except for the widths of the bisimulations that they provide. We conjectured in §2.3.4 that there are subsets of the axioms $C_{mn}$ of unbounded width that are incomplete even together with any finite set of other axioms. One might begin by generalising Proposition 92 below.

**Definition** For $n \geq 1$ an *n-lts* is a finite transition system over $n$ actions, each state having exactly one outgoing transition on each action (and no visibilities). For $k \geq 1$ we write $\sim^k$ for the transitive closure of the union of all bisimulations of width $\leq k$.

The 1-lts' are particularly simple, enabling us to give a precise characterisation of $\sim^k$.

**Proposition 91** *For any* 1*-lts* $S$ *there exist* $m \geq 0, n \geq 1$ *such that* $S \sim^1 S_{n,m} \stackrel{\text{def}}{=} a^{m+n-1}\mu X a^n X$.

**Proposition 92** *If* $m, m' \geq 0$, $n, n' \geq 1$ *and* $k \geq 2$ *then* $S_{n,m} \sim^k S_{n',m'}$ *iff for all primes* $p$*, if* $n$ *and* $n'$ *have different numbers of factors of* $p$ *then* $p \leq k$.

PROOF Straightforward, along the lines of the proof of Lemma 63. □

## 6.3 Unification

In Chapter 5 we showed that for unification problems over $\mu$-expressions (equivalently, over lambda terms containing only zero order variables) there exist finite computable complete sets of unifiers up to bisimulation. It remains, however, to characterise the minimal such sets.

It follows from their existence that the soundness of sequents over $\mu$-expressions is decidable. We would like to have a decidability result for sequents over lambda terms containing first order variables, particularly as these suffice to express the axiomatisation of Bloom and Ésik.

# Appendix A

# Completeness proof

This appendix contains the deferred parts of the proof of Theorem 1, i.e. completeness for the infinite system of §2.3. First we give a little technical result.

**Lemma 93** *If $z_j$ is not free in $\tilde{F}$ for any $j$ and $z_j$ is not free in $E$ for non-zero $j$ then*

$$\vdash \mu \langle X = E, \tilde{Y} = \tilde{F}, \tilde{Z} = \tilde{G} \rangle \quad = \quad \mu \langle X = E[\mu \langle \tilde{Z} = \tilde{G} \rangle / Z_0], \tilde{Y} = \tilde{F} \rangle .$$

PROOF We show the two sides are alpha equivalent by induction on the size of $\tilde{Z}$. The base case is trivial. For the inductive step:

$$\begin{array}{rcl} & \mu \langle X = E, \tilde{Y} = \tilde{F}, \tilde{Z} = \tilde{G}, Z = G \rangle & \\ = & \mu \langle X = E[\mu ZG / Z], \tilde{Y} = \tilde{F}[\mu ZG / Z], \tilde{Z} = \tilde{G}[\mu ZG / Z] \rangle & \text{by def } \mu \\ = & \mu \langle X = E, \tilde{Y} = \tilde{F}, \tilde{Z} = \tilde{G}[\mu ZG / Z] \rangle & \text{by premise} \\ = & \mu \langle X = E[\mu \langle \tilde{Z} = \tilde{G}[\mu ZG / Z] \rangle / Z_0], \tilde{Y} = \tilde{F} \rangle & \text{by ind. hyp.} \\ = & \mu \langle X = E[\mu \langle \tilde{Z} = \tilde{G}, Z = G \rangle / Z_0], \tilde{Y} = \tilde{F} \rangle & \text{by def } \mu. \end{array}$$

□

**Lemma 18** *For axioms $Q$, if $Q \vdash E_i = F_i$ for all $i$ then $Q \vdash \mu \langle \tilde{X} = \tilde{E} \rangle = \mu \langle \tilde{X} = \tilde{F} \rangle$.*

PROOF By induction in the length of $\tilde{X}$. The base case is just $\mu$-congruence. For the inductive step we need to know that if $Q \vdash E = F$ then $Q \vdash E[G/X] = F[G/X]$ which can be shown by induction on proofs, noting that none of the axioms constrain the forms of instantiations of metavariables (up to alpha conversion). □

**Lemma 19** *If* $\pi : n \to n$ *is a permutation with* $\pi(0) = 0$ *then* $B3 \vdash \mu \langle X_i = E_i \mid i \in n\rangle = \mu \langle X_{\pi(i)} = E_{\pi(i)} \mid i \in n\rangle$.

PROOF We show the result for permutations which simply exchange two values, which suffices. Here $\tilde{Y} = \tilde{F}$ may be the empty sequence.

$$\begin{array}{lll} & \mu \langle \tilde{X} = \tilde{E}, X = E, Y = F, \tilde{Y} = \tilde{F}\rangle & \\ = & \mu \langle \tilde{X} = \tilde{E}\rho, X = E\rho, Y = F\rho\rangle & \text{by def } \mu \text{ for some } \rho \text{ with domain } \tilde{Y} \\ = & \mu \langle \tilde{X} = \tilde{E}\rho \left[ \begin{array}{rcl} \mu Y F\rho[\mu X E\rho[\mu Y F\rho / Y] / X] & / & Y \\ \mu X E\rho[\mu Y F\rho / Y] & / & X \end{array} \right]\rangle & \text{by def } \mu \\ = & \mu \langle \tilde{X} = \tilde{E}\rho \left[ \begin{array}{rcl} \mu Y F\rho[\mu X E\rho / X] & / & Y \\ \mu X E\rho[\mu Y F\rho[\mu X E\rho / X] / Y] & / & X \end{array} \right]\rangle & \text{by } B3 \text{ and Lem. 18} \\ = & \mu \langle \tilde{X} = \tilde{E}, Y = F, X = E, \tilde{Y} = \tilde{F}\rangle & \text{by def } \mu \end{array}$$

□

**Lemma 20** *For any* $\mu$*-expression* $E$ *there is a standard accessible sequence* $\langle \tilde{X} = \tilde{E}\rangle$ *such that* $B \vdash E = \mu \langle \tilde{X} = \tilde{E}\rangle$.

PROOF This is immediate from Lemmas 94 and 96 following. □

**Definition** A sequence of equations $\langle \tilde{X} = \tilde{E}\rangle$ is *almost standard* if each $E_i$ is of the form

$$E ::= 0 \mid X \mid aX \mid E + E$$

(where $X$ ranges over all variables) and does not contain a free occurrence of $X_0$.

**Lemma 94** *For any expression* $E$ *there is an almost standard sequence* $\langle \tilde{X} = \tilde{E}\rangle$ *such that* $B \vdash E = \mu \langle \tilde{X} = \tilde{E}\rangle$.

PROOF It is straightforward to show that for any $E$ there is an expression $\mu Y F$ such that:

1. $B \vdash E = \mu Y F$.

2. $Y$ does not occur in $F$.

3. $F$ is of the form $F ::= 0 \mid X \mid a\mu X F \mid \mu X F \mid F + F$.

4. The bound variables of $\mu Y F$ are distinct from each other and from the free variables.

5. $\mu Y F$ has no subexpressions of the form $\mu Z Z + G$.

We define a sequence of equations $\mu^{-1}(\mu Y F)$ for any $\mu Y F$ that satisfies 3–5 as follows.

$$\mu^{-1}\left(\mu Y \sum_{0}^{l-1} a_i \mu X_i F_i + \sum_{l}^{l+m-1} \mu X_i F_i + \sum_{0}^{n} Y_k\right) \stackrel{\text{def}}{=} \left\langle \begin{array}{c} Y = \sum_{0}^{l-1} a_i X_i + \sum_{l}^{l+m-1} X_i + \sum_{0}^{n} Y_k \\ \mu^{-1}(\mu X_i F_i) \mid i \in l+m \end{array} \right\rangle$$

Conditions 3,5 ensure that it is possible to write $F$ in the form appearing in the left hand side. Condition 4 ensures that the result is a well defined equation sequence. Conditions 2,3 ensure that it is almost standard. It therefore suffices to show that $B \vdash \mu\mu^{-1}(\mu Y F) = \mu Y F$, which we do by induction on $F$. The inductive step requires the following.

**Lemma 95** *If:*

- *The sets of bound variables of $\mu X_i F_i$ are disjoint for distinct $i \in n$.*
- *The bound variables of $\mu X_i F_i$ and free variables of $\mu X_j F_j$ are disjoint for $i \neq j$.*
- *The bound variables of $\mu X_i F_i$ and free variables of $E$ are disjoint apart from $X_i$.*

*then*

$$B \vdash \mu \left\langle \begin{array}{c} Y = E \\ \mu^{-1}(\mu X_i F_i) \mid i \in n \end{array} \right\rangle = \mu Y E[\mu X_i F_i \,/X_i]_{i \in n}.$$

PROOF By induction on $n$. For $n = 0$ it is simply the definition of $\mu$. For

$n+1:$

$$
\begin{array}{lll}
& \mu \left\langle \begin{array}{c} Y = E \\ \mu^{-1}(\mu X_i F_i) \mid i \in n \\ \mu^{-1}(\mu X_n F_n) \end{array} \right\rangle & \\
= & \mu \left\langle \begin{array}{c} Y = E[\mu\mu^{-1}(\mu X_n F_n) \,/X_n] \\ \mu^{-1}(\mu X_i F_i) \mid i \in n \end{array} \right\rangle & \text{by Lemma 93} \\
= & \mu \left\langle \begin{array}{c} Y = E[\mu X_n F_n \,/X_n] \\ \mu^{-1}(\mu X_i F_i) \mid i \in n \end{array} \right\rangle & \text{by outer induction} \\
= & \mu Y E[\mu X_n F_n \,/X_n][\mu X_i F_i \,/X_i]_{i \in n} & \text{by inner induction} \\
= & \mu Y E[\mu X_i F_i \,/X_i]_{i \in n+1}. &
\end{array}
$$

□

□

**Lemma 96** *For any almost standard sequence $\langle \tilde{X} = \tilde{E} \rangle$ there is a standard accessible sequence $\langle \tilde{X} = \tilde{F} \rangle$ such that $B \vdash \mu \langle \tilde{X} = \tilde{E} \rangle = \mu \langle \tilde{X} = \tilde{F} \rangle$.*

PROOF Using Lemma 19, $B2$ and Lemma 18 we can ensure that no $E_i$ is of the form $X_i + G$. We proceed by induction on the size of

$$\{ X_i \mid X_i \text{ is a summand of some } E_j \}.$$

Using Lemma 19, if this set is not empty then we can assume it contains $X_{n-1}$. Let $\tilde{F}, F$ be $\tilde{E}$ with each occurrence of $X_{n-1}$ as a summand of some $E_j$ replaced by a new variable $W$. These occurrences of $X_{n-1}$ can be provably replaced with $E_{n-1}$ as follows (writing $\tilde{Y}, Y$ for $\tilde{X}$).

$$
\begin{array}{lll}
& \mu \langle \tilde{Y} = \tilde{F}[Y/W], Y = F \rangle & \\
= & \mu \langle \tilde{Y} = \tilde{F}[Y/W][\mu Y F \,/Y] \rangle & \text{by def } \mu \\
= & \mu \langle \tilde{Y} = \tilde{F} \left[ \begin{array}{c} \mu Y F \,/W \\ \mu Y F \,/Y \end{array} \right] \rangle & \\
= & \mu \langle \tilde{Y} = \tilde{F} \left[ \begin{array}{c} F[\mu Y F \,/W] \,/W \\ \mu Y F \,/Y \end{array} \right] \rangle & \text{by } B1 \text{ and Lemma 18} \\
= & \mu \langle \tilde{Y} = \tilde{F}[F/W][\mu Y F \,/Y] \rangle & \\
= & \mu \langle \tilde{Y} = \tilde{F}[F/W], Y = F \rangle & \text{by def } \mu.
\end{array}
$$

Establishing accessibility is straightforward. □

**Lemma 21** *If $\langle \tilde{X} = \tilde{E} \rangle$ is standard then $\langle \tilde{X} = \tilde{E} \rangle \sim \mu \langle \tilde{X} = \tilde{E} \rangle$.*

PROOF Straightforward. □

# Appendix B

# Loop properties

In this appendix we relate the loop structures of the transition systems $E\rho$ and $E\star\rho$, showing that loops in one can be matched 'on the nose' by loops in the other instead of merely 'up to bisimulation'. We show the following:

**Lemma 97** *If $\rho$ is a substitution for $E_r$ then* $\mathrm{loops}(E_r\rho) = \mathrm{loops}(E_r\star\rho)$.

It is then straightforward to show Lemma 62 of section 3.6. The interesting direction is the inclusion $\subseteq$. It does not follow from the bisimilarity of $E_r\rho$ and $E_r\star\rho$ as the action of $\rho$ may be non-injective — we may have

$$E_r\rho \longrightarrow^* E\rho \longrightarrow^l E'\rho$$

with $E\rho = E'\rho$ but $E \neq E'$. We show below that for any loop of $E_r\rho$ there is another with the same length headed by an instance of $\mathit{fix}$ and then that any such can be matched in $E_r\star\rho$.

**Notation** We write $S$ for the set of states of $E\star\rho$ and for $s \in S$ write $\mathrm{fst}(s)$ for the underlying term of $s$, i.e.

$$\begin{aligned} \mathrm{fst}(\langle E\rangle) &\stackrel{\text{def}}{=} E \\ \mathrm{fst}(\langle E, y\tilde{E}, d\rangle) &\stackrel{\text{def}}{=} E. \end{aligned}$$

We adopt the assumptions on $\rho$ from §3.5, letting $y$ range over $\mathrm{dom}(\rho)$, $z, Z$ over $\mathcal{Z}^P$ and $w$ over $\mathcal{K} - \mathrm{dom}(\rho) - FV(\mathrm{ran}(\rho)) - \mathcal{Z}$ (which we assume contains infinitely many variables at type $P$).

Intuitively if $s \in S$ and $\mathrm{fst}(s)\rho \triangleright \mathit{fix}\ M$ then this instance of $\mathit{fix}$ comes either from the term $\mathrm{fst}(s)$ or from the substitution $\rho$. To capture this we define new visibility predicates:

**Definition** Let $\rhd', \rhd'' \subseteq S \times T_1^P$ be the least relations satisfying the clauses below. The numbering of clauses matches that in the definition of $E{\star}\rho$ in §3.6.

1. $E \rhd F \Rightarrow \langle E\rangle \rhd' F$

2. n/a.

3. $E \rhd y\tilde{E} \wedge H_y \rhd z_j \wedge \langle E_j\rangle \rhd' F \Rightarrow \langle E\rangle \rhd' F$

4. n/a.

5. $E \rhd y\tilde{E} \wedge d : H_y \xrightarrow{l} H' \wedge H' \rhd z_j \wedge \langle E_j\rangle \rhd' F \Rightarrow \langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d\rangle \rhd' F$

1. n/a.

2. $E \rhd y\tilde{E} \wedge H_y \rhd H' \Rightarrow \langle E\rangle \rhd'' H'$

3. $E \rhd y\tilde{E} \wedge H_y \rhd z_j \wedge \langle E_j\rangle \rhd'' F \Rightarrow \langle E\rangle \rhd'' F$

4. $E \rhd y\tilde{E} \wedge d : H_y \xrightarrow{l} H' \wedge H' \rhd H'' \Rightarrow \langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d\rangle \rhd'' H''$

5. $E \rhd y\tilde{E} \wedge d : H_y \xrightarrow{l} H' \wedge H' \rhd z_j \wedge \langle E_j\rangle \rhd'' F \Rightarrow \langle H'[\tilde{E}/\tilde{z}], y\tilde{E}, d\rangle \rhd'' F$

These relations are complementary.

**Lemma 98** *If* $\mathrm{fst}(s)\rho \rhd \mathit{fix}\ \lambda w : P.\ A$ *then one of the following holds.*

1. $\exists E_0\ .\ s \rhd' \mathit{fix}\ \lambda w : P.\ E_0 \wedge E_0\rho = A$

2. $\exists H_0, \tilde{E}\ .\ s \rhd'' \mathit{fix}\ \lambda w : P.\ H_0 \wedge H_0[\tilde{E}\rho/\tilde{z}] = A$

PROOF First suppose $s = \langle E\rangle$ and show the result by induction on $E\rho \rhd \mathit{fix}\ \lambda w : P.\ A$ using Proposition 53. The result for $s$ not of this form is then a simple case analysis. □

Transitions of $E\rho$ can be matched in $E{\star}\rho$.

**Lemma 99** *If one of*

1. $s \stackrel{\text{def}}{=} \langle E[\mathit{fix}\ \lambda w : P.\ F/w]\rangle$ *and* $E\rho \xrightarrow{a} A$

2. $s \stackrel{\text{def}}{=} \langle H'[\tilde{E}[\mathit{fix}\ \lambda w : P.\ \ F\ /w]\ /\tilde{Z}], y\tilde{E}[\mathit{fix}\ \lambda w : P.\ \ F\ /w], d\rangle \in S$ *and* $H'[\tilde{E}/\tilde{Z}]\rho \xrightarrow{a} A$

*hold then there exists* $t \in S$ *such that* $s \xrightarrow{a} t$ *and either there exists* $E'$ *such that*

$$t = \langle E'[\mathit{fix}\ \lambda w : P.\ F\ /w]\rangle \text{ and } A = E'\rho$$

*or there exists* $y', \tilde{E}', H'', d'$ *such that*

$$t = \langle H''[\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\ /w]\ /\tilde{Z}'], y'\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\ /w], d'\rangle \text{ and } A = H''[\tilde{E}'/\tilde{Z}']\rho.$$

PROOF Suppose 1. By Proposition 53 there exists $E_1$ such that $E \xrightarrow{a}_\rho E_1$ and $E_1\rho = A$. The conclusion can be shown by induction on the definition of $\xrightarrow{a}_\rho$. Now suppose 2. By the assumptions on variables $H'[\tilde{E}/\tilde{Z}]\rho = H'[\tilde{E}\rho\ /\tilde{Z}]$ so $H'[\tilde{E}\rho\ /\tilde{Z}] \xrightarrow{a} A$ and either $H' \xrightarrow{a} H'' \wedge H''[\tilde{E}\rho\ /\tilde{Z}] = A$ or $H' \rhd Z_j \wedge E_j\rho \xrightarrow{a} A$. The first case is trivial, the second follows from part 1. □

Transitions of $E\rho$ that are headed by an instance of $\mathit{fix}$ can be matched in $E{\star}\rho$ by transitions which have a destination state that is somewhat independent of their start state.

**Lemma 100** *If* $\langle E\rangle \in S$, $E \rhd \mathit{fix}\ \lambda w : P.\ F$ *and* $F\rho \xrightarrow{a} A$ *then there exists* $t \in S$ *such that*

$$\forall \langle E''\rangle \in S\ .\ E'' \rhd \mathit{fix}\ \lambda w : P.\ F \Rightarrow \langle E''\rangle \xrightarrow{a} t$$

*and either there exists* $E'$ *such that*

$$t = \langle E'[\mathit{fix}\ \lambda w : P.\ F\ /w]\rangle \text{ and } A = E'\rho$$

*or there exists* $y', \tilde{E}', H'', d'$ *such that*

$$t = \langle H''[\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\ /w]\ /\tilde{Z}'], y'\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\ /w], d'\rangle \text{ and } A = H''[\tilde{E}'/\tilde{Z}']\rho.$$

PROOF By Proposition 53 there exists $F'$ such that $F \xrightarrow{a}_\rho F'$ and $F'\rho = A$. The result follows by considering cases of the definition of $\xrightarrow{a}_\rho$, using Lemma 99 for case 3. □

This can be strengthened to deal with all states of $E{\star}\rho$.

**Lemma 101** *If* $s \in S$, $s \rhd' \mathit{fix}\ \lambda w : P.\ F$ *and* $F\rho \xrightarrow{a} A$ *then there exists* $t \in S$ *such that*

$$\forall s' \in S\ .\ s' \rhd' \mathit{fix}\ \lambda w : P.\ F \Rightarrow s' \xrightarrow{a} t$$

*and either there exists $E'$ such that*

$$t = \langle E'[\mathit{fix}\ \lambda w : P.\ F\,/w]\rangle \text{ and } A = E'\rho$$

*or there exists $y', \tilde{E}', H'', d'$ such that*

$$t = \langle H''[\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\,/w]\,/\tilde{Z}'], y'\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\,/w], d'\rangle \text{ and } A = H''[\tilde{E}'/\tilde{Z}']\rho.$$

PROOF By an easy induction there exists $\langle E\rangle \in S$ such that $E \rhd \mathit{fix}\ \lambda w : P.\ F$. By Lemma 100 there exists $t \in S$ such that

$$\forall \langle E''\rangle \in S\ .\ E'' \rhd \mathit{fix}\ \lambda w : P.\ F \Rightarrow \langle E''\rangle \xrightarrow{a} t$$

and the second clause of the conclusion holds. It remains to check by induction on $\rhd'$ that

$$\forall s' \in S\ .\ s' \rhd' \mathit{fix}\ \lambda w : P.\ F \Rightarrow s' \xrightarrow{a} t.$$

□

**Lemma 102** *If one of*

1. $s \stackrel{\text{def}}{=} \langle E[\mathit{fix}\ \lambda w : P.\ F\,/w]\rangle$ *and* $E\rho \rhd w$

2. $s \stackrel{\text{def}}{=} \langle H'[\tilde{E}[\mathit{fix}\ \lambda w : P.\ F\,/w]\,/\tilde{Z}], y\tilde{E}[\mathit{fix}\ \lambda w : P.\ F\,/w], d\rangle \in S$ *and* $H'[\tilde{E}/\tilde{Z}]\rho \rhd w$

*hold then $s \rhd' \mathit{fix}\ \lambda w : P.\ F$.*

PROOF Suppose 1. By Proposition 53 $E \rhd_\rho w$. The conclusion can be shown by induction on the definition of $\rhd_\rho$. The result for 2 is then straightforward. □

Finally we can prove:

**Lemma 97** *If $\rho$ is a substitution for $E_r$ then* $\mathrm{loops}(E_r\rho) = \mathrm{loops}(E_r{\star}\rho)$.

PROOF $\supseteq$: Suppose $\langle E_r\rangle \longrightarrow^* s \longrightarrow^l s$. By tracking transitions along the higher order bisimulation in the proof of Lemma 60 we can see that

$$E_r\rho \longrightarrow^* \mathrm{fst}(s)\rho \longrightarrow^l \mathrm{fst}(s)\rho.$$

$\subseteq$: Suppose $E_r\rho \longrightarrow^* A \longrightarrow^l A$. By Lemma 66 there exists a term $\mathit{fix}\ \lambda w : P.\ B$, a finite set $I$ and $k_i \geq 1 \mid i \in I$ such that $E_r\rho \longrightarrow^* \rhd \mathit{fix}\ \lambda w : P.\ B$, $\sum_I k_i = l$ and $\forall i\ .\ B \longrightarrow^{k_i} \rhd w$.

By tracking transitions along the higher order bisimulation of Lemma 60 we have $\langle E_r\rangle \longrightarrow^* s$ for some $s$ with $\mathrm{fst}(s)\rho \rhd \mathit{fix}\ \lambda w : P.\ B$.

By Lemma 98 one of the following holds.

1. $s \rhd' \mathit{fix}\ \lambda w : P.\ F \wedge F\rho = B$.

   By induction on $n$ if $F\rho \longrightarrow^n A$ then there exist $t \in S$ such that

   $$\forall s' \in S\ .\ s' \rhd' \mathit{fix}\ \lambda w : P.\ F \Rightarrow s' \longrightarrow^n t$$

   and either there exists $E'$ such that

   $$t = \langle E'[\mathit{fix}\ \lambda w : P.\ F\,/w]\rangle \text{ and } A = E'\rho$$

   or there exists $y', \tilde{E}', H'', d'$ such that

   $$t = \langle H''[\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\,/w]\,/\tilde{Z}'], y'\tilde{E}'[\mathit{fix}\ \lambda w : P.\ F\,/w], d'\rangle$$

   and $A = H''[\tilde{E}'/\tilde{Z}']\rho$, using Lemma 101 for the base case and Lemma 99 for the inductive step.

   Finally, by Lemma 102, for all $i \in I$, if $s' \rhd' \mathit{fix}\ \lambda w : P.\ F$ then

   $$s' \longrightarrow^{k_i} \rhd' \mathit{fix}\ \lambda w : P.\ F$$

   so $s$ (and hence $\langle E_r\rangle$) has a loop of length $l$.

2. $s \rhd'' \mathit{fix}\ \lambda w : P.\ H_0 \wedge H_0[\tilde{E}\rho\,/\tilde{z}] = B$.

   We know that $w$ is not free in $\tilde{E}\rho$ so must have for each $i$ that

   $$H_0 \longrightarrow H_i \longrightarrow^{k_i - 1} \rhd w.$$

   By induction on $s \rhd'' \mathit{fix}\ \lambda w : P.\ H_0$ there is some $y\tilde{E}, d$ such that

   $$s \longrightarrow \langle H_i[\mathit{fix}\ \lambda w : P.\ H_0\,/w][\tilde{E}/\tilde{z}], y\tilde{E}, d\rangle.$$

   Further, this state has a loop of length $l$.

$\square$

# Bibliography


[Abr91] Samson Abramsky. A domain equation for bisimulation. *Information and Control*, 92,2:161–218, 1991.

[ABV92] Luca Aceto, Bard Bloom, and Frits Vaandrager. Turning SOS rules into equations. Technical Report CS-R9218, CWI, June 1992. To appear in the LICS '92 special issue of Information and Computation.

[Acz88] Peter Aczel. *Non-well-founded Sets*, volume 14 of *CSLI Lecture Notes*. CSLI, 1988.

[AG87] K. B. Arkhangelskii and P. V. Gorshkov. Implicational axioms for the algebra of regular languages. *Doklady Akad. Nauk, USSR, ser A.*, 10:67–69, 1987. (in Russian).

[AH88] L. Aceto and M. Hennessy. Termination, deadlock and divergence. Technical Report 6/88, Sussex University, 1988.

[AI95] L. Aceto and A. Ingólfsdóttir. A complete equational axiomatization for prefix iteration with silent steps. Research Report RS–95–5, BRICS (Basic Research in Computer Science, Centre of the Danish Research Foundation), Department of Mathematics and Computer Science, Aalborg University, January 1995.

[Avr91] Arnon Avron. Simple consequence relations. *Information and Computation*, 92:105–139, 1991.

[BB88] Franz Baader and Wolfram Büttner. Unification in commutative idempotent monoids. *Theoretical Computer Science*, 56:345–353, 1988.

[BBK87] J. C. M. Baeten, J. A. Bergstra, and J. W. Klop. On the consistency of Koomen's fair abstraction rule. *Theoretical Computer Science*, 51:129–176, 1987.

[BBP94] J. A. Bergstra, I. Bethke, and A. Ponse. Process algebra with iteration. *The Computer Journal*, 37(4):243–258, 1994. Also as University of Amsterdam Programming Research Group report P9314.

[BÉ93a] Stephen L. Bloom and Zoltán Ésik. Equational axioms for regular sets. *Math. Struct. in Comp. Science*, 3:1–24, 1993.

[BÉ93b] Stephen L. Bloom and Zoltán Ésik. *Iteration Theories: The Equational Logic of Iterative Processes*. EATCS Monographs on Theoretical Computer Science. Springer-Verlag, 1993.

[BÉ94] Stephen L. Bloom and Zoltán Ésik. Iteration algebras of finite state process behaviors. Draft, February 1994.

[BK84] J. A. Bergstra and J. W. Klop. A complete inference system for regular processes with silent moves. Technical Report CS-R8420, CWI, 1984.

[Bof90] M. Boffa. Une remarque sur les systèmes complets d'identités rationelles. *Theoret. Inform. Applic.*, 24(4):419–423, 1990.

[BS93] Franz Baader and Jörg H. Siekmann. *Handbook of Logic in Artificial Intelligence and Logic Programming*, chapter 'Unification Theory'. Oxford University Press, 1993.

[BW90] J. C. M. Baeten and W. P. Weijland. *Process Algebra*. Cambridge Tracts in Theoretical Computer Science 18. Cambridge University Press, 1990.

[CEW58] Irving M. Copi, Calvin C. Elgot, and Jesse B. Wright. Realization of events by logical nets. *Journal of the ACM*, 5(2):181–196, April 1958.

[CHM93] Søren Christensen, Yoram Hirshfeld, and Faron Moller. Decomposability, decidability and axiomatisability for bisimulation equivalence on basic parallel processes. In *Proc. 8th IEEE Symposium on Logic in Computer Science*, pages 386–396, New York, 1993. IEEE Computer Society Press.

[CHM94] Søren Christensen, Yoram Hirshfeld, and Faron Moller. Decidable subsets of CCS. *The Computer Journal*, 37(4):233–242, 1994.

[Chr93] Søren Christensen. *Decidability and Decomposition in Process Algebras*. PhD thesis, University of Edinburgh, 1993. Also as CST–105–93.

[Con71] J. H. Conway. *Regular Algebra and Finite Machines*. Chapman and Hall, 1971.

[CS90] Rance Cleveland and Bernhard Steffen. A preorder for partial process specifications. In *CONCUR 90*, pages 141–151, 1990.

[Dro94] N. J. Drost. Unification in an algebra with choice and action prefix. Technical Report P9321, Programming Research Group, University of Amsterdam, July 1994.

[Fok94] W. J. Fokkink. A complete equational axiomatisation for prefix iteration. *Information Processing Letters*, 52(6):333–337, December 1994. Also as CWI report CS-R9415.

[FZ94] Wan Fokkink and Hans Zantema. Basic process algebra with iteration: Completeness of its equational axioms. *The Computer Journal*, 37(4):259–267, 1994. Also as CWI report CS–R9368.

[Gla90] R. J. van Glabeek. The linear time — branching time spectrum. In J. C. M. Baeten and J. W. Klop, editors, *Proceedings CONCUR 90, Amsterdam, LNCS 458,* pages 278–297, 1990.

[Gla93a] R. J. van Glabeek. A complete axiomatization for branching bisimulation congruence of finite-state behaviours. In Andrzej M. Borzyszkowski and Stefan Sokolowski, editors, *Proceedings 18th MFCS, LNCS 711*, pages 473–484, 1993.

[Gla93b] R. J. van Glabeek. Divergence bisimulation. Personal communication, 1993.

[Hen81] Matthew Hennessy. A term model for synchronous processes. *Information and Control*, 51:58–75, 1981.

[Hen88] Matthew Hennessy. *Algebraic Theory of Processes*. The MIT Press, 1988.

[HM80] Matthew Hennessy and Robin Milner. On observing nondeterminism and concurrency. In J. W. de Bakker and J. van Leeuwen, editors, *Proceedings 7th Colloquium on Automata, Languages and Programming, LNCS 85*, pages 299–309. Springer-Verlag, 1980.

[HP80] M. C. B. Hennessy and G. D. Plotkin. A term model for CCS. In *Proceedings 9th MFCS, LNCS 88*, 1980.

[HS91] H. Hüttel and C. Stirling. Actions speak louder than words: proving bisimilarity for context-free processes. In *Proc. 6th IEEE Symposium on Logic in Computer Science*, pages 376–386, New York, 1991. IEEE Computer Society Press.

[Hüt91] H. Hüttel. *Decidability, Behavioural Equivalences and Infinite Transition Graphs*. PhD thesis, University of Edinburgh, 1991. CST–86–91.

[Kle56] S. C. Kleene. Representation of events in nerve nets and finite automata. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 3–41. Princeton University Press, 1956. Annals of Mathematics Studies 34.

[Koz94] Dexter Kozen. A completeness theorem for Kleene algebras and the algebra of regular events. *Information and Computation*, 110:366–390, 1994. Also in LICS '91.

[Kro91] Daniel Krob. Complete systems of B-rational identities. *Theoretical Computer Science*, 89:207–343, 1991.

[Liu93] Xinxin Liu. On decidability and small model property of process equations. Technical Report 93:10, School of Cognitive and Computing Sciences, University of Sussex, 1993.

[Mil80] Robin Milner. *A Calculus of Communicating Systems, LNCS 92*. Springer-Verlag, 1980. Also published as LFCS report ECS-LFCS-86-7.

[Mil81] Robin Milner. A modal characterisation of observable machine-behaviour. In *Proceedings CAAP '81, LNCS 112*, pages 25–34, 1981.

[Mil84] Robin Milner. A complete inference system for a class of regular behaviours. *Journal of Computer and System Sciences*, 28(3):439–466, 1984.

[Mil89] Robin Milner. A complete axiomatisation for observational congruence of finite state behaviours. *Information and Computation*, 81:227–247, 1989.

[Mit90] John C. Mitchell. Type systems for programming languages. In *The Handbook of Theoretical Computer Science*, chapter 8. Elsevier Science, 1990.

[Mol89] Faron Moller. *Axioms for Concurrency*. PhD thesis, University of Edinburgh, 1989.

[MP43] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.

[MT92] K. Meinke and J. V. Tucker. *Handbook of Logic in Computer Science*, volume 1, chapter 'Universal Algebra', pages 189–411. Clarendon press, Oxford, 1992.

[Nie89] Flemming Nielson. The typed $\lambda$-calculus with first-class processes. In *Proc. PARLE '89, LNCS 366*. Springer-Verlag, 1989.

[Par81] D. M. R. Park. Concurrency and automata on infinite sequences. In *Proc. 5th G.I. Conference, LNCS 104*. Springer-Verlag, 1981.

[PT87] R. Paige and R. Tarjan. Three partition refinement algorithms. *SIAM Journal on Computing*, 16(6):973–989, 1987.

[Rab93] Alexander Rabinovich. A complete axiomatisation for trace congruence of finite state behaviors. In Austin and Main, editors, *Proceedings of Mathematical Foundations of Programming Semantics (IX), LNCS*, 1993. (to appear).

[Red64] V. N. Redko. On defining relations for the algebra of regular events. *Ukrain. Mat. Zh.*, 16:120–126, 1964. (in Russian).

[Sal66] Arto Salomaa. Two complete axiom systems for the algebra of regular events. *Journal of the ACM*, 13(1):158–169, January 1966.

[San93] Davide Sangiorgi. *Expressing Mobility in Process Algebras: First-Order and Higher-Order Paradigms*. PhD thesis, University of Edinburgh, 1993.

[San94] Davide Sangiorgi. Bisimulation in higher-order process calculi. In *Proceedings of the IFIP Working Conference on Programming Concepts, Methods and Calculi (PROCOMET '94)*, pages 207–224, 1994.

[Sew94] Peter Sewell. Bisimulation is not finitely (first-order) equationally axiomatisable. In *Proc. 9th IEEE Symposium on Logic in Computer Science*, pages 62–70. IEEE, 1994.

[Sta82] R. Statman. Completeness, invariance and $\lambda$-definability. *The Journal of Symbolic Logic*, 47(1):17–26, March 1982.

[Tur37] A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society (second series)*, 42:230–265, 1937.

[Vra86] J. L. M. Vrancken. The algebra of communicating processes with empty process. Technical Report FVI 86-01, University of Amsterdam, Department of Computer Science, 1986.

[Wal88] D. J. Walker. Bisimulations and divergence. In *Proc. 3rd IEEE Symposium on Logic in Computer Science*, pages 186–192, 1988.

[Yan] Yanov. See [Con71, p. 108].

# Index of axioms

$$
\begin{array}{llcll}
F1 & \mu\left\langle \begin{array}{l} X = E + aY \\ Y = F + \tau G \end{array} \right\rangle & = & \mu\left\langle \begin{array}{l} X = E + aY + aG \\ Y = F + \tau G \end{array} \right\rangle & 75 \\
F2 & \mu\left\langle \begin{array}{l} X = E + \tau Y \\ Y = F + G \end{array} \right\rangle & = & \mu\left\langle \begin{array}{l} X = E + \tau Y + G \\ Y = F + G \end{array} \right\rangle & 75 \\
F3 & a\mu XE & = & a\mu XE + \tau X. & 75 \\
F4 & E \sim F & \rightarrow & E = F & 75
\end{array}
$$

$$
\begin{array}{llcll}
P1 & aE + aF & = & a(E + F) & 22 \\
P2 & a0 & = & 0 & 22
\end{array}
$$

commutative identity for $\mu$-expressions

$$
\begin{array}{rl}
& \tilde{\mu}_{i'} \left\langle X_i = E_i[X_k/Z_{kl}]_{k\in m, l\in n} \mid i \in m \right\rangle \\
= & \tilde{\mu}_{i'j'} \left\langle Y_{ij} = E_i[Y_{k,\rho_{ij}(l)}/Z_{kl}]_{k\in m, l\in n} \mid i \in m, j \in n \right\rangle
\end{array} \qquad 27
$$

commutative identity for $*$-expressions

$$
B^* \cdot \rho \quad = \quad \rho \cdot C^* \qquad 33
$$

functorial implication

$$
\begin{array}{rl}
& \forall i \in m \;.\; E_i[Y_{\rho(j)}/X_j]_{j\in m} = F_{\rho(i)} \\
\rightarrow & \forall i \in m \;.\; \tilde{\mu}_i \tilde{X}\ \tilde{E} = \tilde{\mu}_{\rho(i)} \tilde{Y}\ \tilde{F}
\end{array} \qquad 30
$$

group identity

$$
\left[\begin{array}{cccc} 1 & 0 & \cdots & 0 \end{array}\right] \cdot M^* \cdot \left[\begin{array}{c} 1 \\ \vdots \\ 1 \end{array}\right] \quad = \quad \left(\sum_{g\in G} X_g\right)^* \qquad 34
$$

# Index