Equivalence of Linear, Free, Liberal, Structured Program Schemas is Decidable in Polynomial Time.

Michael R.Laurence [d] Sebastian Danicic [a] Mark Harman [d]
Rob Hierons [b] John Howroyd [c]

[a] *Department of Computing Sciences, Goldsmiths College, University of London, New Cross, London SE14 6NW, UK.*

[b] *Department of Information Systems and Computing, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK.*

[c] *@UK PLC 5 Jupiter House Calleva Park Aldermaston Berks RG7 8NN*

[d] *Department of Computer Science, King's College London, Strand, London WC2R 2LS, United Kingdom*

Abstract

A program schema defines a class of programs, all of which have identical statement structure, but whose expressions may differ. We define a class of syntactic *similarity* binary relations between linear structured schemas and show that these relations characterise schema equivalence for structured schemas which are linear, free and liberal. In this paper we prove that similarity implies equivalence for linear schemas; the proof of a near-converse for schemas that are linear, free and liberal (LFL), which is much longer, is given in a Technical Report, which also contains the results of this paper. Our main result considerably extends the class of program schemas for which equivalence is known to be decidable, and suggests that linearity is a constraint worthy of further investigation.

Key words: structured program schemas, conservative schemas, liberal schemas, free schemas, linear schemas, schema equivalence, static analysis, program slicing

1 Introduction

A program schema represents the statement structure of a program by replacing real functions and predicates with function and predicate symbols taken from sets $\mathcal{F}$ and $\mathcal{P}$ respectively. A schema S thus defines a whole class $[S]$ of programs all of the same structure. Each program in $[S]$ can be obtained from S via a mapping called an *interpretation* which gives meanings to the function and predicate symbols in S. As an example, Figure 1 gives a schema S; and the program P of Figure 2 is in the class $[S]$.

The primary application of the theory of program schemas was as a framework for investigating program transformations; in particular those used by compilers during optimisation. If it could be proved that a certain transformation on schemas preserved equivalence, then this transformation could certainly be safely applied to programs. Surveys on the theory of program schemas can be found in the works of Greibach [1] and Manna [2].

$$\begin{array}{lll} u := h(); & & \\ \mathit{if}\ p(w) & \mathit{then} & v := f(u); \\ & \mathit{else} & v := g(); \end{array}$$

Fig. 1. Schema S

$$\begin{array}{lll} u := 1; & & \\ \mathit{if}\ w > 1 & \mathit{then} & v := u + 1; \\ & \mathit{else} & v := 2; \end{array}$$

Fig. 2. Program P

This paper gives a class of schemas for which equivalence is decidable. Equivalence is defined as follows. Given any variable v in a variable set $\mathcal{V}$, we say that schemas S, T are *v-equivalent*[1], written $S \cong_v T$, if given any interpretation and an initial state (that is, a mapping from the set of variables into some fixed domain) the programs defined by S and T give the same final value to the variable v, provided they both terminate. We also define $S \cong_\omega T$ to mean that given any interpretation and any initial state, the programs defined by S and T either both terminate or both fail to terminate. Thus the schema T of Figure 3 satisfies $S \cong_v T$, with S as in Figure 1; but $S \cong_\omega T$ does not

[1] For the class of *all* schemas the relation $\cong_v$ is not transitive, as an example in Section 3 shows, but it *is* an equivalence relation for the class of free, structured schemas in which we will be working (Proposition 20).

$$
\begin{array}{lll}
\textit{while}\, q(v) & \textit{do} & v := k(v); \\
\textit{if}\, p(w) & \textit{then} & \\
 & & \{ \\
 & & u := h(); \\
 & & v := f(u); \\
 & & \} \\
 & \textit{else} & v := g();
\end{array}
$$

Fig. 3. Schema T

hold. The relation $\cong_V$ for $V \subseteq \mathcal{V} \cup \{\omega\}$ means the conjunction of the relations $\cong_u$ for all $u \in V$. We write $\cong$ ('equivalence') to mean $\cong_{\mathcal{V}\cup\{\omega\}}$. Some researchers use the phrase 'functional equivalence' to refer to the relation $\cong_{\mathcal{V}\cup\{\omega\}}$ and 'weak equivalence' for $\cong_{\mathcal{V}}$.

This definition of equivalence takes no account of relations between the symbols, or requirements that a function or predicate symbol have a certain meaning, although definitions of equivalence for which interpretations are defined in this more restricted way have been considered [3–5].

Traditionally schemas were defined using a set of labelled statements or equivalently a flow diagram. All results proved in this paper only concern *structured* schemas,[2] in which *goto* statements are forbidden, and predicate symbols are only used to build if statements, of the form $\textit{if}\, q(\mathbf{u})\, \textit{then}\, T_1\, \textit{else}\, T_2$, or while statements, of the form $\textit{while}\, p(\mathbf{u})\, \textit{do}\, T$; where in both cases $\mathbf{u}$ is a finite tuple of variables.

It has been shown that it is decidable whether two structured schemas which are Conservative, Free and Linear are equivalent [6]. The main result of this paper is a strengthening of this result; that it can be decided in polynomial time whether two structured schemas which are Liberal, Free and Linear (abbreviated LFL in this paper), are equivalent. We also define the *slice* of a schema and prove that it can be decided whether an LFL schema S is u-equivalent to a given slice T, for any $u \in \mathcal{V} \cup \{\omega\}$.

The proof of the main theorem involves the definition of a binary relation $\textit{simil}_V$ on linear schemas for $V \subseteq \mathcal{V} \cup \{\omega\}$. We will prove in this paper that $S\, \textit{simil}_u\, T \Rightarrow S \cong_u T$ holds for linear schemas S, T and $u \in \mathcal{V} \cup \{\omega\}$. There is a near-converse; $S \cong_{\{v,\omega\}} T \Rightarrow S\, \textit{simil}_{\{v,\omega\}}\, T$ holds for every $v \in \mathcal{V}$ and LFL schemas S, T; the proof of this is given in the Technical Report [7, Theorem 148], on account of its length.

[2] Some authors, for example Manna [2] use the phrase *while* schema for what we call a *structured* schema (except that Manna allows statements like $\textit{while}\, \neg p(\mathbf{u})\, \textit{do}\, T$); in this paper a *while* schema means a structured schema consisting of a while loop (Definition 3).

Since it can be decided in polynomial time whether $S\, simil_u\, T$ holds (Theorem 35), our main theorem follows.

1.1 Organisation of the paper

In Section 2 we give some background to the theory of schemas. In Section 3 we give the basic schema definitions. We also give the formal definitions of free and liberal schemas, and prove that variable equivalence is in fact an equivalence relation for the class of free schemas. We also prove that it is decidable whether a schema is both free and liberal. We then define syntactic relations between the symbols in a linear schema which are required in the statement of the definition of similarity of linear schemas. We then give this definition, and prove that it is decidable in polynomial time whether two linear schemas are u-similar, given any $u \in \mathcal{V} \cup \{\omega\}$. In Section 4 we give the definition of the *slice* of a schema, given by deleting statements from a schema, and discuss conditions under which slicing preserves equivalence. In Section 5 we prove that u-similarity implies u-equivalence for any $u \in \mathcal{V} \cup \{\omega\}$. In Section 6 we give the main theorem and discuss further possibilities for research.

2 Background to schema theory

2.1 Different classes of schemas

Many subclasses of schemas have been defined:

Linear schemas (Definition 4) in which each function and predicate symbol occurs at most once. [3]

Conservative schemas, in which every assignment is of the form $v := f(v_1, \ldots, v_r)$ where $v \in \{v_1, \ldots, v_r\}$.

Free schemas, (Definition 18) where all paths are executable under some interpretation.

Liberal schemas (Definition 18) in which two assignments along any executable path can always be made to assign distinct values to their respective variables.

The last three of these classes were first introduced by Paterson [8]. Of these conditions, the first two can clearly be decided for the class of all schemas. Paterson [8] also proved, using a reduction from the Post Correspondence Problem, that it is not decidable whether a schema is free. He also showed however that it *is* decidable whether a schema is both liberal and free; and since he also gave an algorithm for

[3] Some authors use the phrase ‘non-repeating schemas’ to refer to what we call linear schemas.

transforming a schema S into a schema T such that T is both liberal and free if and only if S is liberal, it is clearly decidable whether a schema is liberal. It is an open problem whether freeness is decidable for the class of linear schemas.

All results on the decidability of equivalence of schemas are either negative or confined to very restrictive classes of schemas. In particular Paterson [8] proved, in effect, that equivalence is undecidable for the class of all (unstructured) schemas. He proved this by showing that the halting problem for Turing machines (which is, of course, undecidable) is reducible to the equivalence problem for the class of all schemas. Ashcroft and Manna showed [9] that an arbitrary schema can be effectively transformed into an equivalent structured schema, provided that statements such as *while* $\neg p(\mathbf{u})$ *do* T are permitted; hence Paterson's result shows that any class of schemas for which equivalence can be decided must not contain this class of schemas. Thus in order to get positive results on this problem, it is clearly necessary to define the relevant classes of schema with great care.

Although the class of linear structured schemas considered in this paper is a highly restrictive one, it has the merit that schemas in this class are the main objects studied in the field of Program Slicing (see Section 4), and that this is therefore a particularly important class.

2.2 Positive results on the decidability of schema equivalence

Besides the result of [6] mentioned above, positive results on the decidability of equivalence of schemas include the following; in an early result in schema theory, Ianov [10] introduced a restrictive class of schemas, the Ianov schemas, for which equivalence is decidable. Ianov schemas are monadic (that is, they contain only a *single* variable) and all function symbols are unary; hence Ianov schemas are conservative.

Paterson [8] proved that equivalence is decidable for a class of schemas called *progressive schemas*, in which every assignment references the variable assigned by the previous assignment along every legal path.

Sabelfeld [11] proved that equivalence is decidable for another class of schemas called *through schemas*. A through schema satisfies two conditions: firstly, that on every path from an accessible predicate p to a predicate q which does not pass through another predicate, and every variable x referenced by p, there is a variable referenced by q which defines a term containing the term defined by x, and secondly, distinct variables referenced by a predicate define distinct terms under any Herbrand interpretation (Definition 9).

Our interest in the theory of program schemas is motivated in part by applications in program slicing. Slicing has many applications including program comprehension [12], software maintenance [13], [14], [15], [16], debugging [17], [18], [19], [20], testing [21], [22], [23], re-engineering [24], [25], component reuse [26], [27], program integration [28], and software metrics [29], [30], [31]. There are several surveys of slicing techniques, applications and variations [32], [33], [34]. All applications of slicing rely on the fact that a slice is faithful to a projection of the original program's semantics, yet it is typically a smaller program.

The field of (static) program slicing is largely concerned with the design of algorithms which given a program and a variable v, eliminate as much code as possible from the program, such that the program (slice) consisting of the remaining code, when executed from the same initial state, will still give the same final value for v as the original program, and preserve termination. One algorithm is thus better than another if it constructs a smaller slice.

Slicing algorithms do not normally take account of the meanings of the functions and predicates occurring in a program, nor do they 'know' when the same function or predicate occurs in more than one place in a program. In effect, therefore, they work with a linear schema defined by the program, and the semantic properties which slices of programs are required to preserve are defined in terms of schema semantics. This motivates the study of schemas, which represent large classes of programs.

Weiser [35] showed that given a program and a variable v, there was a particular set of functions and predicates (corresponding to our set $\mathcal{N}_S(v)$ for schemas in Definition 31) which may affect the final value of v; the symbols not lying in this set may simply be deleted without affecting the final value of v. In Theorem 37 we generalise this by considering ω-equivalence as a slicing criterion. In [36] it was shown that if S is LFL then none of the symbols in $\mathcal{N}_S(u)$ (for $u \in \mathcal{V} \cup \{\omega\}$) can be deleted from S without giving a u-inequivalent schema. This is however false for the class of schemas which are merely linear and free; a counterexample is given in Figure 6 in Section 6.1.

3 Basic definitions

Definition 1 (symbol sets) Throughout this paper, $\mathcal{F}$, $\mathcal{P}$ and $\mathcal{V}$ denote fixed infinite sets of *function symbols*, of *predicate symbols* and of *variables* respectively. We assume a function

$$\mathit{arity} : \mathcal{F} \cup \mathcal{P} \to \mathbb{N}.$$

The arity of a symbol x is the number of arguments referenced by x. We assume that for each $n \in \mathbb{N}$ there are infinitely many elements of $\mathcal{F}$ and $\mathcal{P}$ of arity n, so we never run out of symbols of any required arity. Note that in the case when the arity of a function symbol g is zero, g may be thought of as a constant.

Definition 2 (terms) The set $Term(\mathcal{F}, \mathcal{V})$ of *terms* is defined as follows:

- each variable is a term,
- if $f \in \mathcal{F}$ is of arity n and $t_1, \ldots, t_n$ are terms then $f(t_1, \ldots, t_n)$ is a term.

If each term t_i is a variable, then $f(t_1, \ldots, t_n)$ is called a function expression.
We refer to a tuple $\mathbf{t} = (t_1, \ldots, t_n)$, where each t_i is a term, as a vector term. We call $p(\mathbf{t})$ a predicate term if $p \in \mathcal{P}$ and the number of components of the vector term $\mathbf{t}$ is $arity(p)$. If each component of $\mathbf{t}$ is a variable, then $p(\mathbf{t})$ is called a predicate expression.

Definition 3 (structured schemas) We define the set $Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$ of all *structured schemas* recursively as follows. The empty schema $\Lambda \in Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$. An assignment $y := f(\mathbf{x})$; where $y \in \mathcal{V}$, and $f(\mathbf{x})$ is a function expression, lies in $Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$. From these all schemas in $Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$ may be 'built up' from the following constructs on schemas.

- Sequences; $S' = U_1 U_2 \ldots U_r \in Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$ provided that each schema
$$U_1, \ldots, U_r \in Sch(\mathcal{F}, \mathcal{P}, \mathcal{V}).$$
We define $S\Lambda = \Lambda S = S$ for all schemas S.
- *If* schemas; $S'' = \mathit{if}\ p(\mathbf{x})\ \mathit{then}\ \{T_1\}\ \mathit{else}\ \{T_2\}$ lies in $Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$ whenever $p(\mathbf{x})$ is a predicate expression and $T_1, T_2 \in Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$.
- *While* schemas; $S''' = \mathit{while}\ q(\mathbf{y})\ \mathit{do}\ \{T\}$ lies in $Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$ whenever $q(\mathbf{y})$ is a predicate expression and T is a schema.

The predicate symbols p and q are called the *guards* of the schemas S'' and S''', respectively.
Finally, $|S|$ will denote the total number of function and predicate symbols in S, with n distinct occurrences of the same symbol counting n times.

Thus a schema is a word in a language over an infinite alphabet, for which Λ is the empty word. We normally omit the braces { and } if this causes no ambiguity. Also, we may write $\mathit{if}\ p(\mathbf{x})\ \mathit{then}\ \{T_1\}$ instead of
$\mathit{if}\ p(\mathbf{x})\ \mathit{then}\ \{T_1\}\ \mathit{else}\ \{T_2\}$ if $T_2 = \Lambda$.

Observe that $f(\mathbf{x})$ and $p(\mathbf{x})$ in Definition 3 are always function and predicate *expressions*; that is, the components of the vector term $\mathbf{x}$ are variables.

For the remainder of this paper, the word 'schema' is intended to mean 'structured schema'.

The sets of *if* and *while* predicate symbols occurring in a schema S are denoted by $\mathit{ifPreds}(S)$ and $\mathit{whilePreds}(S)$; their union is $\mathit{Preds}(S)$. We define $\mathit{Funcs}(S) \subseteq \mathcal{F}$ to be the set of function symbols in S and define $Symbols(S) = Funcs(S) \cup Preds(S)$. A schema without predicates is called *predicate-free*; a schema without while predicates is called *while-free*.

Definition 4 (linear schemas) If no element of $\mathcal{F} \cup \mathcal{P}$ appears more than once in a schema S, then S is said to be *linear*. If a linear schema S contains an assignment $y := f(\mathbf{x})$; then we define $assign_S(f) = y$, $\mathbf{refvec}_S(f) = \mathbf{x}$ and the set of components of $\mathbf{x}$ is $Refset_S(f) \subseteq \mathcal{V}$. If $p \in Preds(S)$ then $\mathbf{refvec}_S(p)$ and $Refset_S(p)$ are defined similarly.

3.1 Paths through a schema

The execution of a program defines a (possibly infinite) sequence of assignments and predicates. Each such sequence will correspond to a *path* through the associated schema. The set $\Pi^\omega(S)$ of paths through S is now given.

Definition 5 (the set $alphabet(S)$ and the set $\Pi^\omega(S)$ of paths through S) If σ is a word, or a set of words over an alphabet, then $pre(\sigma)$ is the set of all prefixes of (elements of) σ. If L is any set, then we write L^* for the set of finite words over L and L^ω for the set containing both finite and infinite words over L, and we write Λ to refer to the empty word; recall that Λ is also a particular schema.
For each schema S the alphabet of S, written $alphabet(S)$ is defined by

$$alphabet(S) = A \cup B$$

where

$$A = \{y := f(\mathbf{x}) \mid \; y := f(\mathbf{x}); \text{ is an assignment in } S\},$$
$$B = \{<\!p(\mathbf{x}) = Z\!> \mid \; p(\mathbf{x}) \text{ is a predicate expression in } S,\ Z \in \{\mathsf{T}, \mathsf{F}\}\}.$$

For any letter $l \in alphabet(S)$, we define $symbol(l) \in Symbols(S)$ to be f if l is an assignment with function symbol f, and p if l is $<\!p(\mathbf{x}) = Z\!>$ for $Z \in \{\mathsf{T}, \mathsf{F}\}$. The words in $\Pi(S) \subseteq (alphabet(S))^*$ are formed by concatenation from the words of subschemas as follows:

For Λ,

$$\Pi(\Lambda) = \{\Lambda\}.$$

For assignments,

$$\Pi(y := f(\mathbf{x});) = \{y := f(\mathbf{x})\}.$$

For sequences, $\Pi(S_1 S_2 \ldots S_r) = \Pi(S_1) \ldots \Pi(S_r)$.

For *if* schemas, $\Pi($ *if* $p(\mathbf{x})$ *then* $\{T_1\}$ *else* $\{T_2\})$ is the set of all concatenations of $\lessdot p(\mathbf{x}) = \mathsf{T} \gtrdot$ with a word in $\Pi(T_1)$ and all concatenations of $\lessdot p(\mathbf{x}) = \mathsf{F} \gtrdot$ with a word in $\Pi(T_2)$.

For *while* schemas, $\Pi($ *while* $q(\mathbf{y})$ *do* $\{T\})$ is the set of all words of the form

$$[\lessdot q(\mathbf{y}) = \mathsf{T} \gtrdot \, \Pi(T)]^* \lessdot q(\mathbf{y}) = \mathsf{F} \gtrdot$$

where $[< q(\mathbf{y}) = \mathsf{T} > \Pi(T)]^*$ denotes a finite sequence of words which are the concatenation of $\lessdot q(\mathbf{y}) = \mathsf{T} \gtrdot$ with a word from $\Pi(T)$.

We define the set $\Pi^\omega(S)$ of *paths* through S as

$$\Pi^\omega(S) = \Pi(S) \cup \{\sigma \in (alphabet(S))^\omega - (alphabet(S))^* |\, pre(\sigma) - \{\sigma\} \subseteq pre(\Pi(S))\}.$$

When referring to a linear schema S, we will sometimes omit the reference to $\mathbf{refvec}_S(p)$ for $p \in Preds(S)$ when denoting elements of $alphabet(S)$; that is, we will write $\lessdot p = Z \gtrdot$ to refer to $\lessdot p(\mathbf{x}) = Z \gtrdot$. Since the schema S is linear, this is unambiguous.

Lemma 6 *Let S be a schema.*

(1) *If $\sigma \in pre(\Pi(S))$, the set $\{l \in alphabet(S) |\, \sigma l \in pre(\Pi(S))\}$ is one of the following; the empty set, a singleton containing an assignment, or a pair $\{\lessdot p(\mathbf{x}) = \mathsf{T} \gtrdot, \lessdot p(\mathbf{x}) = \mathsf{F} \gtrdot\}$ where $p \in Preds(S)$.*
(2) *An element of $\Pi(S)$ cannot be a strict prefix of another.*

Proof. Both assertions follow by induction on $|S|$. $\square$

Lemma 6 reflects the fact that at any point in the execution of a program, there is never more than one 'next step' which may be taken.

Definition 7 (paths *passing through* a symbol) We say that a path $\sigma \in \Pi^\omega(S)$ *passes* through a function symbol f (or a predicate p) if it contains an assignment with function symbol f (or $\lessdot p(\mathbf{x}) = Z \gtrdot$ for $Z \in \{\mathsf{T}, \mathsf{F}\}$). We may strengthen this by saying that σ passes through an element $l \in alphabet(S)$ if l occurs in σ.

Definition 8 (segments of a schema and of segments) Let S be a schema and let $\mu \in alphabet(S)^*$. We say that μ is a segment (in S) if there are words μ_1, μ_2 such that $\mu_1 \mu \mu_2 \in \Pi(S)$. If μ, σ are segments in S, then we say that μ is a segment of σ in S if we can write $\sigma = \mu_1 \mu \mu_2$.
We say that a segment μ starts (ends) at $x \in Symbols(S)$ if $\tilde{x} \in alphabet(S)$ is the first (last) letter of μ, with $x = symbol(\tilde{x})$.

The symbols upon which schemas are built are given meaning by defining the notions of a state and of an interpretation. It will be assumed that 'values' are given in a single set D, which will be called the *domain*. We are mainly interested in the case in which $D = \mathit{Term}(\mathcal{F}, \mathcal{V})$ (the Herbrand domain) and the function symbols represent the 'natural' functions with respect to $\mathit{Term}(\mathcal{F}, \mathcal{V})$.

Definition 9 (states, (Herbrand) interpretations and the natural state e) Given a domain D, a *state* is either $\perp$ (in the case of non-terminating programs) or a function $\mathcal{V} \to D$. The set of all such states will be denoted by $\mathrm{State}(\mathcal{V}, D)$. An interpretation i defines, for each function symbol $f \in \mathcal{F}$ of arity n, a function $f^i : D^n \to D$, and for each predicate symbol $p \in \mathcal{P}$ of arity m, a function $p^i : D^m \to \{\mathsf{T}, \mathsf{F}\}$. The set of all interpretations with domain D will be denoted $\mathit{Int}(\mathcal{F}, \mathcal{P}, D)$. When the domain used is $\mathit{Term}(\mathcal{F}, \mathcal{V})$, an interpretation i is said to be *Herbrand* if the functions $f^i : \mathit{Term}(\mathcal{F}, \mathcal{V}) \to \mathit{Term}(\mathcal{F}, \mathcal{V})$ for each $f \in \mathcal{F}$ are defined as

$$f^i(t_1, \ldots, t_n) = f(t_1, \ldots, t_n)$$

for all n-tuples of terms $(t_1, \ldots, t_n)$.
In the case when the domain is $\mathit{Term}(\mathcal{F}, \mathcal{V})$, the *natural state* $e : \mathcal{V} \to \mathit{Term}(\mathcal{F}, \mathcal{V})$ is defined by $e(v) = v$ for all $v \in \mathcal{V}$.

Note that an interpretation i being Herbrand places no restriction on the mappings $p^i : (\mathit{Term}(\mathcal{F}, \mathcal{V}))^m \to \{\mathsf{T}, \mathsf{F}\}$ defined by i for each $p \in \mathcal{P}$.
It is well known [2, Section 4-14] that Herbrand interpretations, on the domain of terms, are the only ones that need to be considered when considering equivalence of schemas. This fact is stated more precisely in Theorem 17.

A program is obtained from a schema S and an interpretation i by replacing all symbols $f \in \mathcal{F}$ and $p \in \mathcal{P}$ in S by f^i and p^i; and given an initial state $d \in \mathrm{State}(\mathcal{V}, D)$, this program defines a final state

$$\mathcal{M}[\![S]\!]^i_d \in \mathrm{State}(\mathcal{V}, D)$$

in the obvious way, which will be given formally in Definition 13. (If the program fails to terminate for an initial state d, or if $d = \perp$, then we define $\mathcal{M}[\![S]\!]^i_d = \perp$.)

Given a schema $S \in \mathit{Sch}(\mathcal{F}, \mathcal{P}, \mathcal{V})$ and a domain D, an initial state $d \in \mathrm{State}(\mathcal{V}, D)$ with $d \neq \perp$ and an interpretation $i \in \mathit{Int}(\mathcal{F}, \mathcal{P}, D)$ we now define the final state $\mathcal{M}[\![S]\!]^i_d \in \mathrm{State}(\mathcal{V}, D)$ and the associated path $\pi_S(i, d) \in \Pi^\omega(S)$.

Definition 10 (the schema $\mathit{schema}(\sigma)$) Given a word $\sigma \in (\mathit{alphabet}(S))^*$, the predicate-free schema $\mathit{schema}(\sigma)$ consists of all the assignments along σ in the same order as in σ; and $\mathit{schema}(\sigma) = \Lambda$ if σ has no assignments.

Definition 11 (semantics of predicate-free schemas) Given a state $d \neq \perp$, the final state $\mathcal{M}[\![S]\!]_d^i$ and associated path $\pi_S(i,d) \in \Pi^\omega(S)$ of a schema S are defined as follows:

For Λ,

$$\mathcal{M}[\![\Lambda]\!]_d^i = d$$

and

$$\pi_\Lambda(i,d) = \Lambda.$$

For assignments,

$$\mathcal{M}[\![y := f(\mathbf{x});]\!]_d^i(v) \quad = \quad \begin{cases} d(v) & \text{if } v \neq y, \\ f^i(d(\mathbf{x})) & \text{if } v = y \end{cases}$$

(where $d(x_1, \ldots, x_r)$ is defined to be the tuple $(d(x_1), \ldots, d(x_r)))$

and

$$\pi_{y := f(\mathbf{x});}(i,d) \quad = \quad y := f(\mathbf{x}),$$

and for sequences $S_1 S_2$ of predicate-free schemas,

$$\mathcal{M}[\![S_1 S_2]\!]_d^i \quad = \quad \mathcal{M}[\![S_2]\!]^i_{\mathcal{M}[\![S_1]\!]_d^i}$$

and

$$\pi_{S_1 S_2}(i,d) \quad = \quad \pi_{S_1}(i,d)\pi_{S_2}(i, \mathcal{M}[\![S_1]\!]_d^i).$$

This uniquely defines $\mathcal{M}[\![S]\!]_d^i$ and $\pi_S(i,d)$ if S is predicate-free.

In order to give the semantics of a general schema S, first the path, $\pi_S(i,d)$, of S with respect to interpretation, i, and initial state d is defined.

Definition 12 (the path $\pi_S(i,d)$) Given a schema S, an interpretation i, and a state, $d \neq \perp$, the path $\pi_S(i,d) \in \Pi^\omega(S)$ is defined by the following condition; for all $\sigma \lessdot p(\mathbf{x}) = X \gtrdot \in \mathit{pre}(\pi_S(i,d))$, the equality $p^i(\mathcal{M}[\![\mathit{schema}(\sigma)]\!]_d^i(\mathbf{x})) = X$ holds.

In other words, the path $\pi_S(i,d)$ has the following property; if a predicate expression $p(\mathbf{x})$ along $\pi_S(i,d)$ is evaluated with respect to the predicate-free schema consisting of the sequence of assignments preceding that predicate in $\pi_S(i,d)$, then the value of the resulting predicate term given by i 'agrees' with the value given in $\pi_S(i,d)$.

By Lemma 6, this defines the path $\pi_S(i,d) \in \Pi^\omega(S)$ uniquely.

Definition 13 (semantics of arbitrary schemas) If $\pi_S(i,d)$ is finite, we define

$$\mathcal{M}[\![S]\!]_d^i = \mathcal{M}[\![\mathit{schema}(\pi_S(i,d))]\!]_d^i$$

(which is already defined, since $\mathit{schema}(\pi_S(i,d))$ is predicate-free) otherwise $\pi_S(i,d)$ is infinite and we define $\mathcal{M}[\![S]\!]_d^i = \perp$. In this last case we may say that $\mathcal{M}[\![S]\!]_d^i$ is not terminating. For convenience, if S is predicate-free and $d : \mathcal{V} \to \mathit{Term}(\mathcal{F}, \mathcal{V})$ is a state then we define unambiguously $\mathcal{M}[\![S]\!]_d = \mathcal{M}[\![S]\!]_d^i$. Also, for schemas S, T and interpretations i and j we write $\mathcal{M}[\![S]\!]_d^i(\omega) = \mathcal{M}[\![T]\!]_d^j(\omega)$ to mean $\mathcal{M}[\![S]\!]_d^i = \bot \iff \mathcal{M}[\![T]\!]_d^j = \bot$.

Observe that $\mathcal{M}[\![S_1 S_2]\!]_d^i = \mathcal{M}[\![S_2]\!]^i_{\mathcal{M}[\![S_1]\!]_d^i}$ and

$$\pi_{S_1 S_2}(i, d) = \pi_{S_1}(i, d)\pi_{S_2}(i, \mathcal{M}[\![S_1]\!]_d^i)$$

hold for all schemas (not just predicate-free ones).

Definition 14 (termination from the natural state e) If $\mathcal{M}[\![S]\!]_e^i \neq \bot$, then we say that i is a *terminating* interpretation for S.

Definition 15 (changing an interpretation) Given an interpretation i and $X \in \{\mathsf{T}, \mathsf{F}\}$ and $p \in \mathcal{P}$, the Herbrand interpretation $i(p = X)$ is given by

$$q^{i(p=X)}(\mathbf{t}) = \begin{cases} q^i(\mathbf{t}) & q \neq p \\ X & q = p \end{cases}$$

for every vector of terms $\mathbf{t}$ of the appropriate length. We generalise this by defining the interpretation $i(p(\mathbf{t}) = X)$ to mean the interpretation j satisfying $p^j(\mathbf{t}) = X$ and $q^j(\mathbf{t}') = q^i(\mathbf{t}')$ for all predicate terms $q(\mathbf{t}') \neq p(\mathbf{t})$.

Definition 16 (u-equivalence of schemas) Given any $u \in \mathcal{V} \cup \{\omega\}$, we say that schemas $S, T \in \mathit{Sch}(\mathcal{F}, \mathcal{P}, \mathcal{V})$ are *u-equivalent,* written $S \cong_u T$, if for every domain D and state $d : \mathcal{V} \to D$ and every $i \in \mathit{Int}(\mathcal{F}, \mathcal{P}, D)$, the following holds; either $u \in \mathcal{V} \wedge \bot \in \{\mathcal{M}[\![S]\!]_d^i, \mathcal{M}[\![T]\!]_d^i\}$, or

$$\mathcal{M}[\![S]\!]_d^i(u) = \mathcal{M}[\![T]\!]_d^i(u).$$

If $V \subseteq \mathcal{V} \cup \{\omega\}$, we write $S \cong_V T$ to mean $S \cong_u T \;\; \forall u \in V$ and we write $S \cong T$ to mean $S \cong_{\mathcal{V} \cup \{\omega\}} T$.

The following theorem, which is a restatement of [2, Theorem 4-1], ensures that we only need to consider Herbrand interpretations.

Theorem 17 *Let S be a schema and let D be a domain.*

- *For any state $d : \mathcal{V} \to D$ and any interpretation $i \in \mathit{Int}(\mathcal{F}, \mathcal{P}, D)$, there exists a Herbrand interpretation j such that $\pi_S(j, e) = \pi_S(i, d)$.*
- *If T is a schema, $u \in \mathcal{V} \cup \{\omega\}$ and for all Herbrand interpretations j we have $\mathcal{M}[\![S]\!]_e^j(u) = \mathcal{M}[\![T]\!]_e^j(u)$ (provided both sides terminate if $u \in \mathcal{V}$), then $S \cong_u T$ holds.*

Throughout the remainder of the paper, all interpretations will be assumed to be Herbrand.

Definition 18 Let $S \in \mathit{Sch}(\mathcal{F}, \mathcal{P}, \mathcal{V})$.

- If for every $\sigma \in \mathit{pre}(\Pi(S))$ there is a Herbrand interpretation i such that $\sigma \in \mathit{pre}(\pi_S(i, e))$, then S is said to be *free*.
- If for every prefix $\sigma = \mu < y := f(\mathbf{a}) > \nu < z := g(\mathbf{b}) > \in \mathit{pre}(\Pi(S))$ such that there is a Herbrand interpretation i such that $\sigma \in \mathit{pre}(\pi_S(i, e))$, we have

$$\mathcal{M}[\![\mathit{schema}(\mu)]\!]_e(f(\mathbf{a})) \neq \mathcal{M}[\![\mathit{schema}(\mu < y := f(\mathbf{a}) > \nu)]\!]_e(g(\mathbf{b})),$$

  then S is said to be *liberal*. (If $f \neq g$ then of course this condition is trivially satisfied.)

Thus a schema S is said to be free if for every path through S, there is a Herbrand interpretation which follows it with the natural state e as the initial state, and a schema S is said to be liberal if given any path through S passing through two assignments and a Herbrand interpretation which follows it with e as the initial state, the assignments give distinct values to the variables to which they assign.

Observe that if a schema S is free, and

$$\mu < p(\mathbf{x}) = X > \mu' < p(\mathbf{y}) = Y > \in \mathit{pre}(\pi_S(i, e))$$

for some Herbrand interpretation i, then

$$\mathcal{M}[\![\mathit{schema}(\mu)]\!]_e(\mathbf{x}) \neq \mathcal{M}[\![\mathit{schema}(\mu\mu')]\!]_e(\mathbf{y})$$

holds, since otherwise there would be no Herbrand interpretation whose path (for e) has the prefix $\mu < p(\mathbf{x}) = X > \mu' < p(\mathbf{y}) = \neg X >$. Thus a path through a free schema cannot pass twice (for initial state e) through the same predicate term.

As mentioned in the introduction, it was proved in [8] that it is decidable whether a schema is liberal, or liberal and free. Theorem 19 gives the essential result for linear schemas.

Theorem 19 (syntactic condition for being liberal and free)
Let S be a linear schema. Then S is both liberal and free if and only if for every segment $\tilde{x}\mu\tilde{y}$ in S with $\tilde{x}, \tilde{y} \in \mathit{alphabet}(S)$, $\mathit{symbol}(\tilde{x}) = \mathit{symbol}(\tilde{y})$ and such that the same symbol does not occur more than once in $\tilde{x}\mu$ or $\mu\tilde{y}$, then the segment $\tilde{x}\mu$ contains an assignment to a variable referenced by $\tilde{y}$.
In particular, it is decidable whether a linear schema is both liberal and free.

Proof [8]. Assume that S is both liberal and free. Then for any segment $\tilde{x}\mu\tilde{y}$ satisfying the conditions given, there is a prefix Θ and an interpretation i such that $\Theta\tilde{x}\mu\tilde{y} \in \mathit{pre}(\pi_S(i, e))$, and distinct (predicate) terms are defined when $\tilde{x}$ and $\tilde{y}$ are reached,

thus proving the condition.
To prove sufficiency, first observe that the 'non-repeating' condition on the letters of the segments $\tilde{x}\mu$ and $\mu\tilde{y}$ may be ignored, since segments that begin and end with letters having the same symbol can be removed from within $\tilde{x}\mu$ or $\mu\tilde{y}$ until it is satisfied. Consider the set of prefixes of $\Pi(S)$ of the form $\Theta\tilde{x}\mu\tilde{y}$ with $symbol(\tilde{x}) = symbol(\tilde{y})$ such that $\tilde{x}\mu\tilde{y}$ satisfies the condition given. By induction on the length of such prefixes, it can be shown that every assignment encountered along such a prefix defines a different term (for initial state e), and the result follows immediately from this.
Since there are finitely many segments in S which contain no repeated symbols except at the endpoints, and these can be enumerated, the decidability of liberality and freeness for the set of linear schemas follows easily. $\Box$

Theorem 19 can easily be generalised to apply to arbitrary unstructured schemas; we state it in restricted form in order to simplify the notation used.

Clearly the relation $\cong_\omega$ is an equivalence relation. For the relation $\cong_v$ with $v \in \mathcal{V}$ we have the following result.

Proposition 20 (transitivity of $\cong_v$ for free schemas) *Let $v \in \mathcal{V}$; then the relation $\cong_v$ is an equivalence relation when restricted to the class of free schemas.*

Proof. Only transitivity is at issue. Suppose $S' \cong_v S''$ and $S'' \cong_v S'''$ hold for free schemas S', S'', S'''. Let i be an interpretation and assume that

$$\perp \notin \{\mathcal{M}[\![S']\!]^i_e, \mathcal{M}[\![S''']\!]^i_e\}$$

holds. Let the interpretation j map every predicate term $p(\mathbf{t})$ to F unless $\pi_{S'}(i,e)$ or $\pi_{S'''}(i,e)$ passes through $p(\mathbf{t})$, in which case let $p^j(\mathbf{t}) = p^i(\mathbf{t})$. Thus $\mathcal{M}[\![S']\!]^i_e = \mathcal{M}[\![S']\!]^j_e$ and $\mathcal{M}[\![S''']\!]^i_e = \mathcal{M}[\![S''']\!]^j_e$ hold and j maps finitely many predicate terms to T, hence $\mathcal{M}[\![S'']\!]^j_e \neq \perp$ holds. Thus

$$\mathcal{M}[\![S']\!]^j_e(v) = \mathcal{M}[\![S'']\!]^j_e(v) = \mathcal{M}[\![S''']\!]^j_e(v)$$

holds, giving the result. $\Box$

Proposition 20 is false for the set of all linear schemas. To see this, consider the three linear schemas

$$\begin{aligned} S' = \ & \mathit{if}\ p(u)\ \ \mathit{then}\ \ v := f_1(); \\ & \qquad\qquad\ \mathit{else}\ \ v := g(); \end{aligned}$$

$$\begin{aligned} S'' = \ & \mathit{while}\, p(u)\ \mathit{do}\ \Lambda; \\ & v := g(); \end{aligned}$$

$$\begin{aligned} S''' = \ & \mathit{if}\ p(u)\ \ \mathit{then}\ \ v := f_2(); \\ & \qquad\qquad\ \mathit{else}\ \ v := g(); \end{aligned}$$

of which S'' is not free. Clearly $S' \cong_v S''$ and $S'' \cong_v S'''$ hold, but not $S' \cong_v S'''$.

We will henceforth refer to a schema which is liberal, free and linear as an LFL schema.

3.4 Subschemas of linear schemas

The *subschemas* of a schema are defined as follows; the empty sequence Λ is a subschema of every schema; if $S \in Sch(\mathcal{F}, \mathcal{P}, \mathcal{V})$ is an assignment or Λ then the only subschemas of S are S itself and Λ; the subschemas of the schema $U_1 \dots U_r$ are those of each U_j for $1 \leq j \leq r$ and also the schemas $U_i U_{i+1} \dots U_j$ for $i \leq j$; the subschemas of $S'' = \mathit{if}\ p(\mathbf{x})\ \mathit{then}\ \{T_1\}\ \mathit{else}\ \{T_2\}$ are S itself and those of T_1 and T_2; the subschemas of $S''' = \mathit{while}\ q(\mathbf{y})\ \mathit{do}\ \{T\}$ are S''' itself and those of T. The subschemas T_1 and T_2 of S'' are called the *true* and *false* parts of p (or of S''). In the *while* schema the subschema T is called the *body* of q (or of S''').

Definition 21 (the subschemas $S(p)$, $\mathit{part}_S^X(p)$ and $\mathit{body}_S(p)$) Let S be a linear schema. If $p \in \mathit{Preds}(S)$ then we sometimes write $S(p)$ for the while or if subschema of S of which p is the guard.
Also, if $p \in \mathit{ifPreds}(S)$ and $X \in \{\mathsf{T}, \mathsf{F}\}$ then we may write $\mathit{part}_S^X(p)$ for the X-part of p in S.
If $p \in \mathit{whilePreds}(S)$ then $\mathit{body}_S(p)$ is the body of p in S.

Definition 22 (the $\searrow_S$ 'lying below' relation, 'immediately below') Let S be a linear schema. If $p \in \mathit{Preds}(S)$, we write $p \searrow_S x$ to mean $x \in \mathit{Symbols}(\mathit{body}_S(p))$ if $p \in \mathit{whilePreds}(S)$ and $x \in \mathit{Symbols}(\mathit{part}_S^{\mathsf{T}}(p)) \cup \mathit{Symbols}(\mathit{part}_S^{\mathsf{F}}(p))$ if $p \in \mathit{ifPreds}(S)$. We may strengthen this to $p \searrow_S x\,(X)$ to mean that either $x \in \mathit{Symbols}(\mathit{part}_S^X(p))$ (if $p \in \mathit{ifPreds}(S)$), or $x \in \mathit{Symbols}(\mathit{body}_S(p))$ (if $X = \mathsf{T}$ and $p \in \mathit{whilePreds}(S)$).
Also, if $A \subseteq \mathit{Symbols}(S)$, then we say that A lies *immediately* below S (or equivalently, S lies immediately above A) if $A \subseteq \mathit{Symbols}(S)$ and there is no $p \in \mathit{whilePreds}(S)$ such that $A \subseteq \mathit{Symbols}(\mathit{body}_S(p))$. In this case, if $S = \mathit{body}_T(q)$ for some linear schema T and $q \in \mathit{whilePreds}(T)$, we may also say that A lies immediately below q in T.

Definition 23 (main subschemas of a linear schema) Let S be a linear schema. The set of *main* subschemas of S contains S itself and the bodies of all while subschemas of S.

Observe that there is exactly one main subschema of a linear schema S lying immediately above a set $A \subseteq \mathit{Symbols}(S)$.

Definition 24 (the $\rightsquigarrow_S$ 'data dependence' relation) Let S be a linear schema. We write $f \rightsquigarrow_S x$ for $f \in \mathit{Funcs}(S)$, $x \in \mathit{Symbols}(S)$ if there is a segment $\tilde{f}\sigma\tilde{x}$ in S such that $\tilde{f}$ is an assignment to f and $\tilde{x} \in \mathit{alphabet}(S)$ satisfies $\mathit{symbol}(\tilde{x}) = x$, and there is no assignment to the variable $\mathit{assign}_S(f)$ along σ. We call $\tilde{f}\sigma\tilde{x}$ an fx-segment in this case. We generalise this by defining $f \rightsquigarrow_S v$ for $f \in \mathit{Funcs}(S)$, $v \in \mathcal{V}$ if $f \rightsquigarrow_{S\,w:=g(v);}\, g$ holds for any linear schema $S\,w:=g(v);$, in which case we define an fv-segment in S to be any segment σ of S such that $\sigma\, w:=g(v)$ is a fg-segment in the schema $S\,w:=g(v);$. Lastly, we write $v \rightsquigarrow_S x$ for $v \in \mathcal{V}$, $x \in \mathit{Symbols}(S)$ if $h \rightsquigarrow_{v:=h();\,S}\, x$ holds for any linear schema $v:=h();\, S$, in which case we define a vx-segment in S to be any $\sigma \in \mathit{pre}(\Pi(S))$ such that $v:=h()\,\sigma$ is an hx-segment in the schema $v:=h();\, S$.
In all cases, we may strengthen the relation $x \rightsquigarrow_S y$ by writing $x \rightsquigarrow_S y\,(n)$ for $n \in \mathbb{N}$ if either $y \in \mathcal{V}$ or the nth component of $\mathbf{refvec}_S(y)$ is x or $\mathit{assign}_S(x)$.

Thus $f \rightsquigarrow_S x$ holds for $f \in \mathit{Funcs}(S)$, $x \in \mathit{Symbols}(S)$ if and only if there exists a path in S along which a (predicate) term $x(\mathbf{t})$ such that $\mathbf{t}$ has a component $f(\mathbf{t}')$ is created; and we may define an fx-segment to be any segment in S which 'witnesses' such a creation. Similar characterisations can be given for the statements $f \rightsquigarrow_S v$ and $v \rightsquigarrow_S x$ for $v \in \mathcal{V}$.
As an example, if T is the linear schema of Figure 3, the relations $v \rightsquigarrow_T q$, $k \rightsquigarrow_T q$, $v \rightsquigarrow_T k$, $k \rightsquigarrow_T k$ (but not $k \rightsquigarrow_T v$), $w \rightsquigarrow_T p$, $h \rightsquigarrow_T f$, $h \rightsquigarrow_T u$, $f \rightsquigarrow_T v$, and $g \rightsquigarrow_T v$ hold.
Note that the relation $\rightsquigarrow_S$ denotes a purely syntactic property of a linear schema S; $f \rightsquigarrow_S x$ may hold even if there is no interpretation defining a path passing through the fx-segment whose existence is asserted.

3.6 Other relations between schema symbols

Definition 25 gives three relations which strengthen the data dependence relation.

Definition 25 (the *outif, thru* and *back* relations) Let S be a linear schema and let $x \in \mathcal{F} \cup \mathcal{V}$ and $y \in \mathcal{F} \cup \mathcal{P} \cup \mathcal{V}$. Let $p \in \mathcal{P}$. Assume that $x \rightsquigarrow_S y$ holds. Then we make the following definitions.

- If $p \in \mathit{whilePreds}(S)$ and both x and y are symbols in $\mathit{body}_S(p)$ but $\neg(x \rightsquigarrow_{\mathit{body}_S(p)} y)$ holds (a backward data dependence) then we write $\mathit{back}_S(p, x, y)$.
- If $Y \in \{\mathsf{T}, \mathsf{F}\}$ and $p \in \mathit{ifPreds}(S)$ and $x \in \mathit{Funcs}(\mathit{part}_S^Y(p))$ and $\neg(x \rightsquigarrow_{\mathit{part}_S^Y(p)} y) \vee (y \in \mathcal{V})$ holds, then we write $\mathit{outif}_S(p, Y, x, y)$. If $Y \in \{\mathsf{T}, \mathsf{F}\}$ and $p \in \mathit{ifPreds}(S)$ and neither x nor y is a symbol in either of the schemas $\mathit{part}_S^{\mathsf{T}}(p)$ or $\mathit{part}_S^{\mathsf{F}}(p)$ and every xy-segment contains the letter $<p = Y>$, then we write $\mathit{thru}_S(p, Y, x, y)$. (Note that $\mathit{thru}_S(p, Y, x, p)$ is always false.)

$$\begin{array}{l} \textit{while}\ p(v)\ \textit{do} \\ \qquad \{ \\ \qquad u := g(v); \\ \qquad v := f(); \\ \qquad \} \end{array}$$

Fig. 4. $back_S(p, f, g)$ holds here

As an example, $back_S(p, f, g)$ holds if S is the linear schema in Figure 4.

Definition 26 (q-competing function symbols and variables) Let S be a linear schema and assume that $f \rightsquigarrow_S x\,(n)$ and $g \rightsquigarrow_S x\,(n)$ for $f, g, x \in Symbols(S) \cup \mathcal{V}$ and $n \in \mathbb{N}$. Let $q \in ifPreds(S)$. We say that f and g are q-competing for x in S if for $\{X, Y\} = \{\mathsf{T}, \mathsf{F}\}$, we have both $outif_S(q, X, f, x) \vee thru_S(q, X, f, x)$ and $outif_S(q, Y, g, x) \vee thru_S(q, Y, g, x)$.

Thus f and g are p-competing for v in the schemas of Figures 1 and 5. Proposition 27 shows that if $thru_S(p, Y, x, y)$ holds for suitable p, Y, x, y then $outif_S(p, \neg Y, f', y)$ holds for some function symbol f'.

Proposition 27 (connection between $outif_S$ and $thru_S$) *Let S be a linear schema and assume that $thru_S(p, Y, x, y)$ holds for some $p \in ifPreds(S)$, $Y \in \{\mathsf{T}, \mathsf{F}\}$ and $x, y \in Symbols(S) \cup \mathcal{V}$. Assume that $x \rightsquigarrow_S y\,(n)$ holds for $n \in \mathbb{N}$. Then every path in $\Pi(part_S^{\neg Y}(p))$ passes through some $f' \in \mathcal{F}$ satisfying $outif_S(p, \neg Y, f', y)$ and $f' \rightsquigarrow_S y\,(n)$.*

Proof. We may assume that $x \in \mathcal{F}$ holds, otherwise we may replace S with a linear schema $x := f();\, S$. Since $thru_S(p, Y, x, y)$ holds, there is an xy-segment $\gamma = \mu <\!p = Y\!> \mu'\mu''$ in S with $\mu' \in \Pi(part_S^Y(p)))$ and $\mu'' \neq \Lambda$. We may assume that p occurs only once in the segment γ; otherwise we could delete a segment from within γ. Let $\sigma \in \Pi(part_S^{\neg Y}(p))$. The segment $\mu <\!p = \neg Y\!> \sigma\mu''$ does not enter the Y-part of p and so is not an xy-segment, by the definition of $thru_S(p, Y, x, y)$. Thus the variable assigned by x is 'killed' along σ, giving the result. $\square$

$$\begin{array}{lll} u := h(); & & \\ v := f(u); & & \\ \textit{if}\ p(w) & \textit{then} & \Lambda \\ & \textit{else} & v := g(); \end{array}$$

Fig. 5. $thru_S(p, \mathsf{T}, f, v) \wedge outif_S(p, \mathsf{F}, g, v)$ holds here

Definition 28 (the $above_S$ function) Let S be a linear schema and let x be a symbol in S. If x lies immediately below S, then we define $above_S(x) = x$; otherwise we define $above_S(x)$ to be the while predicate lying immediately below S and containing x in its body.

Proposition 29 (connection between $above_S$ and $back_S$ relations) *Let S be a linear schema and assume that $back_S(p, f, x)$ holds for some $f \in \mathcal{F}$, $x \in Symbols(S)$ and $p \in whilePreds(S)$. If $f \neq x$ then $above_{body_S(p)}(f) \neq above_{body_S(p)}(x)$ holds.*

Proof. Since $f \neq x$ holds, $above_{body_S(p)}(f) = above_{body_S(p)}(x)$ implies

$$above_{body_S(p)}(f) = q \in whilePreds(body_S(p))$$

holds. Thus $f \rightsquigarrow_{S(q)} x$ and hence $f \rightsquigarrow_{body_S(p)} x$ holds, contradicting $back_S(p, f, x)$. $\square$

Definition 30 (the $\ll_S$ relation) Let S be a linear schema and let

$$\{x, y\} \subseteq Symbols(S).$$

Assume that S lies immediately above the set $\{x, y\}$. We define $x \ll_S y$ if $above_S(x) \neq above_S(y)$ and there is a segment in S which begins at $above_S(x)$ and ends at $above_S(y)$.

Observe the following; if $x \ll_S y$ then every segment in S which begins at x and ends at y passes through every occurrence of x before any occurrence of y, and x and y do not lie in opposite parts of any if predicate. Also, $\ll_S$ is transitive; and $x \ll_S y \wedge y \ll_S x$ never holds, since otherwise S would contain a while predicate containing both $above_S(x)$ and $above_S(y)$ in its body.
It can be shown (see [7, Lemma 134]) that if $back_S(q, f, x)$ holds for $q \in whilePreds(S)$, then $\neg(f \ll_{body_S(q)} x)$ holds.

3.7 The $\mathcal{N}_S$ and Inv_S sets

The symbol and variable sets of Definition 31 are purely syntactically defined, and contain all the symbols and (initial) variables which can influence the final value of a variable. This is stated precisely in Theorem 37.

Definition 31 (symbols *needed* by variables) Let S be a linear schema and let $u \in \{\omega\} \cup \mathcal{V}$. Then we define the set $\mathcal{N}_S(u)$ to be the minimal subset of $Symbols(S)$ satisfying the following closure conditions; if $f \in \mathcal{F}$, $x \in (\mathcal{V} \cap \{u\}) \cup \mathcal{N}_S(u)$ and $f \rightsquigarrow_S x$ then $f \in \mathcal{N}_S(u)$; and if $u = \omega$ then $whilePreds(S) \subseteq \mathcal{N}_S(u)$; and if $p \searrow_S x$ for $x \in \mathcal{N}_S(u)$ then $p \in \mathcal{N}_S(u)$.
We also define $Inv_S(u) \subseteq \mathcal{V}$ to contain all variables v satisfying $v \rightsquigarrow_S v$ if $v = u \in \mathcal{V}$ or $v \rightsquigarrow_S y$ for some $y \in \mathcal{N}_S(u)$.

We generalise this by defining $\mathcal{N}_S(V) = \cup_{u \in V} \mathcal{N}_S(u)$ for a set V, and similarly with Inv_S.

The functions $\mathcal{N}_S$, Inv_S have more restricted domains in Definition 31 above than in [7, Definition 35], in which $\mathcal{N}_S(x)$ and $\mathit{Inv}_S(x)$ for $x \in \mathit{Symbols}(S)$ are also defined.

Note that $\mathcal{N}_S(y)$ is a set of symbols of S, whereas $\mathit{Inv}_S(y)$ is a subset of $\mathcal{V}$.

It can easily be proved that if $v \in \mathcal{V}$ and a linear schema $S = AB$, then $\mathit{Inv}_S(v) = \mathit{Inv}_A(\mathit{Inv}_B(v))$.

Observe that if any of the relations given in Definition 25 hold, and $y \in \mathcal{N}_S(u)$ for some $u \in \mathcal{V} \cup \{\omega\}$, then $x \in \mathcal{N}_S(u)$ holds; in the case that $\mathit{thru}_S(p, Y, x, y)$ holds, this follows from Proposition 27.

Definition 32 (dependence sequences, $\mathit{depnum}_S(x, u)$**)** Let S be a linear schema and assume that $x \in \mathcal{N}_S(u)$ or $x \in \mathit{Inv}_S(u)$. Then a u-dependence sequence for x in S is a word $w \in (\mathcal{P} \cup \mathcal{F} \cup \mathcal{V})^* - \mathcal{V}$ beginning in x and ending in u (or in a element of $\mathit{whilePreds}(S)$ if $u = \omega$) which 'witnesses' this fact; that is, only the first and last letters of w may be variables; also, if $f \in \mathcal{F}$ is in w, and z is the next letter in w after f, then $f \rightsquigarrow_S z$; if $p \in \mathcal{P}$ is in w, and z is the next letter in w after p, then $p \searrow_S z$. We define $\mathit{depnum}_S(x, u) \in \mathbb{N}$ to be the minimal length of any u-dependence sequence for x in S.

Observe that if $v \in \mathcal{V}$ and $x \in \mathcal{N}_S(v) \cup \mathit{Inv}_S(v)$ then $\mathit{depnum}_S(x, v) \geq 2$ holds, since v is not a dependence sequence.
Clearly if w is a u-dependence sequence for x of minimal length, then w contains no repeated letters in $\mathcal{P} \cup \mathcal{F}$, and $\neg(p \searrow_S z)$ holds for any p in w and any letter z occurring after p in w, unless z occurs immediately after p.

Proposition 33 (form of dependence sequences) *Let S be a linear schema and let $u \in \mathcal{V} \cup \{\omega\}$ and $x \in \mathcal{N}_S(u) \cup \mathit{Inv}_S(u)$.*
If $\mathit{depnum}_S(x, u) \geq 2$, then there exists $y \in \mathcal{N}_S(u) \cup \{u\}$ such that either $x \rightsquigarrow_S y$ or $x \searrow_S y$, and $y = u \vee \mathit{dep}(y, u) = \mathit{depnum}_S(x, u) - 1$ holds.
Furthermore, if also $\mathit{depnum}_S(x, u) \geq 3$ holds, then this $y \in \mathit{Symbols}(S)$, and if $x \searrow_S y(Y)$ holds then $y \rightsquigarrow_S z$ holds for some z such that $(z = u \in \mathcal{V}) \vee (z \in \mathcal{N}_S(u) \wedge \mathit{depnum}_S(z, u) = \mathit{depnum}_S(x, u) - 2 \wedge \neg(x \searrow_S z))$.

Proof. This follows immediately from the definition of a minimal-length u-dependence sequence for x in S. $\square$

Definition 34 (u-similar and u-congruent linear schemas) Let S, T be linear schemas and let $u \in \{\omega\} \cup \mathcal{V}$. Then $S\, \mathit{simil}_u\, T$ (S is u-similar to T) if and only if the following hold:

(1) $\mathcal{N}_S(u) = \mathcal{N}_T(u)$;
(2) $\mathcal{N}_S(u) \cap \mathit{ifPreds}(S) = \mathcal{N}_T(u) \cap \mathit{ifPreds}(T)$;
(3) $\mathcal{N}_S(u) \cap \mathit{whilePreds}(S) = \mathcal{N}_T(u) \cap \mathit{whilePreds}(T)$;
(4) $f \rightsquigarrow_S x\,(n) \wedge x \in \mathcal{N}_S(u) \iff f \rightsquigarrow_T x\,(n) \wedge x \in \mathcal{N}_T(u)$, for all $f \in \mathcal{F}$ and $n \geq 1$;
(5) $f \rightsquigarrow_S u \iff f \rightsquigarrow_T u$ if $u \in \mathcal{V}$ and $f \in \mathcal{F}$;
(6) $v \rightsquigarrow_S x\,(n) \iff v \rightsquigarrow_T x\,(n)$ for all $v \in \mathcal{V}$ and $x \in \mathcal{N}_S(u)$ and $n \geq 1$;
(7) $q \searrow_S p\,(Z) \iff q \searrow_T p\,(Z)$ if $u = \omega$ and $p \in \mathit{whilePreds}(S)$ and q is any predicate and $Z \in \{\mathsf{T}, \mathsf{F}\}$;
(8) $\mathit{Symbols}(\mathit{body}_S(p)) \cap \mathcal{N}_S(u) = \mathit{Symbols}(\mathit{body}_T(p)) \cap \mathcal{N}_T(u)$ if $p \in \mathit{whilePreds}(S)$;
(9) $\mathit{back}_S(p, f, x) \wedge x \in \mathcal{N}_S(u) \iff \mathit{back}_T(p, f, x) \wedge x \in \mathcal{N}_T(u)$;
(10) If $q \in \mathit{ifPreds}(S)$ and $Z \in \{\mathsf{T}, \mathsf{F}\}$ and $f \in \mathcal{F}$ and $x \in \mathcal{N}_S(u) \cup (\mathcal{V} \cap \{u\})$ then

$$\mathit{outif}_S(q, Z, f, x) \vee \mathit{thru}_S(q, Z, f, x) \iff \mathit{outif}_T(q, Z, f, x) \vee \mathit{thru}_T(q, Z, f, x);$$

(11) If $f, f' \in \mathcal{F}$ and $f, f' \rightsquigarrow_S x\,(r)$ for $x \in \mathcal{N}_S(u) \cup (\{u\} \cap \mathcal{V})$, and $r \in \mathbb{N}$, and $\bar{S}, \bar{T}$ are the main subschemas of S and T respectively lying immediately above $\{f, f'\}$, then either $\neg(f \ll_{\bar{S}} f' \wedge f' \ll_{\bar{T}} f)$ holds, or there exists $q \in \mathit{ifPreds}(S)$ such that f and f' are q-competing for x in S;
(12) If $p \in \mathit{whilePreds}(S)$, $f \in \mathcal{F}$ and $f \rightsquigarrow_S x \wedge x \in \mathcal{N}_S(u)$ and $v = \mathit{assign}_S(f)$ and $w = \mathit{assign}_T(f)$, then

$$f \rightsquigarrow_{\mathit{body}_S(p)} v \wedge v \rightsquigarrow_{\mathit{body}_S(p)} x$$
$$\iff$$
$$f \rightsquigarrow_{\mathit{body}_T(p)} w \wedge w \rightsquigarrow_{\mathit{body}_T(p)} x$$

holds.
(13) If $p \in \mathit{whilePreds}(S)$, $q \in \mathit{ifPreds}(S)$, $f \in \mathit{Funcs}(S)$, $x \in \mathcal{N}_S(u)$, $Z \in \{\mathsf{T}, \mathsf{F}\}$ and $f \rightsquigarrow_S x$, with $v = \mathit{assign}_S(f)$ and $w = \mathit{assign}_T(f)$ and $v \rightsquigarrow_{\mathit{body}_S(p)} x$, then

$$\mathit{outif}_{\mathit{body}_S(p)}(q, Z, f, v) \vee \mathit{thru}_{\mathit{body}_S(p)}(q, Z, f, v)$$
$$\iff$$
$$\mathit{outif}_{\mathit{body}_T(p)}(q, Z, f, w) \vee \mathit{thru}_{\mathit{body}_T(p)}(q, Z, f, w)$$

holds.

If $S\, \mathit{simil}_u\, T$ and also $\mathbf{refvec}_S(x) = \mathbf{refvec}_T(x)$ for all $x \in \mathcal{N}_S(u)$ and $\mathit{assign}_S(f) = \mathit{assign}_T(f)$ for all $f \in \mathcal{N}_S(u) \cap \mathcal{F}$, then we say that S and T are u-congruent, written $S\, \mathit{cong}_u\, T$.

We also write $S\, \mathit{simil}_V\, T$ to mean that $S\, \mathit{simil}_u\, T$ for all $u \in V$, and $S\, \mathit{simil}\, T$ to mean that $S\, \mathit{simil}_{\mathcal{V} \cup \{\omega\}}\, T$ holds. Also $S\, \mathit{cong}_V\, T$ has a similar meaning.

Observe that the two linear predicate-free schemas

$$u := f();$$
$$v := g(u);$$

and

$$u' := f();$$
$$v := g(u');$$

are v-similar but not v-congruent if $u \neq u'$; thus congruence is a stronger condition than similarity.

Informally, for two linear structured schemas S, T to satisfy $S\,simil_u\,T$, the following must hold;

- S and T have the same set of u-needed function symbols, if predicate symbols and while predicate symbols. (Conditions (1), (2), (3) of $S\,simil_u\,T$).
- S and T have the same *data dependence* relations among those symbols in $\mathcal{N}_S(u)$. (Conditions (4), (5),(6) of $S\,simil_u\,T$).
- S and T have the same set of u-needed symbols lying in the body of each while predicate (Condition (8) of $S\,simil_u\,T$). If $u = \omega$ a weaker statement also holds for while predicates lying under if predicates (Condition (7) of $S\,simil_u\,T$).
- Also, the bodies of while predicates in S and T satisfy the same data dependence conditions between symbols lying in $\mathcal{N}_S(u)$ (Conditions (9), (12) of $S\,simil_u\,T$).
- Conditions (10), (11) and (13) of $S\,simil_u\,T$ are a kind of counterpart for function symbols lying under if predicates to Condition (8) for symbols lying under while predicates, showing that change of ordering with respect to $\ll_S$ of function symbols (as with f, g in the v-equivalent schemas given in Figures 1 and 5) can only occur in connection with an if predicate.

Theorem 35 ($S\,simil_u\,T$ is decidable in polynomial time) *Given linear schemas S and T and $u \in \mathcal{V} \cup \{\omega\}$, it is decidable in polynomial time whether $S\,simil_u\,T$ holds.*

Proof. Given a linear schema S, encoded as indicated in Definition 3, with the braces $\{\ \}$, the truth of the relations $p \searrow_S x\,(Z)$ for each $p \in Preds(S)$, $x \in Symbols(S)$, $Z \in \{\mathsf{T}, \mathsf{F}\}$ can be established in polynomial time. Given two elements $v, w \in alphabet(S)$, with symbols v', w', we can decide in polynomial time whether w occurs immediately after v in any word in $\Pi(S)$, since this holds if and only if either w' occurs after v' in S without there being any other symbol between them, and $p \searrow_S v' \iff p \searrow_S w'$ for all $p \in whilePreds(S)$, or $v' \in whilePreds(S)$ and v' lies immediately above w' and there are no symbols occurring after v' in S before the closing brace $\}$ defined by v'. Thus we can construct in polynomial time a directed graph G_S, whose vertices are the elements of $alphabet(S)$ and such that there is an edge from vertex v to w in the graph

G_S if and only if w occurs immediately after v in a word in $\Pi(S)$. Given $f \in \mathit{Funcs}(S)$ and $x \in \mathit{Symbols}(S)$, we can establish whether $f \rightsquigarrow_S x$ holds by deleting all vertices in G_S that are assignments to $\mathit{assign}_S(f)$ except the one with function symbol f or x, if $x \in \mathcal{F}$, and edges adjacent to deleted vertices, and establishing whether the letter containing x is reachable from the f-assignment in the resulting directed graph. This latter problem is well-known to be polynomial-time decidable in the size of G_S. The values of n for which $f \rightsquigarrow_S x\,(n)$ also holds can also be easily established, as can the truth of the assertions $v \rightsquigarrow_S x\,(n)$ for $v \in \mathcal{V}$ and $f \rightsquigarrow_S u$. Also, the truth of the relations above_S and $\ll_S$ for appropriate arguments can be decided in polynomial time by studying S. Having obtained this information, we can test the truth of the relations back_S, outif_S, thru_S (and hence the q-competing condition) for appropriate arguments. By comparing this information with that obtained from T and the graph G_T, it can be decided in polynomial time whether S and T satisfy $S\,\mathit{simil}_u\,T$. $\square$

4 Slices of schemas

An important special case of the equivalence problem for schemas S, T is that in which T is a *slice* of S.

Definition 36 A *slice* of a structured schema S may be obtained recursively by the following rules;

- if $S = S_1S_2S_3$ then S_1S_3, S_1S_2 and S_2S_3 are slices of S;
- if T' is a slice of T then $\mathit{while}\, p(\mathbf{u})\, \mathit{do}\, T'$ is a slice of $\mathit{while}\, p(\mathbf{u})\, \mathit{do}\, T$;
- if T' is a slice of T then the if schema $\mathit{if}\; q(\mathbf{u})\; \mathit{then}\, S\; \mathit{else}\, T'$ is a slice of $\mathit{if}\; q(\mathbf{u})\; \mathit{then}\, S\; \mathit{else}\, T$ (the true and false parts may be interchanged in this example);
- a slice of a slice of S is itself a slice of S.

The following facts are easily proved. All slices of a linear schema are also linear. If a set $\Sigma \subseteq \mathit{Symbols}(S)$ (for linear S) satisfies $(x \in \Sigma \wedge p \searrow_S x) \Rightarrow p \in \Sigma$, then there is a unique slice T of S satisfying $\mathit{Symbols}(T) = \Sigma$; the slice T can be obtained from S by successively removing all assignments whose function symbols do not lie in Σ, and every if and while subschema of S whose guard does not lie in Σ.

A special case is given by $\Sigma = \mathcal{N}_S(V)$ for $V \subseteq \mathcal{V} \cup \{\omega\}$. In this case every slice T of S containing all symbols in $\mathcal{N}_S(V)$ satisfies $\mathit{Inv}_T(V) = \mathit{Inv}_S(V)$ and $S\,\mathit{cong}_V\,T$, since deletion from S of symbols not lying in $\mathcal{N}_S(V)$ does not affect the schema properties defining these statements. We will show in Part (2) of Theorem 37 that $S \cong_V T$ also holds.

A slice of an LFL schema need not be free or liberal; for example, the schema $\mathit{while}\, p(v)\, \mathit{do}\, \Lambda$, which is not free, is a slice of the LFL schema below;

while $p(v)$ *do*

$$\{$$
$$u := h(u);$$
$$w := k(u);$$
$$v := g(v);$$
$$\}$$

Also, deleting the assignment $u := h(u)$; gives a schema which is free but not liberal. However the slice of an LFL schema S which contains precisely the symbols in $\mathcal{N}_S(V)$ for any $V \subseteq \mathcal{V}$ is itself LFL; this follows from Theorem 19 and the 'backward data dependence' property of $\mathcal{N}_S(V)$.

Theorem 37 was proved by Weiser in [35], using different terminology.

Theorem 37 *Let S be a (not necessarily free or liberal) linear schema and let T be a slice of S. Let $u \in \mathcal{V} \cup \{\omega\}$, let i, j be interpretations differing only on predicates not lying in $\mathcal{N}_S(u)$, and let c, d be states such that $c(v) = d(v)$ for all $v \in Inv_S(u)$. Assume that T contains every symbol of $\mathcal{N}_S(u)$.*

(1) *If $Symbols(T) = \mathcal{N}_S(u)$, then $\mathcal{M}[\![S]\!]^i_c \neq \perp \Rightarrow \mathcal{M}[\![T]\!]^j_d \neq \perp$.*
(2) *If $u \in \mathcal{V}$ and $\mathcal{M}[\![S]\!]^i_c$ and $\mathcal{M}[\![T]\!]^j_d$ both terminate, then $\mathcal{M}[\![S]\!]^i_c(u) = \mathcal{M}[\![T]\!]^j_d(u)$; and if $u = \omega$ then $\mathcal{M}[\![S]\!]^i_c \neq \perp \iff \mathcal{M}[\![T]\!]^j_d \neq \perp$.*

In particular, $S \cong_u T$ holds.

Proof. This is proved in [7, Theorem 42]. $\Box$

Part (1) of Theorem 37 may fail for a slice T whose symbol set strictly contains $\mathcal{N}_S(u)$; for example, if S is

$v = f()$;
while $p(v)$ *do* Λ

and T is the slice

while $p(v)$ *do* Λ

for a variable $v \neq u$. If the interpretation i maps every predicate term $p(t)$ to T unless $t = f()$ then $\mathcal{M}[\![S]\!]^i_e$ terminates whereas $\mathcal{M}[\![T]\!]^i_e$ does not.

Theorem 50 is the main result of this section. The proof of this theorem procedes along the following lines. By Part (2) of Theorem 37, we may assume that u-similar linear schemas S and T contain only the symbols in $\mathcal{N}_S(u)$, since replacing schemas S and T by their respective slices containing only these symbols preserves u-similarity and congruence. We will prove that we may also assume that S and T are u-congruent. We will prove that we can make a further simplifying assumption about S and T, namely that there are no if predicates lying immediately below S or T. In order to do this, we have to define the schema $S/(q = Z)$ with the $\neg Z$-part of q deleted, where $q \in \mathit{ifPreds}(S)$ and $Z \in \{\mathsf{T}, \mathsf{F}\}$. In essence, a linear schema S which has m if predicates lying immediately below it may be thought of as a set of 2^m linear schemas, one for each choice from $\{\mathsf{T}, \mathsf{F}\}$ for each of these if predicates. We will show that u-congruence of S and T implies u-congruence for each of these schemas with its counterpart for T, and that any interpretation defines paths through S and T which make the same choice from $\{\mathsf{T}, \mathsf{F}\}$ in S and T for each of these if predicates. Thus we need to define the truncated schema $\mathit{trunc}_S\,(q)$, which is the 'bit' of S occurring before the if predicate q is reached, and prove that $\mathit{trunc}_S\,(q)$ and $\mathit{trunc}_T\,(q)$ are v-congruent for all variables v referenced by q. Thus we reduce the problem to the special case in which S and T are sequences of assignments and while schemas.

5.1 Deleting parts of if schemas

Definition 38 (deleting a part of an if schema) Let S be a linear schema and let $q \in \mathcal{P}$ and $Z \in \{\mathsf{T}, \mathsf{F}\}$. Assume that $q \in \mathit{ifPreds}(S)$ holds. The linear schema $S/(q = Z)$ is obtained by replacing the if subschema $S(q)$ of S by its Z-part, $\mathit{part}^Z_S(q)$.

Proposition 39 (characterising $\mathit{pre}(\Pi(S/(q = Z)))$) *Let S be a linear schema. Let $q \in \mathit{ifPreds}(S)$ and $Z \in \{\mathsf{T}, \mathsf{F}\}$. Let $\mu \in \mathit{alphabet}(S)^*$. Then $\mu \in \mathit{pre}(\Pi(S/(q = Z)))$ if and only if there exists $\rho \in \mathit{pre}(\Pi(S))$ not passing through $<q = \neg Z>$ and such that μ is obtained from ρ by deleting all occurrences of $<q = Z>$.*
$\square$

Lemma 40 (relating S to $S/(q = Z)$) *Let S be a linear schema. Let $q \in \mathit{ifPreds}(S)$ and $Z \in \{\mathsf{T}, \mathsf{F}\}$. Then the following hold.*

(1) *Let $f \in \mathit{Funcs}(S)$ and $y \in \mathit{Symbols}(S/(q = Z)) \cup \mathcal{V}$. Then*

$$f \rightsquigarrow_{S/(q=Z)} y \iff f \rightsquigarrow_S y \wedge \neg(\mathit{outif}_S(q, \neg Z, f, y) \vee \mathit{thru}_S(q, \neg Z, f, y))$$

holds.

(2) *Let $y \in \mathit{Symbols}(S/(q = Z))$, $Y \in \{\mathsf{T}, \mathsf{F}\}$ and $p \in \mathit{Preds}(S) - \{q\}$. Then $p \searrow_{S/(q=Z)} y\,(Y) \iff p \searrow_S y\,(Y)$.*

(3) *Let* $f \in \mathit{Funcs}(S)$ *and* $y \in \mathit{Symbols}(S/(q=Z)) \cup \mathcal{V}$. *Assume* $f \rightsquigarrow_{S/(q=Z)} y$ *holds. Let* $p \in \mathit{ifPreds}(S) - \{q\}$ *and* $Y \in \{\mathsf{T}, \mathsf{F}\}$. *Then*

$$\mathit{outif}_S(p, Y, f, y) \Rightarrow \mathit{outif}_{S/(q=Z)}(p, Y, f, y)$$

and

$$\mathit{thru}_S(p, Y, f, y) \Rightarrow \mathit{thru}_{S/(q=Z)}(p, Y, f, y)$$

hold, and if $\mathit{above}_S(q) = q$ *and* $\mathit{above}_S(p) \neq p$, *then the converse statements are true.*

Proof.

(1) Assume that $f \rightsquigarrow_{S/(q=Z)} y$ holds. Thus $\neg(q \searrow_S f(\neg Z))$ and so $\neg \mathit{outif}_S(q, \neg Z, f, y)$ holds. Let $\mu \in \mathit{pre}(\Pi(S/(q=Z)))$ end in an fy-segment. By Proposition 39, there is a prefix ρ in S not passing through $\lessdot q = \neg Z \gtrdot$ such that μ is obtained from ρ by deleting all occurrences of $\lessdot q = Z \gtrdot$; hence ρ ends in an fy-segment in S and so $f \rightsquigarrow_S y \wedge \neg \mathit{thru}_S(q, \neg Z, f, y)$ holds.
Conversely, assume $f \rightsquigarrow_S y \wedge \neg(\mathit{outif}_S(q, \neg Z, f, y) \vee \mathit{thru}_S(q, \neg Z, f, y))$ holds. By the hypotheses, $y \neq q \wedge \neg(q \searrow_S y(\neg Z))$ and hence $\neg(q \searrow_S f(\neg Z))$ holds. Thus there exists a prefix ρ in S which ends in an fy-segment and does not pass through $\lessdot q = \neg Z \gtrdot$. By Proposition 39, deleting all occurrences of $\lessdot q = Z \gtrdot$ from ρ gives an fy-segment in $S/(q=Z)$.

(2) This follows easily from the definition of $S/(q=Z)$.

(3) We have four cases.

- Assume $\mathit{outif}_S(p, Y, f, y)$ holds. Thus $p \searrow_S f(Y)$ and since $f \in \mathit{Funcs}(S/(q = Z))$ by the hypotheses, $p \searrow_{S/(q=Z)} f(Y)$ by Part (2) of this Lemma; in particular, $p \in \mathit{Symbols}(S/(q = Z))$. Assume that $\neg \mathit{outif}_{S/(q=Z)}(p, Y, f, y)$; thus since $f \rightsquigarrow_{S/(q=Z)} y$ holds by the hypotheses, $f \rightsquigarrow_{\mathit{part}^Y_{S/(q=Z)}(q)} y$ holds. If $\neg(p \searrow_S q(Y))$ then $\mathit{part}^Y_{S/(q=Z)}(q) = \mathit{part}^Y_S(p)$, and if $p \searrow_S q(Y)$ then $\mathit{part}^Y_{S/(q=Z)}(q) = \mathit{part}^Y_S(p)/(q = Z)$ can be easily shown. Thus in either case $f \rightsquigarrow_{\mathit{part}^Y_S(p)} y$ holds, contradicting $\mathit{outif}_S(p, Y, f, y)$.
- Now assume $\mathit{thru}_S(p, Y, f, y)$ holds. Thus

$$f, x \notin \mathit{Symbols}(S(p)) \supseteq \mathit{Symbols}(S/(q = Z)(p)).$$

  If $\neg \mathit{thru}_{S/(q=Z)}(p, Y, f, y)$ holds, then since $f \rightsquigarrow_{S/(q=Z)} y$ holds, there is a prefix μ in the schema $S/(q = Z)$ which ends in an fy-segment does not pass through $\lessdot p = Y \gtrdot$. Thus by Proposition 39, there is a prefix ρ in S not passing through $< q = \neg Z >$ such that μ is obtained from ρ by deleting all occurrences of $\lessdot q = Z \gtrdot$; hence ρ ends in an fy-segment and so $f \rightsquigarrow_S y \wedge \neg \mathit{thru}_S(q, \neg Z, f, y)$ holds. Since $q \neq p$, ρ also does not pass through $< p = Y >$, contradicting $\mathit{thru}_S(p, Y, f, y)$.
- Assume $\neg \mathit{outif}_S(p, Y, f, y) \wedge \mathit{outif}_{S/(q=Z)}(p, Y, f, y)$ and $\mathit{above}_S(q) = q$ and $\mathit{above}_S(p) \neq p$ hold. Thus $p \searrow_{S/(q=Z)} f(Y)$ and so $p \searrow_S f(Y)$ by Part (2) of this Lemma. Hence there is an fy-segment μ in the schema $\mathit{part}^Y_S(p)$. Clearly

$\neg(p \searrow_S q)$ holds under the hypotheses given, and since $p \in \mathit{Preds}(S/(q=Z))$, μ is also an an fy-segment in $S/(q=Z)$, giving a contradiction.

- Assume $\neg \mathit{thru}_S(p,Y,f,y) \wedge \mathit{thru}_{S/(q=Z)}(p,Y,f,y)$ and $\mathit{above}_S(q)=q$ and $\mathit{above}_S(p) \neq p$ hold. Let $\mathit{above}_S(p)=p' \neq p$. Clearly $p' \in \mathit{whilePreds}(S)$. Since $p \in \mathit{ifPreds}(S/(q=Z))$ and $\mathit{above}_S(q)=q \neq p'$, $S(p')$ is a while schema in $S/(q=Z)$; thus since
$$\mathit{thru}_{S/(q=Z)}(p,Y,f,y)$$
holds, either $p' \searrow_S f$ or $p' \searrow_S y$ holds, since otherwise an fy-segment in $S/(q=Z)$ need never enter the body of p', contradicting $\mathit{thru}_{S/(q=Z)}(p,Y,f,y)$. If both hold then $\mathit{thru}_{S(p')}(p,Y,f,y)$ and hence $\mathit{thru}_S(p,Y,f,y)$ follows, giving a contradiction. If $p' \searrow_S f \wedge \neg(p' \searrow_S y)$ holds, then
$$\mathit{thru}_{S(p')}(p,Y,f,\mathit{assign}_S(f))$$
holds, and if $p' \searrow_S y \wedge \neg(p' \searrow_S f)$ holds, then
$$\mathit{thru}_{S(p')}(p,Y,\mathit{assign}_S(f),y)$$
holds, in both cases contradicting $\neg\mathit{thru}_S(p,Y,f,y)$.

$\square$

Corollary 41 (iterating Lemma 40) *Let S be a linear schema. Let $\{q(1),\ldots,q(m)\}$ be a set of if predicates in S which all lie immediately below S and let $Z(1),\ldots,Z(m) \in \{\mathsf{T},\mathsf{F}\}$. Define $S' = S/(q(1)=Z(1))/\ldots/(q(m)=Z(m))$. Then the following hold.*

(1) *Let $f \in \mathit{Funcs}(S)$ and $y \in \mathit{Symbols}(S') \cup \mathcal{V}$. Then*
$$f \rightsquigarrow_{S'} y$$
$$\Longleftrightarrow$$
$$f \rightsquigarrow_S y \wedge \Big(\bigwedge_{i \le m} \neg(\mathit{outif}_S(q(i),\neg Z,f,y) \vee \mathit{thru}_S(q(i),\neg Z,f,y))\Big)$$
holds.

(2) *Let $y \in \mathit{Symbols}(S')$, $Y \in \{\mathsf{T},\mathsf{F}\}$ and $p \in \mathit{Preds}(S) - \{q(1),\ldots,q(m)\}$. Then $p \searrow_{S'} y\,(Y) \iff p \searrow_S y\,(Y)$.*

(3) *Let $f \in \mathit{Funcs}(S)$ and $y \in \mathit{Symbols}(S') \cup \mathcal{V}$. Assume $f \rightsquigarrow_{S'} y$ holds. Let $p \in \mathit{ifPreds}(S) - \{q(1),\ldots,q(m)\}$ and $Y \in \{\mathsf{T},\mathsf{F}\}$. Then*
$$\mathit{outif}_S(p,Y,f,y) \Rightarrow \mathit{outif}_{S'}(p,Y,f,y)$$
and $\mathit{thru}_S(p,Y,f,y) \Rightarrow \mathit{thru}_{S'}(p,Y,f,y)$ hold, and if $\mathit{above}_S(p) \neq p$, then the converse statements are true.

Proof. All parts follow straightforwardly from Lemma 40, using induction on m. $\square$

Theorem 42 (u-similarity is inherited by part deletion) *Let S, T be linear schemas and assume $S\, simil_u\, T$ for $u \in \mathcal{V} \cup \{\omega\}$. Assume $Symbols(S) = Symbols(T) = \mathcal{N}_S(u)$. Let $q(1), \ldots, q(m)$ be the set of all if predicates in $\mathcal{N}_S(u)$ lying immediately below S and let $Z(1), \ldots, Z(m) \in \{\mathsf{T}, \mathsf{F}\}$. Define*

$$S' = S/(q(1) = Z(1))/\ldots/(q(m) = Z(m))$$

and define T' similarly. Then $S'\, simil_u\, T'$ holds.

Proof. Let $x \in \mathcal{N}_{S'}(u)$. Then $x \in \mathcal{N}_{T'}(u)$ follows from Parts (1) and (2) of Corollary 41 and Proposition 33, Condition (10) of $S\, simil_u\, T$ and Proposition 27, using induction on $depnum_S(x, u)$. Thus $\mathcal{N}_{S'}(u) \subseteq \mathcal{N}_{T'}(u)$ holds. By interchanging S and T and using Part (1) of $S\, simil_u\, T$ we get $\mathcal{N}_{S'}(u) = \mathcal{N}_{T'}(u)$, thus proving Condition (1) of $S'\, simil_u\, T'$. Conditions (2)–(7) of $S'\, simil_u\, T'$ follow at once from this, using Corollary 41. Condition (8),(9), (12) and (13) of $S'\, simil_u\, T'$ follow from the fact that if $p \in whilePreds(S')$ then the while schema $S'(p) = S(p)$. Condition (10) of $S'\, simil_u\, T'$ is given by Corollary 41, Part (3) and Condition (10) of $S\, simil_u\, T$. Lastly, Condition (11) follows from Condition (10) of $S'\, simil_u\, T'$ and the fact that $f \ll_{S'} g \iff f \ll_S g$ for any $f, g \in Funcs(S')$, and Part (3) of Corollary 41. $\square$

5.2 Truncated schemas

Definition 43 (q-truncated schema) Let S be a linear schema and let $q \in ifPreds(S)$ lie immediately below S. We define the schema $trunc_S(q)$ as follows. Let $\{p(1), \ldots, p(m)\}$ be the set of all if predicates containing q in one part; say $p(i) \searrow_S q\,(Z(i))$ for $Z(i) \in \{\mathsf{T}, \mathsf{F}\}$. Define $S' = S/(p(1) = Z(1))/\ldots/(p(m) = Z(m))$. Write $S' = S''TS'''$ where T is the if subschema of S guarded by q. Then define $trunc_S(q) = S''$.

Observe that for any linear schema S, $Symbols(trunc_S(q)) = \{x \in Symbols(S) | x < <_S q\}$; and if $x \in Symbols(S)$ then $f \rightsquigarrow_{trunc_S(q)} x\,(r) \iff (f \rightsquigarrow_S x\,(r) \wedge x \ll_S q)$ and $x \searrow_{trunc_S(q)} y\,(Z) \iff (x \searrow_S y\,(Z) \wedge x \ll_S q)$ hold.

Lemma 44 (inheritance of congruence in truncated schemas) *Let $u \in \mathcal{V} \cup \{\omega\}$ and let S, T be u-congruent linear schemas. Assume $Symbols(S) = Symbols(T) = \mathcal{N}_S(u)$. Let $q \in ifPreds(S)$ lie immediately below S and let $v \in Refset_S(q)$. Then $trunc_S(q)\, cong_v\, trunc_T(q)$ holds.*

Proof. Write $S' = trunc_S(q)$ and define T' similarly. We prove only $\mathcal{N}_{S'}(v) = \mathcal{N}_{T'}(v)$; all other conditions of $S'\, simil_v\, T'$ follow easily from this; hence $S'\, simil_v\, T'$ and so $S'\, cong_v\, T'$ follows.
Let $x \in \mathcal{N}_{S'}(v)$. We show $x \in \mathcal{N}_{T'}(v)$ by induction on $depnum_{S'}(x, v)$; the result will then follow by interchanging S and T.
If $x \in \mathcal{F}$ then $x \rightsquigarrow_{S'} y$ for some $y \in \mathcal{N}_{S'}(v) \cup \{v\}$ with $depnum_{S'}(y, v) < depnum_{S'}(x, v)$.

If $y = v$ then $x \rightsquigarrow_{T'} y$ and so $x \in \mathcal{N}_{T'}(v)$ is easily shown. If $y \in \mathcal{N}_{S'}(v)$ then $x \rightsquigarrow_S y$ and $y \in \mathcal{N}_{T'}(v)$ by the inductive hypothesis applied to y and $y \ll_T q$ holds; thus $x \rightsquigarrow_T y$ since $S\,cong_u\,T$ and so $x \rightsquigarrow_{T'} y$ follows, proving $x \in \mathcal{N}_{T'}(v)$.
If $x \in \mathcal{P}$ then $x \searrow_{S'} y$ for some $y \in \mathcal{N}_{S'}(v)$ with $depnum_{S'}(y, v) < depnum_{S'}(x, v)$; and by Proposition 33, $y \rightsquigarrow_{S'} z$ for some $z \in \{v\} \cup \mathcal{N}_{S'}(v)$ with $depnum_{S'}(z, v) < depnum_{S'}(y, v)$ and $\neg(y \searrow_{S'} z)$. Thus $y \in \mathcal{N}_{T'}(v)$ by the inductive hypothesis applied to y. If $x \in whilePreds(S')$ then $x \searrow_{T'} y$ follows from Condition (8) of $S\,simil_u\,T$, so assume $x \in ifPreds(S')$. If $z = v$ then $outif_S(x, X, y, q)$ for some $X \in \{\mathsf{T}, \mathsf{F}\}$ and so by Condition (10) of $S\,simil_u\,T$, $outif_T(x, X, y, q) \vee thru_T(x, X, y, q)$ holds; hence by Proposition 27, $outif_T(x, Y, y', q)$ for some $y' \in \mathcal{F}$ and $Y \in \{\mathsf{T}, \mathsf{F}\}$ with $depnum_{S'}(y', v) \leq depnum_{S'}(y, v)$, and so $y' \in \mathcal{N}_{T'}(v)$ by the inductive hypothesis and so $x \in \mathcal{N}_{T'}(v)$ follows. If instead $z \in \mathcal{N}_{S'}(v)$ then $x \in \mathcal{N}_{T'}(v)$ can be proved similarly. $\Box$

5.3 Relating bodies of u-similar while schemas

Lemma 45 (relating while schemas to their bodies) *Let the linear schema $S =$ while $p(\mathbf{x})$ do S'. Then the following hold.*

(1) $v \rightsquigarrow_{S'} z \iff v \rightsquigarrow_S z$ *and* $f \rightsquigarrow_{S'} v \iff f \rightsquigarrow_S v$ *hold for all* $v \in \mathcal{V}$, $z \in Symbols(S')$ *and* $f \in Funcs(S')$.
(2) $f \rightsquigarrow_{S'} z \iff (f \rightsquigarrow_S z \wedge \neg back_S(p, f, z))$ *for all* $f \in \mathcal{F}$ *and* $z \in Symbols(S')$.
(3) *Let* $u \in \mathcal{V} \cup \{\omega\}$*; then* $\mathcal{N}_{S'}(Inv_S(u)) = \mathcal{N}_S(u) - \{p\}$ *and* $Inv_{S'}(Inv_S(u)) \subseteq Inv_S(u)$.

Proof. We prove only (3); (1) and (2) are straightforward.
Let $v \in Inv_S(u)$ and $x \in \mathcal{N}_{S'}(v)$. Clearly $x \neq p$. We show $x \in \mathcal{N}_S(u)$ by induction on $depnum_{S'}(x, v)$. If $depnum_{S'}(x, v) = 2$ then $x \rightsquigarrow_S v$ and either $v = u$ (and so $x \in \mathcal{N}_S(v) = \mathcal{N}_S(u)$) or $v \rightsquigarrow_S y$ for $y \in \mathcal{N}_S(u)$; hence $x \rightsquigarrow_S y$ and so $x \in \mathcal{N}_S(u)$. If $depnum_{S'}(x, v) \geq 3$ then by Proposition 33 $x \rightsquigarrow_{S'} y \vee x \searrow_{S'} y$ for some $y \in \mathcal{N}_{S'}(v)$ with $depnum_{S'}(y, v) < depnum_{S'}(x, v)$; thus $x \in \mathcal{N}_S(u)$ follows by the inductive hypothesis applied to y. Thus we have shown $\mathcal{N}_S(u) - \{p\} \supseteq \mathcal{N}_{S'}(Inv_S(u))$.
Conversely, let $x \in \mathcal{N}_S(u) - \{p\}$. We show $x \in \mathcal{N}_{S'}(Inv_S(u))$ by induction on $depnum_S(x, u)$. If $x \rightsquigarrow_S u \in \mathcal{V}$ then this is immediate since $u \in Inv_S(u)$. If not, then $depnum_S(x, u) \geq 2$ and so by Proposition 33 $x \rightsquigarrow_S y \vee x \searrow_S y$ for some $y \in \mathcal{N}_S(u)$ with $depnum_S(y, u) < depnum_S(x, u)$. If $y = p \vee back_S(p, x, y)$ then $x \rightsquigarrow_S v$ for some $v \in Refset_S(y)$ and $v \rightsquigarrow_S y$, hence $v \in Inv_S(u)$ and so $x \in \mathcal{N}_{S'}(Inv_S(u))$. Lastly, if $x \searrow_S y$ or $x \rightsquigarrow_{S'} y$ then $x \in \mathcal{N}_{S'}(Inv_S(u))$ follows from the inductive hypothesis applied to y.
Thus we have shown $\mathcal{N}_S(u) - \{p\} = \mathcal{N}_{S'}(Inv_S(u))$. Since $v \in Inv_{S'}(Inv_S(u))$ if and only if $v \rightsquigarrow_{S'} v \in Inv_S(u) \vee v \rightsquigarrow_S x \in \mathcal{N}_{S'}(Inv_S(u))$ and $v \in Inv_S(u)$ if and only if $v \rightsquigarrow_{S'} v \in Inv_S(u) \vee v \rightsquigarrow_S x \in \mathcal{N}_S(u)$, we get $Inv_{S'}(Inv_S(u)) \subseteq Inv_S(u)$. $\Box$

Lemma 46 (congruent while schemas have congruent bodies) *Let* $u \in \mathcal{V} \cup \{\omega\}$ *and let* $S = \mathit{while}\, p(\mathbf{x})\, \mathit{do}\, S'$ *and* $T = \mathit{while}\, p(\mathbf{x})\, \mathit{do}\, T'$ *be linear schemas. Assume* $S\, \mathit{cong}_u\, T$*. Then* $S'\, \mathit{cong}_{\mathit{Inv}_S(u)}\, T'$ *holds. Also, if* $u = \omega$ *then* $S'\, \mathit{cong}_\omega\, T'$ *holds.*

Proof. Let $v \in \mathit{Inv}_S(u)$. We will show that $S'\, \mathit{simil}_v\, T'$ holds. Part (3) of Lemma 45 proves $\mathcal{N}_{S'}(v) \cup \mathcal{N}_{T'}(v) \subseteq \mathcal{N}_S(u) = \mathcal{N}_T(u)$. Thus $x \in \mathcal{N}_{S'}(v) \iff x \in \mathcal{N}_{T'}(v)$ can be proved by induction on $\mathit{depnum}_{S'}(x, v)$ using Proposition 27, Part (2) of Lemma 45 and Conditions (1),(9) of $S\, \mathit{simil}_u\, T$, proving Conditions (1) and hence (4) and (6) of $S'\, \mathit{simil}_v\, T'$. All other conditions of $S'\, \mathit{simil}_v\, T'$ except (5) and (10) follow straightforwardly from this.
We now prove Condition (5) of $S'\, \mathit{simil}_v\, T'$. Assume that $f \rightsquigarrow_{S'} v$ for $f \in \mathcal{F}$. Thus $f \in \mathcal{N}_{S'}(v) \subseteq \mathcal{N}_S(u)$. We will show $f \rightsquigarrow_{T'} v$. If $v = u$ then $f \rightsquigarrow_{T'} v$ is immediate. If not, then $v \rightsquigarrow_S y$ for $y \in \mathcal{N}_S(u)$; hence $f \rightsquigarrow_{T'} v$ follows from Condition (12) of $S\, \mathit{simil}_u\, T$ and so Condition (5) of $S'\, \mathit{simil}_v\, T'$ follows by interchanging S and T.
Lastly we show Condition (10) of $S'\, \mathit{simil}_v\, T'$. Assume
$\mathit{outif}_{S'}(q, Z, f, x) \vee \mathit{thru}_{S'}(q, Z, f, x)$ holds for $q \in \mathit{ifPreds}(S)$ and $Z \in \{\mathsf{T}, \mathsf{F}\}$ and $f \in \mathcal{F}$ and $x \in \mathcal{N}_S(v) \cup \{v\}$. If $x \in \mathcal{N}_{S'}(v)$ then $\mathit{outif}_{T'}(q, Z, f, x) \vee \mathit{thru}_{T'}(q, Z, f, x)$ follows from Condition (10) of $S\, \mathit{simil}_u\, T$. If not, then $x = v$ holds, and so $v \rightsquigarrow_S y$ for some $y \in \mathcal{N}_S(u)$; thus $\mathit{outif}_{T'}(q, Z, f, x) \vee \mathit{thru}_{T'}(q, Z, f, x)$ follows from Condition (12) and (13) of $S\, \mathit{simil}_u\, T$. Thus Condition (10) of $S'\, \mathit{cong}_v\, T'$ follows by interchanging S and T.
Hence we have shown $S'\, \mathit{simil}_v\, T'$ and $\mathcal{N}_{S'}(v) = \mathcal{N}_{T'}(v) \subseteq \mathcal{N}_S(u) = \mathcal{N}_T(u)$ for every $v \in \mathit{Inv}_S(u)$ and so $S'\, \mathit{simil}_{\mathit{Inv}_S(u)}\, T'$ and hence $S'\, \mathit{cong}_{\mathit{Inv}_S(u)}\, T'$ holds.
It is a similar exercise to show that $S\, \mathit{cong}_\omega\, T \Rightarrow S'\, \mathit{cong}_\omega\, T'$ holds. $\square$

Lemma 47 (schema body equivalence implies schema equivalence) *Let* $S = \mathit{while}\, p(\mathbf{x})\, \mathit{do}\, S'$ *and* $T = \mathit{while}\, p(\mathbf{x})\, \mathit{do}\, T'$ *be linear schemas. Let* $V \subseteq \mathcal{V}$ *and assume*

$$\mathit{Inv}_{S'}(V) \cup \mathit{Inv}_{T'}(V) \cup \mathit{Refset}_S(p) \subseteq V$$

holds. Assume that $S' \cong_V T'$*.*

(1) *Then* $S \cong_V T$ *holds.*
(2) *If* $\mathit{Inv}_{S'}(\omega) \cup \mathit{Inv}_{T'}(\omega) \subseteq V$ *and* $S' \cong_\omega T'$ *holds then* $S \cong_\omega T$ *holds.*

Proof.

(1) Let i be an interpretation and assume that $\mathcal{M}[\![S]\!]^i_e$ and $\mathcal{M}[\![T]\!]^i_e$ both terminate. Suppose that $<p = \mathsf{T}>$ occurs exactly n times in the path $\pi_S(i, e)$. By induction on r, and using Part (2) of Theorem 37 and the fact that $\mathit{Inv}_{S'}(V) \cup \mathit{Inv}_{T'}(V) \subseteq V$, $\mathcal{M}[\![(S')^r]\!]^i_e(v) = \mathcal{M}[\![(T')^r]\!]^i_e(v)$ for all $v \in V$ and $r \in \mathbb{N}$, and hence since $\mathit{Refset}_S(p) \subseteq V$, $<p = \mathsf{T}>$ occurs exactly n times in $\pi_T(i, e)$. Thus $\mathcal{M}[\![S]\!]^i_e = \mathcal{M}[\![(S')^{n+1}]\!]^i_e(v) = \mathcal{M}[\![(T')^{n+1}]\!]^i_e(v) = \mathcal{M}[\![T]\!]^i_e(v)$ for all $v \in V$.
(2) Similar to Part (1) of this Lemma.

$\square$

Lemma 48 *Let S be a linear schema and let $u \in \mathcal{V} \cup \{\omega\}$. Suppose T is another linear schema and $S \, cong_u \, T$ holds, and $Symbols(S) = Symbols(T) = \mathcal{N}_S(u) = \mathcal{N}_T(u)$, and $S = S_1 \dots S_m$ and $T = T_1 \dots T_n$ holds such that each S_j and T_j is either a while schema or an assignment. Then $m = n$ and there is a permutation χ of the set $\{1, \dots, m\}$ such that $Symbols(S_j) = Symbols(T_{\chi(j)})$, and if S_j is an assignment, then $T_{\chi(j)} = S_j$ for all j, and if any S_j is a while schema then so is $T_{\chi(j)}$, with the same guard; and $u \in \mathcal{V} \Rightarrow \chi(m) = m$ holds.*

Proof. The existence of χ linking subschemas of the same type and symbol set, implying $m = n$, follows straightforwardly from the congruence definition. To show $u \in \mathcal{V} \Rightarrow \chi(m) = m$, assume that $u \in \mathcal{V}$ and $\chi(m) = l < m$. Since the symbols of $T_{l+1} \dots T_m$ all lie in $\mathcal{N}_S(u)$, $T_{l+1} \dots T_m$ must contain an assignment $u := f(\mathbf{x})$ such that $f \rightsquigarrow_T u$. Similarly S_m (and hence T_l) contains an assignment $u := g(\mathbf{y})$ with $f \neq g$ such that $g \rightsquigarrow_S u$. Thus $u := f(\mathbf{x})$ occurs in $S_1 \dots S_{m-1}$. By Conditions (5), (10) and (11) of $S \, simil_u \, T$, there is some $q \in \mathit{ifPreds}(S) \cap \mathcal{N}_S(u)$ such that f and g are q-competing for u. But q would have to occur in S_m and $T_{l+1} \dots T_m$, clearly contradicting the other conditions on χ, giving a contradiction. $\square$

Lemma 49 (replacing similar schemas by congruent schemas) *Let $u \in \mathcal{V} \cup \{\omega\}$ and let S, T be u-similar linear schemas. Then there are linear schemas S', T' which are u-equivalent and u-similar to S and T respectively and such that $S' \, cong_u \, T'$.*

Proof. Assume that S and T are not already u-congruent. Then there exists $x \in \mathcal{N}_S(u)$ and $r \leq arity(x)$ such that the rth component of $\mathbf{refvec}_S(x)$ is v, whereas the rth component of $\mathbf{refvec}_T(x)$ is $w \neq v$; or a similar situation holds in which S and T differ on an assigned variable; this case can be treated analogously. We will proceed by replacing v and w in S and T respectively by a new variable v' to give schemas $\hat{S}$, $\hat{T}$ which are 'more congruent' than S and T are, but in order to do this, we must replace v and w not just in $\mathbf{refvec}_S(x)$ and $\mathbf{refvec}_T(x)$, but at other points in both schemas which are affected by the replacement in $\mathbf{refvec}_S(x)$ and $\mathbf{refvec}_T(x)$. Thus we define symbol sets

$$\Gamma_{assign}, \Gamma_{ref}(n) \subseteq \mathcal{N}_S(u)$$

for $n \in \mathbb{N}$ which give, respectively, the symbols at which the assigned variable and the nth referenced variable must be changed from v or w to v'. Γ_{assign} and $\Gamma_{ref}(n)$ are defined to be the minimal sets satisfying $x \in \Gamma_{ref}(r)$ and the following closure condition; if $g \rightsquigarrow_S y\,(r')$ for $y \in \mathcal{N}_S(u)$, $r' \leq arity(y)$ and $assign_S(g) = v$, then $g \in \Gamma_{assign} \iff y \in \Gamma_{ref}(r')$ holds. Let v' be a variable not occurring in either S or T. Let $\hat{S}$ be the schema obtained from S by setting $assign_{S'}(h) = v'$ for all $h \in \Gamma_{assign}$ and changing the sth component of $\mathbf{refvec}_S(z)$ from v to v' for all $z \in \Gamma_{ref}(s)$ and $s \leq arity(z)$. We similarly replace 'corresponding' occurrences of w in T by v' to get $\hat{T}$. $\hat{S} \, simil_u \, S$ follows from the conditions of Definition 34, since the relation $\rightsquigarrow_S$ is unaffected by this replacement, and similarly $\hat{S} \cong_u S$ follows by observing that for every interpretation i, the path $\pi_{\hat{S}}(i, e)$ can be obtained from $\pi_S(i, e)$ by replacing some occurrences of v by v'; and the same relations hold for T and $\hat{T}$. We repeat this process for every $x \in \mathcal{N}_S(u)$. The resulting schemas satisfy the conditions stated. $\Box$

Theorem 50 *Let $u \in \mathcal{V} \cup \{\omega\}$ and let S and T be u-similar linear schemas. Then S and T are u-equivalent.*

Proof. This follows by induction on $|S| + |T|$. By Lemma 49 we may assume that $S\,cong_u\,T$, and by Part (2) of Theorem 37, and the transitivity of u-congruence, we may assume that $Symbols(S) = Symbols(T) = \mathcal{N}_S(u) = \mathcal{N}_T(u)$, since S and T may be replaced by their slices containing only the symbols in $\mathcal{N}_S(u)$. Let i be an interpretation and assume that if $u \in \mathcal{V}$ then for each $U \in \{S, T\}$, $\pi_U(i, e)$ is finite. We will show that $\mathcal{M}[\![S]\!]^i_e(u) = \mathcal{M}[\![T]\!]^i_e(u)$. Let Q be the set of all $q \in ifPreds(S)$ for which $above_S(q) = q$. For each $q \in Q$ and $U \in \{S, T\}$, if $\pi_U(i, e)$ passes through q then the prefix of $\pi_U(i, e)$ preceding the sole occurrence of q contains precisely the same sequence of assignments as $\pi_{trunc_U(q)}(i, e)$. Thus by Lemma 44 and the inductive hypothesis applied to $trunc_S(q)$ and $trunc_T(q)$, for each $q \in Q$ there exists $Z(q) \in \{\mathsf{T}, \mathsf{F}\}$ such that neither $\pi_S(i, e)$ nor $\pi_T(i, e)$ passes through $<q = \neg Z(q)>$. Let S' be obtained from S by deleting the $Z(q)$-part of each $q \in Q$ and define T' similarly. For each $U \in \{S, T\}$, let $\mu(U) \in alphabet(U)^*$ be obtained from $\pi_U(i, e)$ by deleting all occurrences of $<q = Z(q)>$ for $q \in Q$. Thus each $\mu(U) = \pi_{U'}(i, e)$ and so if $Q \neq \emptyset$ then $\mathcal{M}[\![S]\!]^i_e(u) = \mathcal{M}[\![T]\!]^i_e(u)$ follows from Theorem 42 and the inductive hypothesis applied to S' and T'. Hence we may assume that $Q = \emptyset$.
Thus we may write $S = S_1 \ldots S_m$ and $T = T_1 \ldots T_n$ such that each S_j and each T_j is either a while schema or an assignment.
By Lemma 48, $m = n$ holds. We now consider two cases.

- Suppose $u \in \mathcal{V}$. By Lemma 48,

$$Symbols(S_m) = Symbols(T_m) \neq \emptyset$$

  holds. Let $V = Inv_{S_m}(u)$. Clearly $V = Inv_{T_m}(u)$, $S_m\,cong_u\,T_m$ and

$$S_1 \ldots S_{m-1}\,cong_V\,T_1 \ldots T_{m-1}$$

  can be proved using $S\,cong_u\,T$. Thus $S_1 \ldots S_{m-1} \cong_V T_1 \ldots T_{m-1}$ follows from the inductive hypothesis, and $S_m \cong_u T_m$ holds, using Lemma 46 and Part (1) of Lemma 47 if S_m and T_m are while schemas. Thus $S \cong_u T$ follows from Part (2) of Theorem 37.
- Suppose $u = \omega$. We may assume (after interchanging S and T if necessary) that for some $k \leq m$, the path $\pi_S(i, e)$ reaches S_k and fails to terminate in S_k. Thus it suffices to prove that $\pi_T(i, e)$ is infinite. By Lemma 48, there is a schema T_l such that

$$Symbols(S_k) = Symbols(T_l) \neq \emptyset$$

  holds. Let $V = Inv_{S_k}(\omega)$. As in Case (1), it is easy to show that $V = Inv_{T_l}(\omega)$, $S_k \cong_\omega T_l$, $S_1 \ldots S_{k-1}\,cong_V\,T_1 \ldots T_{l-1}$ and hence $S_1 \ldots S_{k-1} \cong_V T_1 \ldots T_{l-1}$ hold using

$S\,cong_u\,T$, the inductive hypothesis and Lemmas 46 and 47. Thus either the path $\pi_T(i,e)$ fails to reach T_l (and thus is infinite) or it reaches T_l and so $\mathcal{M}[\![T_1 \ldots T_l]\!]^i_e(\omega) = \perp$ follows from Part (2) of Theorem 37.

□

6 The main theorem and further directions

Our main result is the following.

Theorem 51 *Let S, T be LFL schemas. Then*

$$S \cong T \iff S\,simil\,T$$

holds. If $V \subseteq \mathcal{V} \cup \{\omega\}$ and $\omega \in V$ then

$$S \cong_V T \iff S\,simil_V\,T$$

holds. In particular, it is decidable in polynomial time whether S and T are equivalent.

Proof. The first assertion is a special case of the second (where V is the set containing all variables assigned in either S or T, plus ω). For any set $V \subseteq \mathcal{V} \cup \{\omega\}$, $S\,simil_V\,T \Rightarrow S \cong_V T$ follows from Theorem 50. The proof of the converse statement for sets V containing ω is given as part of [7, Theorem 148]. The polynomial time bound follows from Theorem 35. □

An overview of the full proof of Theorem 51 is given in [7, Section 1.2].

Of the various related problems which seem worth studying (besides the 'missing'result $S \cong_v T \Rightarrow S\,simil_v\,T$ for $v \in \mathcal{V}$, which we have failed to prove), two strike us as being particularly promising.

6.1 Computing minimal slices of schemas

For the purpose of program slicing, given a schema S and variable u, it is of interest to be able to compute those minimal slices of S (with minimality defined by symbol sets) which are v-equivalent to S and which preserve termination. By Part (1) of Theorem 37 and [7, Theorem 76], it follows that for any $u \in \mathcal{V}$, the minimal slice T of an LFL schema S such that $S \cong_u T$ and $\mathcal{M}[\![S]\!]^j_d \neq \perp \Rightarrow \mathcal{M}[\![T]\!]^j_d \neq \perp$ always holds is precisely the slice of S such that $Symbols(T) = \mathcal{N}_S(u)$ holds. The first author has proved in [37] that this also holds if the linearity hypothesis is replaced by function-linearity

$$
\begin{array}{l}
\textit{while}\; q(v)\; \textit{do} \\
\qquad \{ \\
\qquad v := k(v); \\
\qquad w := h(w); \\
\qquad \textit{if}\; p(w) \quad \textit{then} \\
\qquad\qquad\qquad \{ \\
\qquad\qquad\qquad u := g(); \\
\qquad\qquad\qquad w := f(w); \\
\qquad\qquad\qquad \} \\
\qquad\qquad \textit{else} \quad \Lambda \\
\qquad \}
\end{array}
$$

Fig. 6. Deleting the f-assignment gives a u-equivalent slice of this schema

(a schema is function-linear if it does not contain more than one occurrence of the same function symbol), provided that the definition of $\mathcal{N}_S(u)$ is generalised to allow for multiple occurrences of predicate symbols.

If S is merely free and linear then $S \cong_u T$ need not imply $\mathit{Symbols}(T) \supseteq \mathcal{N}_S(u)$, as the example of Figure 6 shows. Owing to the constant g-assignment, S is not liberal, though it is free. Clearly $f \in \mathcal{N}_S(u)$ holds, but the slice of S obtained by deleting the f-assignment, which is also free, is u-equivalent to S. It is also ω-equivalent to S, and hence satisfies the termination requirement for slices.

It would be of interest to find a method of computing the minimal slice of S satisfying these conditions under weaker hypotheses than the assumption that S is liberal, free and function-linear.

6.2 Using schema transformations to construct equivalent schemas

Given a linear schema S and $u \in \mathcal{V} \cup \{\omega\}$, it can be shown using Theorem 50 that the following transformations of S preserve u-equivalence.

- Changing the variables mentioned in S in any way that preserves u-similarity.
- Replacing S by a slice T of S, such that T contains every element of $\mathcal{N}_S(u)$.
- Pulling out a subschema from an if subschema of S; that is, replacing a subschema

if $p(\mathbf{v})$ *then* $S_1 S_2$

else S_3

of S by the schema

S_1

if $p(\mathbf{v})$ *then* S_2

else S_3

provided that this does not create a new fx-segment μ for $f \in \mathit{Funcs}(S_1)$ and $x \in \mathcal{N}_S(u) \cup \{u\}$ such that either $x = p$ or μ passes through $<\!p = \mathsf{F}\!>$. Also, if $u = \omega$ then S_1 must not contain a while predicate, otherwise Condition (7) of simil_u is violated. Clearly the true and false parts of p may be interchanged.

- Changing the order of 'towers' of if predicates; that is, interchanging $p(\mathbf{u})$ and $q(\mathbf{v})$ in a subschema

if $p(\mathbf{u})$ *then*

{

if $q(\mathbf{v})$ *then* T

else Λ

}

else Λ

of S. Again, the true and false parts of p or q may be interchanged.

- Replacing a subschema $S_1 S_2$ of S by $S_2 S_1$ to give a schema T, provided that no variable is assigned in both S_1 and S_2, and $S_1 S_2$ contains no fx-segment with $f \in \mathit{Funcs}(S_1)$ and $x \in \mathit{Symbols}(S_2)$, and the same statement holds with (S, 1, 2) replaced by (T, 2, 1).

We conjecture that given any LFL schema S and $u \in \mathcal{V} \cup \{\omega\}$, all u-similar LFL schemas can be obtained from S by a sequence of these transformations and their inverses.

It may also be possible to prove that given an LFL schema S, any u-equivalent LFL schema may be reached from S by a finite sequence of such transformations without using Theorem 51, thus giving an alternative (and possibly shorter) way of proving this theorem than the one we have given in the Technical Report [7].

References

[1] S. Greibach, Theory of program structures: schemes, semantics, verification, Vol. 36 of Lecture Notes in Computer Science, Springer-Verlag Inc., New York, NY, USA, 1975.

[2] Z. Manna, Mathematical Theory of Computation, McGraw–Hill, 1974.

[3] A. K. Chandra, On the decision problems of program schemas with commutative and invertable functions, in: Conference Record of the ACM Symposium on Principles of Programming Languages, ACM SIGACT-SIGPLAN, Boston, Massachusetts, 1973, pp. 235–242.

[4] A. K. Chandra, Z. Manna, Program schemas with equality, in: Conference Record, Fourth Annual ACM Symposium on Theory of Computing, Denver, Colorado, 1972, pp. 52–64.

[5] A. K. Chandra, Z. Manna, On the power of programming features, Computer Languages 1 (3) (1975) 219–232.

[6] M. R. Laurence, S. Danicic, M. Harman, R. M. Hierons, J. D. Howroyd, Equivalence of conservative, free, linear schemas is decidable, Theoretical Computer Science 290 (1) (2002) 831–862.

[7] M. R. Laurence, S. Danicic, M. Harman, R. Hierons, J. Howroyd, Equivalence of linear, free, liberal, structured program schemas is decidable in polynomial time, Tech. Rep. ULCS-04-014, University of Liverpool, electronically available at `http://www.csc.liv.ac.uk/research/techreports/` (2004).

[8] M. S. Paterson, Equivalence problems in a model of computation, PhD thesis, University of Cambridge, UK (1967).

[9] E. Ashcroft, Z. Manna, Translating program schemas to while-schemas, SIAM Journal on Computing 4 (2) (1975) 125–146.

[10] Y. I. Ianov, The logical schemes of algorithms, in: Problems of Cybernetics, Vol. 1, Pergamon Press, New York, 1960, pp. 82–140.

[11] V. K. Sabelfeld, An algorithm for deciding functional equivalence in a new class of program schemes, Journal of Theoretical Computer Science 71 (1990) 265–279.

[12] A. De Lucia, A. R. Fasolino, M. Munro, Understanding function behaviours through program slicing, in: 4^{th} IEEE Workshop on Program Comprehension, IEEE Computer Society Press, Los Alamitos, California, USA, Berlin, Germany, 1996, pp. 9–18.

[13] G. Canfora, A. Cimitile, A. De Lucia, G. A. D. Lucca, Software salvaging based on conditions, in: International Conference on Software Maintenance (ICSM'96), IEEE Computer Society Press, Los Alamitos, California, USA, Victoria, Canada, 1994, pp. 424–433.

[14] A. Cimitile, A. De Lucia, M. Munro, A specification driven slicing process for identifying reusable functions, Software maintenance: Research and Practice 8 (1996) 145–178.

[15] K. B. Gallagher, J. R. Lyle, Using program slicing in software maintenance, IEEE Transactions on Software Engineering 17 (8) (1991) 751–761.

[16] K. B. Gallagher, Evaluating the surgeon's assistant: Results of a pilot study, in: Proceedings of the International Conference on Software Maintenance, IEEE Computer Society Press, Los Alamitos, California, USA, 1992, pp. 236–244.

[17] M. Weiser, J. R. Lyle, Experiments on slicing–based debugging aids, Empirical studies of programmers, Soloway and Iyengar (eds.), Molex, 1985, Ch. 12, pp. 187–197.

[18] H. Agrawal, R. A. DeMillo, E. H. Spafford, Debugging with dynamic slicing and backtracking, Software – Practice and Experience 23 (6) (1993) 589–616.

[19] M. Kamkar, Interprocedural dynamic slicing with applications to debugging and testing, PhD Thesis, Department of Computer Science and Information Science, Linköping University, Sweden, available as Linköping Studies in Science and Technology, Dissertations, Number 297 (1993).

[20] J. R. Lyle, M. Weiser, Automatic program bug location by program slicing, in: 2^{nd} International Conference on Computers and Applications, IEEE Computer Society Press, Los Alamitos, California, USA, Peking, 1987, pp. 877–882.

[21] D. W. Binkley, The application of program slicing to regression testing, in: M. Harman, K. Gallagher (Eds.), Information and Software Technology Special Issue on Program Slicing, Vol. 40, Elsevier, 1998, pp. 583–594.

[22] R. Gupta, M. J. Harrold, M. L. Soffa, An approach to regression testing using slicing, in: Proceedings of the IEEE Conference on Software Maintenance, IEEE Computer Society Press, Los Alamitos, California, USA, Orlando, Florida, USA, 1992, pp. 299–308.

[23] M. Harman, S. Danicic, Using program slicing to simplify testing, in: 2^{nd} EuroSTAR conference on Software Testing Analysis and Review, Brussels, 1994.

[24] G. Canfora, A. Cimitile, M. Munro, RE^2: Reverse engineering and reuse re-engineering, Journal of Software Maintenance : Research and Practice 6 (2) (1994) 53–72.

[25] D. Simpson, S. H. Valentine, R. Mitchell, L. Liu, R. Ellis, Recoup – Maintaining Fortran, ACM Fortran forum 12 (3) (1993) 26–32.

[26] J. Beck, D. Eichmann, Program and interface slicing for reverse engineering, in: IEEE/ACM 15^{th} Conference on Software Engineering (ICSE'93), IEEE Computer Society Press, Los Alamitos, California, USA, 1993, pp. 509–518.

[27] A. Cimitile, A. De Lucia, M. Munro, Identifying reusable functions using specification driven program slicing: a case study, in: Proceedings of the IEEE International Conference on Software Maintenance (ICSM'95), IEEE Computer Society Press, Los Alamitos, California, USA, Nice, France, 1995, pp. 124–133.

[28] S. Horwitz, J. Prins, T. Reps, Integrating non–interfering versions of programs, ACM Transactions on Programming Languages and Systems 11 (3) (1989) 345–387.

[29] J. M. Bieman, L. M. Ott, Measuring functional cohesion, IEEE Transactions on Software Engineering 20 (8) (1994) 644–657.

[30] J. J. Thuss, An investigation into slice–based cohesion metrics, Master's thesis, Michigan Technological University (1988).

[31] A. Lakhotia, Rule–based approach to computing module cohesion, in: Proceedings of the 15^{th} Conference on Software Engineering (ICSE-15), 1993, pp. 34–44.

[32] D. W. Binkley, K. B. Gallagher, Program slicing, in: M. Zelkowitz (Ed.), Advances of Computing, Volume 43, Academic Press, 1996, pp. 1–50.

[33] A. De Lucia, Program slicing: methods and applications, in: 1st IEEE International workshop on Source Code Analysis and Manipulation, IEEE Computer Society Press, Los Alamitos, California, USA, Florence, Italy, 2001, pp 142-149.

[34] F. Tip, A survey of program slicing techniques, Tech. Rep. CS-R9438, Centrum voor Wiskunde en Informatica, Amsterdam (1994).

[35] M. Weiser, Program slices: Formal, psychological, and practical investigations of an automatic program abstraction method, PhD thesis, University of Michigan, Ann Arbor, MI (1979).

[36] S. Danicic, M. Harman, R. M. Hierons, J. howroyd, M. R. Laurence, Static program slicing algorithms are minimal for free liberal program schemas, The Computer Journal 48 (6) (2006) 737–748.

[37] M. R. Laurence, Characterising minimal semantics-preserving slices of function-linear, free, liberal program schemas, submitted to Journal of Logic and Algebraic Programming in July 2005 .