# A Program Refinement Framework Supporting Reasoning about Knowledge and Time

Kai Engelhardt, Ron van der Meyden, and Yoram Moses

[1] School of Computing Sciences
University of Technology, Sydney, Australia
[ke|ron]@socs.uts.edu.au
[2] Department of Electrical Engineering
Technion, Haifa, Israel
moses@ee.technion.ac.il

**Abstract.** This paper develops a highly expressive semantic framework for program refinement that supports both temporal reasoning and reasoning about the knowledge of a single agent. The framework generalizes a previously developed temporal refinement framework by amalgamating it with a logic of quantified local propositions, a generalization of the logic of knowledge. The combined framework provides a formal setting for development of knowledge-based programs, and addresses two problems of existing theories of such programs: lack of compositionality and the fact that such programs often have only implementations of high computational complexity. Use of the framework is illustrated by a control theoretic example concerning a robot operating with an imprecise position sensor.

## 1 Introduction

The *knowledge-based* approach to the design and analysis of distributed systems, introduced by Halpern and Moses [6] involves the use of modal logics of knowledge. One of the key contributions of this approach is the notion of *knowledge-based programs* [5, 4], which generalize standard programs by allowing the tests in conditional constructs to be formulas in the logic of knowledge. Such programs contain statements of the form "if you know that $X$ then do $A$ else $B$". This provides a high level abstraction of distributed programs that allows for perspicuous descriptions of how an agent's actions are related to its state of information (which, in a distributed system, is typically incomplete) about its environment.

In its current state of development, the knowledge-based approach has a number of limitations, among them that:

1. The formal methodology for developing and reasoning about knowledge-based programs is at present only weakly developed.
2. The existing semantics for knowledge-based programs is based on a particular interpretation of knowledge that requires a complete description of the implementing program. This prevents the compositional development of program fragments.
3. Knowledge-based programs often have only implementations of unacceptably high computational complexity.

This paper is a step in the direction of the formulation of the knowledge-based approach that addresses these limitations.

One of the starting points for our work is the observation that knowledge-based programs are in one respect more like specifications than like standard programs. They cannot be directly executed — instead, their meaning is defined by a relation of "implementation" between knowledge based programs and standard programs: a given knowledge-based program may have no, one, or many concrete programs as its implementations. As a specification formalism, however, knowledge-based programs are unbalanced, abstracting only the tests performed by agents, but providing no abstraction mechanism for their actions [11].

Action abstraction is handled much better in *refinement calculi* [1, 9, 10], also known as "broad spectrum" languages. Such calculi view programs and specifications as having the same semantic type, and support a formal methodology for the development of programs that are "correct by design", where one begins with a specification and transforms it to an implementation by means of a sequence of correctness preserving refinement steps. The focus in this area has been on sequential programs and atemporal assertions but recently some approaches to refinement admitting the expressive power of temporal logics have been developed [13, 7].

A first step in the direction of a refinement calculus suited to the knowledge-based development of programs was taken in van der Meyden and Moses [17, 15], where it is shown how to develop a refinement approach capturing certain types of temporal reasoning that will be critical in knowledge-based program development. We further develop these ideas in the present paper, by showing how they may be extended to accommodate knowledge-based reasoning. Significantly, the framework we define admits compositional program development.

In developing the extension, we also seek to address the final limitation of knowledge-based programs alluded to above. To implement the statement "if you know that $X$ then do $A$ else $B$", a concrete program must do $A$ exactly when it is in a local state (captured by the values of the variables and storage it maintains locally) that carries the information that $X$ is true. The difficulty with this is that computing whether a local state bears the information that $X$ may have very high computational complexity [12, 14, 18]. As argued by Engelhardt, van der Meyden and Moses [3], in practice, it may often be sufficient to use conditions on the agent's state of information that are sound, but not complete, tests of its knowledge. Such tests may be expressed in the *Logic of Local Propositions* (LLP) [3].

The present paper integrates the temporal refinement framework of van der Meyden and Moses [15] with the logic of local propositions. Although our ultimate aim is a framework for the development of distributed systems, we deal in this paper with a single agent operating synchronously with its environment: asynchrony and multiple agents introduce complexities that we plan to address in the future. The main novelty is the introduction of a programming/specification construct that resembles a quantification over local propositions. This construct makes it possible to write specifications stating that the agent conditions its behaviour on a *local* test for some property of interest, without stating explicitly what test is used. The introduction of this construct necessitates an adaptation of the semantics of the temporal refinement of [15].

The paper is structured as follows. Section 2 defines an assertion language that adapts the LLP semantics to the richer temporal setting required for reasoning about programs. Section 3 defines the syntax and semantics of our broad spectrum programming and specification language that incorporates the assertion language from Sect. 2. Section 4 defines the semantic refinement relation we use for this class of programs and develops a number of refinement rules valid for this relation. Section 5 illustrates the use of the framework by presenting a formal development of a control theoretic example previously treated informally in the literature on knowledge-based programs.

## 2 A Semantics for Reasoning about Knowledge and Time

We begin by presenting a semantic framework for a single agent and its environment, inspired by [4], to which we refer the reader for motivation.

Let $L_e$ be a set of possible states for the environment and let $L_1$ be a set of possible local states for agent 1. We take $\mathcal{G} = L_e \times L_1$ to be the set of *global states*. Let $A_1$ and $A_e$ be nonvoid sets of *actions* for agent 1 and for the environment, respectively. (These sets usually contain a special *null action* $\Lambda$.) A *joint action* is a pair $(a_e, a_1) \in \mathcal{A} = A_e \times A_1$. A *run* over $\mathcal{G}$ and $\mathcal{A}$ is a pair $r = (h, \alpha)$ of infinite sequences: a *state history* $h : \mathbb{N} \longrightarrow \mathcal{G}$, and an *action history* $\alpha : \mathbb{N} \longrightarrow \mathcal{A}$. Intuitively, for $c \in \mathbb{N}$, $h(c)$ is the global state of the system at time $c$ and $\alpha(c)$ is the joint action occurring at time $c$. (We say more about the transition relation connecting states and actions later.) A *system* over $\mathcal{G}$ and $\mathcal{A}$ is a set of runs over $\mathcal{G}$ and $\mathcal{A}$, intuitively representing all possible histories. A pair $(r, c)$ consisting of a run $r$ (in system $S$) and a time $c \in \mathbb{N}$ is called a *point (in S)*. We write $\mathrm{Points}(S)$ for the set of points of $S$. Let *Prop* be a set of propositional variables. An *interpretation* of a system $S$ is a mapping $\pi : \mathit{Prop} \longrightarrow 2^{\mathrm{Points}(S)}$ associating a set of points with each propositional variable. Intuitively, proposition $p \in \mathit{Prop}$ is true exactly at the points contained in $\pi(p)$. An *interpreted system (over $\mathcal{G}$ and $\mathcal{A}$)* is a pair $\mathfrak{I} = (S, \pi)$ where $S$ is a system over $\mathcal{G}$ and $\mathcal{A}$ and $\pi$ is an interpretation of $S$.

The structure in the above definitions supports the following notions used to define the agent's knowledge. We say two points $(r,c), (r',c')$ in a system $S$ are 1-*indistinguishable*, denoted $(r,c) \sim_1 (r',c')$, if the local components of the global states at these points are equal, i.e., if there exists a local state $s_1 \in L_1$ and states of the environment $s_e, s_e'$ such that $h(c) = (s_e, s_1)$ and $h'(c') = (s_e', s_1)$, where $r = (h, \alpha)$ and $r' = (h', \alpha')$. A set $P$ of points of $S$ is 1-*local* if it is closed under $\sim_1$, in other words, when for all points $(r,c), (r',c')$ of $S$, if $(r,c) \in P$ and $(r,c) \sim_1 (r',c')$ then $(r',c') \in P$. Intuitively, 1-local sets of points correspond to properties that the agent is able to determine entirely on the basis of its local state. If $\pi$ and $\pi'$ are interpretations and $p \in \mathit{Prop}$, then $\pi'$ is said to be a 1-*local p-variant* of $\pi$, denoted $\pi \simeq_p^1 \pi'$, if $\pi$ and $\pi'$ differ at most in the value of $p$ and $\pi'(p)$ is 1-local. If $\mathfrak{I} = (S, \pi)$ and $\mathfrak{I} = (S', \pi')$ are two interpreted systems over $\mathcal{G}$ and $\mathcal{A}$, then $\mathfrak{I}'$ is said to be 1-*local p-variant* of $\mathfrak{I}$, denoted $\mathfrak{I} \simeq_p^1 \mathfrak{I}'$, if $S = S'$ and $\pi \simeq_p^1 \pi'$.

The logical language $\mathcal{L}$ we use in this paper resembles a restricted monadic second order logic with two additions: (a) an S5-modality for necessity and (b) operators from the linear time temporal logic LTL [8]. Its syntax is given by:

$$\mathcal{L} \ni \phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid \mathsf{Nec}\,\phi \mid \forall_1 p\,(\phi) \mid \bigcirc\phi \mid \phi\,\mathsf{U}\,\phi \mid \ominus\phi \mid \phi\,\mathsf{S}\,\phi$$

where $p \in Prop$. Intuitively, $\mathsf{Nec}\,\phi$ says that $\phi$ is true at all points in the interpreted system, and its dual $\mathsf{Poss}\,\phi = \neg\,\mathsf{Nec}\,\neg\phi$ states that $\phi$ is true at some point. The formula $\forall_1 p\,(\phi)$ says that $\phi$ is true for all assignments of a 1-local proposition (set of points) to the propositional variable $p$. We write $\exists_1 p\,(\phi)$ for its dual $\neg\forall_1 p\,(\neg\phi)$. The remaining connectives have their standard interpretations from linear time temporal logic: $\bigcirc$ ("next"), $\mathsf{U}$ ("until"), $\ominus$ ("previously") and $\mathsf{S}$ ("since"). We employ parenthesis to indicate aggregation and use standard abbreviations such as *true*, *false*, $\vee$, and definable future time operators like $\Box$ ("henceforth") and $\Diamond$ ("eventually"), as well as their past time counterparts $\boxminus$ ("until now") and $\diamondminus$ ("once").

Formulae of $\mathcal{L}$ are interpreted at a point $(r,c)$ of an interpreted system $\mathfrak{I} = (S,\pi)$ by means of the satisfaction relation $\models$, defined inductively by:

- $\mathfrak{I},(r,c) \models p$ iff $(r,c) \in \pi(p)$.
- $\mathfrak{I},(r,c) \models \neg\phi$ iff $\mathfrak{I},(r,c) \not\models \phi$.
- $\mathfrak{I},(r,c) \models \phi\wedge\psi$ iff $\mathfrak{I},(r,c) \models \phi$ and $\mathfrak{I},(r,c) \models \psi$.
- $\mathfrak{I},(r,c) \models \mathsf{Nec}\,\phi$ iff $\mathfrak{I},(r',c') \models \phi$, for all $(r',c') \in \mathrm{Points}(S)$.
- $\mathfrak{I},(r,c) \models \forall_1 p\,(\phi)$ iff $\mathfrak{I}',(r,c) \models \phi$ for all $\mathfrak{I}'$ such that $\mathfrak{I} \simeq^1_p \mathfrak{I}'$
- $\mathfrak{I},(r,c) \models \bigcirc\phi$ iff $\mathfrak{I},(r,c+1) \models \phi$
- $\mathfrak{I},(r,c) \models \phi\,\mathsf{U}\,\psi$ iff there exists a $d \geq c$ such that $\mathfrak{I},(r,d) \models \psi$ and $\mathfrak{I},(r,e) \models \phi$ for all $e$ with $c \leq e < d$.
- $\mathfrak{I},(r,c) \models \ominus\phi$ iff $c > 0$ and $\mathfrak{I},(r,c-1) \models \phi$
- $\mathfrak{I},(r,c) \models \phi\,\mathsf{S}\,\psi$ iff there exists a $d \leq c$ such that $\mathfrak{I},(r,d) \models \psi$ and $\mathfrak{I},(r,e) \models \phi$ for all $e$ with $d < e \leq c$.

Given these constructs, it is possible to express many operators from the literature on reasoning about knowledge. For example, consider the standard knowledge operator $K_1$, defined by $\mathfrak{I},(r,c) \models K_1\phi$ if $\mathfrak{I},(r',c') \models \phi$ for all $(r',c') \in \mathrm{Points}(S)$ such that $(r,c) \sim_1 (r',c')$. This is expressible as $\exists_1 p\,(p \wedge \mathsf{Nec}(p \to \phi))$. We refer to [3] for further examples and discussion.

## 3 Sequential Programs with Quantification over Local Propositions

In this section we define our wide spectrum programming language, and discuss its semantics. We also define a refinement relation on programs.

### 3.1 Syntax

The programming language describes the structure of segments of runs. Let *CV* be a set of *constraint variables* and *PV* a set of *program variables*. Define the syntactic category *Prg* of *programs* by

$$Prg \ni P ::= \varepsilon \mid Z \mid a \mid P * P \mid P + P \mid P^{\omega} \mid \exists_1 p\,(P) \mid [\phi,\psi]^X \mid [\phi]^X \mid \{\phi\}_C$$

where $Z \in PV$, $a \in A_1$, $p \in Prop$, $\phi,\psi \in \mathcal{L}$, $X \in CV$, and $C \subseteq CV$. The intuitive meaning of these constructs is as follows. The symbol $\varepsilon$ denotes the *empty program*, which takes no time to execute, and has no effects. Program variables $Z$ are placeholders used to allow substitution of programs. Note that a program may refer directly to actions $a$ of the agent, but the actions of the environment are left implicit. The operation $*$ represents *sequential composition*. The symbol $+$ denotes nondeterministic choice, while $P^{\omega}$ denotes zero or more (possibly infinitely many) repetitions of $P$. The construct $\exists_1 p\,(P)$ can also be understood as a kind of nondeterministic choice: it states that $P$ runs with respect to some assignment of a 1-local proposition to the propositional variable $p$. The last three constructs are like certain constructs found in refinement calculi. Intuitively, the *specification* $[\phi,\psi]^X$ states that some program runs in this location that has the property that, if started at a point satisfying $\phi$, eventually terminates at a point satisfying $\psi$.[1] The *coercion* $[\phi]^X$ is a program that takes no time to execute, but expresses a constraint on the surrounding program context: this must guarantee that $\phi$ holds at this location. The constraint variable $X$ in specifications and coercions acts as a label that allows references by other pieces of program text. Specifically, this is done in the *assertions* $\{\phi\}_C$, which act like program annotations: such a statement takes no time to execute, and, intuitively, asserts that $\phi$ can be proved to hold at this program location, with the proof depending only on concrete program fragments and on specification and coercion statements whose labels are in $C$. We may omit the constraint variables when it is not necessary to make such references.

In programs we employ parentheses to indicate aggregation wherever necessary and tend to omit $*$ near coercions and assertions. Moreover, we use the following abbreviations.

$$\textbf{if}^X\ \phi\ \textbf{then}\ P\ \textbf{else}\ Q\ \textbf{fi} \stackrel{\text{def}}{=} ([\phi]^X P) + ([\neg\phi]^X Q)$$
$$\textbf{while}^X\ \phi\ \textbf{do}\ P\ \textbf{od} \stackrel{\text{def}}{=} ([\phi]^X P)^{\omega}\,[\neg\phi]^X$$

### 3.2 Semantics

We note that the semantics presented in this section treats assertions $\{\phi\}_C$ as equivalent to the null program $\varepsilon$ — the role of assertions in the framework will be explained later. Our semantics will treat programs like specifications of certain sets of run segments in a system, intuitively, the sets of run segments that can be viewed as having been generated by executing the program. We first define execution trees, which represent unfoldings of the nondeterminism in a program.

It is convenient to represent these trees as follows. A *binary tree domain* is a prefix-closed subset of the set $\{0,1\}^* \cup \{0,1\}^{\omega}$. So, each nonvoid tree domain contains the empty sequence $\lambda$. Let $A$ be a set. An *A-labelled binary tree* is a function $T$ from a binary tree domain $D$ to $A$. The *nodes* of $T$ are the elements of $D$. The node $\lambda$ is called the *root* of $T$. If $n \in D$ we call $T(n)$ the *label* at node $n$. If $n \in D$ then the *children of n in T* are the nodes of $T$ (if any) of the form $n \cdot i$ where $i \in \{0,1\}$. Finite maxima in the prefix order on $D$ are called *leaves* of $T$.

[1] In refinement calculi, such statements are typically associated with *frame variables*, representing the variables allowed to change during the execution - we could add these, but omit them for brevity.

An *execution tree* is a *Prg*-labelled binary tree, subject to the following constraints on the nodes $n$:

1. If $n$ is labelled by $\varepsilon$, a program variable $Z \in PV$, a basic action $a$, a specification $[\phi,\psi]^X$, a coercion $[\phi]^X$, or an assertion $\{\phi\}_C$, then $n$ is a leaf.
2. If $n$ is labelled by $\exists_1 p\,(P)$ then $n$ has exactly one child $n\cdot 0$, labelled by $P$.
3. If $n$ is labelled by $P * Q$ or $P + Q$ then $n$ has exactly two children $n\cdot 0, n\cdot 1$, labelled by $P$ and $Q$ respectively.
4. If $n$ is labelled by $P^\omega$ then $n$ has exactly two children, $n\cdot 0, n\cdot 1$, labelled by $\varepsilon$ and $P*(P^\omega)$, respectively.

With each program $P$ we associate a particular execution tree, $T_P$, namely the unique execution tree labelled with $P$ at the root $\lambda$.

We now define the semantic constructs specified by programs. An *interval over a system S* is a triple $r[c,d]$ consisting of a run $r$ of $S$ and two elements $c$ and $d$ of $\mathbb{N}_+ = \mathbb{N} \cup \{\infty\}$ such that $c \le d$. We say that the interval is *finite* if $d < \infty$. A set $I$ of intervals is *run-unique* if $r[c,d], r[c',d'] \in I$ implies $c = c'$ and $d = d'$. An *interpreted interval set over S* (or *iis* for short) is a pair $(\pi, I)$ consisting of an interpretation $\pi$ of $S$ and a run-unique set $I$ of intervals over $S$.

We will view programs as specifying, or executing over, interpreted interval sets, by means of certain mappings from execution trees to interpreted interval sets. To facilitate the definition in the case of sequential composition, we introduce a shorthand for the two sets obtained by splitting each interval in a given set $I$ of intervals of $S$ in two. Say that $f : I \longrightarrow \mathbb{N}_+$ *divides I* whenever $c \le f(r[c,d]) \le d$ holds for all $r[c,d] \in I$. Given some $f$ dividing $I$, we write $f_{\blacktriangleleft}(I)$ for the set of intervals $r[f(r[c,d]),d]$ such that $r[c,d] \in I$. Analogously, we write $f_{\blacktriangleright}(I)$ for $\{\, r[c, f(r[c,d])] \mid r[c,d] \in I \,\}$.

Let $S$ be a system, let $(\pi, I)$ be an iis w.r.t. $S$, and let $P$ be a program. A function $\theta$ mapping each node $n$ of $T_P$ to an iis $(\pi_\theta(n), I_\theta(n))$, respectively, is an *embedding* of $T_P$ in $(\pi, I)$ w.r.t. $S$ whenever the following conditions are satisfied:

1. $\theta(\lambda) = (\pi, I)$.
2. If $n$ is labelled $\varepsilon$ or $\{\phi\}_C$, then $c = d$ for all $r[c,d] \in I_\theta(n)$.
3. If $n$ is labelled $a$ then, for all $(h,\alpha)[c,d] \in I_\theta(n)$, if $c < \infty$ then both $d = 1 + c$ and $a = a_1$, where $\alpha(c) = (a_e, a_1)$.
4. If $n$ is labelled $[\phi,\psi]$, then, for all $r[c,d] \in I_\theta(n)$, whenever $c < \infty$ and $(S, \pi_\theta(n)), (r,c) \models \phi$, then both $d < \infty$ and $(S, \pi_\theta(n)), (r,d) \models \psi$.
5. If $n$ is labelled $[\phi]$, then $c < \infty$ implies that $c = d$ and $(S, \pi_\theta(n)), (r,c) \models \phi$, for all $r[c,d] \in I_\theta(n)$.
6. If $n$ is labelled $\exists_1 p\,(Q)$ then $\pi_\theta(n) \simeq^1_p \pi_\theta(n\cdot 0)$ and $I_\theta(n\cdot 0) = I_\theta(n)$.
7. If $n$ is labelled $Q_1 + Q_2$, then $\pi_\theta(n\cdot 0) = \pi_\theta(n\cdot 1) = \pi_\theta(n)$ and $I_\theta(n)$ is the disjoint union of $I_\theta(n\cdot 0)$ and $I_\theta(n\cdot 1)$.
8. If $n$ is labelled $Q_1 * Q_2$, then $\pi_\theta(n\cdot 0) = \pi_\theta(n\cdot 1) = \pi_\theta(n)$ and there is an $f$ dividing $I_\theta(n)$ such that $I_\theta(n\cdot 0) = f_{\blacktriangleright}(I_\theta(n))$ and $I_\theta(n\cdot 1) = f_{\blacktriangleleft}(I_\theta(n))$.
9. If $n$ is labelled $Q^\omega$ then $\pi_\theta(n\cdot 0) = \pi_\theta(n\cdot 1) = \pi_\theta(n)$ and $I_\theta(n)$ is the disjoint union of $I_\theta(n\cdot 0)$ and $I_\theta(n\cdot 1)$ (as in case 7) and, for all $r[c,d] \in I_\theta(n)$:

$$d = \bigsqcup \left\{\, d' \;\middle|\; r[c',d'] \in I_\theta(n\cdot m) \;\text{ for some leaf } n\cdot m \text{ of } T_P \text{ below } n \,\right\} \quad .$$

We write $S,(\pi,I) \Vdash_\theta P$ whenever $\theta$ is an embedding of $T_P$ in $(\pi,I)$ w.r.t. $S$. Say that $P$ *occurs over* $(\pi,I)$ *w.r.t.* $S$ if there exists a $\theta$ such that $S,(\pi,I) \Vdash_\theta P$.

Clauses 1 to 8 formalize the intuitive understanding given above for each of the program constructs. Concerning clause 9 of this definition, we remark that, by run-uniqueness and the other clauses, if $n \cdot m_0, n \cdot m_1 \ldots$ are the leaves $n \cdot m$ below $n$ for which $I_\theta(n \cdot m)$ contains an interval on $r$, in left to right order, and these intervals are $r[c_0,d_0], r[c_1,d_1], \ldots$, respectively, then we have $d_i = c_{i+1}$ for each index $i$ in the sequence. (We may have $c_i = d_i$.) If $d$ were not the least upper bound $d'$ of the $d_i$, then this sequence of intervals would amount to an execution of $Q^\omega$ over $r[c,d']$ rather than over $r[c,d]$.

### 3.3 Refinement

The semantics just presented can be shown to be a generalization of the semantics of [15] for a similar language without the local propositional quantifier. That semantics, however, dealt with *single* intervals where we have used a set of intervals. The motivation for the change is that certain undesirable refinement rules involving the local propositional quantifier would be valid under the earlier semantic approach. We now present two definitions of refinement and an example that motivates the richer semantics.

Intuitively, a program $P$ refines $Q$ if, whenever $P$ executes, so does $Q$. A refinement relation of this type, when transitive and preserved under program composition, allows us to start with a high level specification and derive a concrete implementation through a sequence of refinement steps.

One refinement relation definable using our semantics as is follows: $P$ refines $Q$, denoted $P \sqsubseteq Q$ when for all systems $S$, and interpreted interval sets $(\pi,I)$ over $S$, if $S,(\pi,I) \Vdash P$ then $S,(\pi,I) \Vdash P$. For the semantics using single intervals, the corresponding relation would be defined by $P \sqsubseteq_{\text{run}} Q$ when for all systems $S$, interpretations $\pi$ and intervals $r[c,d]$ of $S$, if $S,(\pi,\{r[c,d]\}) \Vdash P$ then $S,(\pi,\{r[c,d]\}) \Vdash P$. Clearly, if $P \sqsubseteq Q$ then $P \sqsubseteq_{\text{run}} Q$. As the following example demonstrates, the converse is false.

*Example 1.* Let $\phi \in \mathcal{L}$ be any formula and consider the following two programs.

$$P = \textbf{if}\ \phi\ \textbf{then}\ a\ \textbf{else}\ a * a\ \textbf{fi}$$
$$Q = \exists_1 p\,(\textbf{if}\ p\ \textbf{then}\ a\ \textbf{else}\ a * a\ \textbf{fi})$$

We shall first show that $P \sqsubseteq_{\text{run}} Q$ and then argue that this is not desirable. Suppose $S,(\pi,\{r[c,d]\}) \Vdash P$. Recall that an **if** statement abbreviates a non-deterministic choice. Thus, there are two cases to be considered:

**Case 1:** $S,(\pi,\{r[c,d]\}) \Vdash [\phi]\,a$. Define the 1-local $p$-variant $\pi'$ of $\pi$ by $\pi'(p) = \text{Points}(S)$, that is, $p$ is everywhere true under $\pi'$. It follows that $S,(\pi',\{r[c,d]\}) \Vdash [p]\,a$, and thus, $S,(\pi',\{r[c,d]\}) \Vdash$ **if** $p$ **then** $a$ **else** $a * a$ **fi**. By definition, $S,(\pi,\{r[c,d]\}) \Vdash Q$.
**Case 2:** $S,(\pi,\{r[c,d]\}) \Vdash [\neg\phi]\,a * a$. This is handled analogously by defining $\pi'(p) = \emptyset$.

To see that it is not the case that $P \sqsubseteq Q$, take $\phi$ to be a propositional variable $q$. It is straightforward to construct a system $S$, finite intervals $i = r[c,d]$ and $i' = r'[c',d']$, and interpretation $\pi$ such that $S,(\pi,\{i\}) \Vdash [q]\, a$ and $S,(\pi,\{i'\}) \Vdash [\neg q]\, a * a$ (hence $S,(\pi,\{i,i'\}) \Vdash$ **if** $q$ **then** $a$ **else** $a*a$ **fi**), but $(r,c)$ and $(r',c')$ are 1-indistinguishable. If we were to have $S,(\pi,\{i,i'\}) \Vdash \exists_1 p\,($**if** $p$ **then** $a$ **else** $a*a$ **fi**$)$, then we would have a 1-local $p$-variant $\pi'$ of $\pi$ such that $S,(\pi',\{i,i'\}) \Vdash$ **if** $p$ **then** $a$ **else** $a*a$ **fi**. But by assumption $(r,c) \in \pi'(p)$ iff $(r',c') \in \pi'(p)$, so we have either $S,(\pi',\{i,i'\}) \Vdash a$ or $S,(\pi',\{i,i'\}) \Vdash a*a$. But neither of these is possible, since one or the other interval has the wrong length.

Our intuition in writing $Q$ is that it specifies a program that chooses to do either $a$ or $a*a$ on the basis of some locally computable test $p$. The refinement $P \sqsubseteq_{\text{run}} Q$ is contrary to this intuition: it states that $Q$ may be implemented by using in place of $p$ *any* test, even one not locally computable. Intuitively, this result is obtained by using a different 1-local test in different executions of the program. Our semantics has been designed so as to avoid this: it ensures that a *uniform* test $p$ is used in every execution of the program. Thereby, the undesirable refinement is blocked.

We remark that a slight variant of the example is a valid, and desired refinement: $[\exists_1 p\,(\mathsf{Nec}(p \equiv \phi))]\, P \sqsubseteq Q$. Here, the coercion states that $\phi$ is in fact equivalent to a 1-local proposition. We will use this rule below. □

## 4 Validity and Valid Refinement

We now briefly discuss the role of assertions $\{\phi\}_C$ in the framework and define the associated semantic notions. The reader is referred to [15] for a more detailed explanation of these ideas in a simpler setting.

Intuitively, an assertion $\{\phi\}_C$ is like an annotation at a program location stating that $\phi$ is guaranteed to hold whenever the program execution reaches this location. Moreover, such an assertion states that this fact "depends" only on constraints in the program (specifications and coercions) labelled with constraint variables in the set $C$, as well as on concrete program fragments. (We do not include labels for these because they cannot be "refined away".) The reason we include the justification $C$ for the assertion is that it proves to be necessary to track such information in order to be able to formulate a number of desirable refinement rules. These rules refine a program fragment in ways that depend upon the larger program context within which the fragment occurs.

One typical example of this is a rule concerning the elimination of coercions. Suppose a coercion $[\phi]$ occurs at a program location where $\phi$ is guaranteed to hold. Intuitively, we would like to say that the coercion can be eliminated (replaced by $\varepsilon$) in such circumstances. However, the attempt to formulate this by the refinement rule $\varepsilon \leq \{\phi\}\,[\phi]$ is not quite correct, for the reason the assertion holds could be the very coercion we seek to eliminate. (It may seem a little odd at first to say that the justification for the assertion is some part of the program text that follows, but consider the case of $\phi = \Diamond\psi$. See [15] for an example that makes essential use of assertions justified by later pieces of program text.) The use of justifications enables us to formulate the rule as $\varepsilon \leq \{\phi\}_C\,[\phi]^X$, *provided* $X$ is not in $C$, i.e., provided the assertion does not rely upon the coercion. This blocks the circular reasoning.

The semantics of assertions is formalized as follows. In order to capture constraint dependencies, we first define for each program $P$ and constraint set $C \subseteq CV$ a program $\text{relax}(P,C)$ that is like $P$, except that only constraints whose labels are in $C$ are enforced: all other constraints are relaxed. Formally, we obtain $\mathrm{relax}(P,C)$ from $P$ by replacing each occurrence of a coercion $[\phi]^X$ where $X \notin C$ by $\varepsilon$, and also replacing each occurrence of a specification $[\phi, \psi]^X$ where $X \notin C$ by $[\mathit{false}, \mathit{true}]^X$ in $P^C$.

We may now define a program $P$ to be *valid* with respect to a set of interpreted systems $\mathcal{S}$ when for all assertions $\{\phi\}_C$ in $P$, all interpreted systems $(S, \pi) \in \mathcal{S}$ and all intervals sets $I$ over $S$, all embeddings $\theta$ of $T_{\mathrm{relax}(P,C)}$ into $S, (I, \pi)$ have the property that for all nodes $n$ of $T_{\mathrm{relax}(P,C)}$ labelled with $\{\phi\}_C$ in $P$, we have $S, \theta(n) \Vdash [\phi]$. Intuitively, the embedding represents an execution of $P$ in which only constraints in $C$ are enforced, and we check that the associated assertions hold at the appropriate points in the execution. Note that when $n$ is labelled by an assertion, $I_\theta(n)$ must be a set of intervals of length 0. Moreover, the semantics of $S, (I, \pi) \Vdash [\phi]$ checks $\phi$ only at finite points in this set. Thus, validity can be understood as a kind of generalized partial correctness. We define validity with respect to a set of interpreted systems $\mathcal{S}$ to allow assumptions concerning the environment to be modelled: e.g., $S$ might be the set of all interpreted systems in which actions have specific intended interpretations. We give an example of this in the next section.

Clearly, we want to avoid programs that are not valid (such as $[p]^X \{\neg p\}_{\{X\}}$). Thus, we would now like a notion of refinement that preserves validity, so that we derive only valid programs from valid programs by refinement. The refinement relation $\sqsubseteq$ defined above does not have this property. We now define a notion that does. In order to do so, we first need to define a technical notion. A *justification transformation* is a mapping $\eta : 2^{CV} \longrightarrow 2^{CV}$ that is increasing, i.e., satisfies $C \subseteq \eta(C)$ for all $C \subseteq CV$. The result of applying a justification transformation $\eta$ to a program $P$ is the program $P\eta$ obtained by replacing each instance of an assertion $\{\phi\}_C$ in $P$ by the assertion $\{\phi\}_{\eta(C)}$. When $R(Z)$ is a program containing a program variable $Z$ and $P$ is a program, we write $R\eta(P)$ for the result of first applying $\eta$ to $R(Z)$ and then substituting $P$ for $Z$. We need such transformations for refinements such as replacing $\{\phi\}_C[\phi]^X$ by $\varepsilon$ when $X \notin C$ within some large program context. Intuitively, when we do this, any assertion in the larger context that depended on the coercion labelled $X$ is still valid, but its justification should now include $C$ in place of $X$.

The identity justification transformation is denoted by $\iota$. We will also represent justification transformations using expressions of the form $[X \hookrightarrow S]$, where $X \in CV$ and $S \subseteq CV$. Such an expression denotes the justification transformation $\eta$ such that $\eta(C) = C \cup S$ if $X \in C$ and $\eta(C) = C$ otherwise.

Let $\mathcal{S}$ be a set of interpreted systems, let $\eta$ be a justification transformation and let $P$ and $Q$ be programs. Say that *$P$ validly refines $Q$ in $\mathcal{S}$ under $\eta$*, and write $P \leq^{\mathcal{S}}_{\eta} Q$, if for all programs $R(Z)$ with $Z$ a program variable, if $R(Q)$ is valid with respect to $\mathcal{S}$ then $R\eta(P)$ is valid with respect to $\mathcal{S}$, and for all $(S, \pi) \in \mathcal{S}$ and interval sets $I$ over $S$, if $S, (I, \pi) \Vdash R\eta(P)$ then $S, (I, \pi) \Vdash R(Q)$.

We remark that other definitions of valid refinement are possible. In particular, the definition above is very sensitive to the syntax of the programming language. We consider some more semantic alternatives elsewhere.

### 4.1 Valid Refinement Rules

We now present a number of rules concerning valid refinement that are sound with respect to the semantics just presented. making no attempt at completeness. We focus on rules concerning the existential quantifiers, and refer to [15] for additional rules concerning the other constructs, which are also sound in the framework of the present paper.[2]

The following rules make it possible for refinement to broken down into a sequence of steps that operate on small program fragments. (Only justification transformation operates globally, but this can also be managed locally by means of appropriate data structures.)

$$\frac{P \leq^{\mathcal{S}}_{\eta} Q,\, Q \leq^{\mathcal{S}}_{\eta'} R}{P \leq^{\mathcal{S}}_{\eta\circ\eta'} R} \qquad \textbf{trans}$$

$$\frac{P \leq^{\mathcal{S}}_{\eta} Q}{R\eta(P) \leq^{\mathcal{S}}_{\eta} R(Q)} \qquad \textbf{mon}$$

Reducing the amount of nondeterminism and introducing a coercion are sound refinement steps.

$$P \leq^{\mathcal{S}}_{\iota} P + Q \qquad \textbf{red-ndet}$$

$$[\phi] \leq^{\mathcal{S}}_{\iota} \varepsilon \qquad \textbf{i-coerc}$$

Quantification over local propositional variables can be introduced, extracted from a coercion, and lifted to contexts.

$$\exists_1 p\,(P) \leq^{\mathcal{S}}_{\iota} P \quad \textit{if } p \textit{ not free in } P \qquad \textbf{i-lq}$$

$$\exists_1 p\,([\phi]) \leq^{\mathcal{S}}_{\iota} [\exists_1 p\,(\phi)] \qquad \textbf{ext-lq}$$

$$\exists_1 p\,(R(P)) \leq^{\mathcal{S}}_{\iota} R(\exists_1 p\,(P)) \quad \textit{if } p \textit{ not free in } R(Z) \qquad \textbf{lift-lq}$$

Let $P_\phi$ denote the program obtained from $P$ by substituting formula $\phi$ for all free occurrences of $p$ in $P$, while taking the usual care of free variables in $\phi$ by renaming clashing bound variables in $P$.

$$[\exists_1 p\,(\mathsf{Nec}(\phi \equiv p))]\, P_\phi \leq^{\mathcal{S}}_{\iota} \exists_1 p\,(P) \qquad \textbf{inst-lp}$$

### 4.2 Single-Stepping Programs and Loops

Reasoning about termination of a loop, say, **while** $g$ **do** $P$ **od** becomes easier when strict bounds on the running time of $P$ are known. We present here a simple example of this phenomenon that is useful for the example we present next. More general rules can be formulated than the one we develop here.

[2] For the benefit of the reviewer, these rules are included in the appendix.

Say that program $P$ is *single-stepping*, if $S,(\pi,I) \Vdash P$ and $r[c,d] \in I$ and $c < \infty$ imply that $d = 1 + c$, for all $S$, $\pi$, and $I$. Syntactically, the fact that $P$ is single-stepping is expressed by:

$$P \leq_{\eta}^{\mathcal{S}} \exists p\,([\bigcirc\,\mathsf{first}\,p]\;[true, \mathsf{first}\,p]) \quad .$$

where $\mathsf{first}\,\phi$ is an abbreviation for $\phi \wedge \neg \circleddash \diamondminus \phi$, which holds exactly at the first point in a run that makes $\phi$ true. This can be combined with the usual pre/post-condition style of specifying $P$'s behaviour, e.g., to specify that $P$ is single-stepping and terminates in points satisfying $\psi$ when started in points satisfying $\phi$:

$$P \leq_{\eta}^{\mathcal{S}} \exists p \left( [\bigcirc\,\mathsf{first}\,p]^X \;[true, \mathsf{first}\,p \wedge (\circleddash \phi \rightarrow \psi)]^X \right)$$

Call the RHS of the above $\mathbf{ss}[\phi,\psi]^X$. Observe that $\mathbf{ss}[\phi,\psi]^X$ takes a single step regardless of whether $\phi$ holds initially. Adding the single-stepping requirement yields a valid refinement:

$$\mathbf{ss}[\phi,\psi]^X \leq_{\eta}^{\mathcal{S}} [\phi,\psi]^X \qquad \textbf{ss-imp}$$

The following for single-stepping loop bodies will be used in Section 5.

$$\mathbf{while}^X\; g\; \mathbf{do}\; \mathbf{ss}[\alpha \wedge g, \alpha]^X\; \mathbf{od} \leq_{Z \hookrightarrow C \cup \{X\}}^{\mathcal{S}} \{\diamondsuit \neg g\}_C * [\alpha, \alpha \wedge \neg g]^Z \quad \text{if } Z \notin C \qquad \textbf{i-ss-loop}$$

## 5 Example: Autonomous Robot

In this section we discuss an example that closely resembles Example 7.2.2 in [4] which in turn has been inspired by the 1994 conference version of [2].

A robot travels along an endless corridor, which in this example is identified with the natural numbers. The robot starts at 0 and has the task to stop in the goal region $\{2,3,4\}$. To judge when to stop the robot has a sensor that reads the current position. (See Fig. 1.) Unfortunately, this sensor is not very accurate, i.e., the readings may be wrong by at most 1. The only action the robot can actively take is halting, the effect of which is instantaneous stopping. Unless this action is taken, the robot may move by steps of length 1 to higher numbers. Unless it has taken its halting action it is beyond its control whether it moves in a step. Our task is now to design a control program for the robot that initiates the halting action such that:

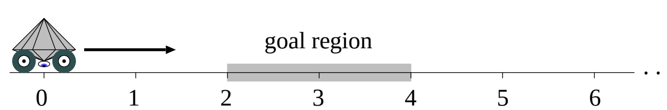


**Fig. 1.** Autonomous Robot

**(safety)** The robot only stops in the goal region.
**(liveness)** The robot is guaranteed to stop eventually.

A modest assumption about the environment is needed for the latter to be achievable. We insist that it is not the case that the robot sits still forever without moving forward or taking the halting action.

To model these assumptions we introduce a system constraint reflecting the following conditions. Strictly speaking, our specification language $\mathcal{L}$ only contains variables that are interpreted as Boolean values but none for natural numbers. It is possible to present this example only using propositions by sacrificing legibility. An extension of our framework to typed variables is straightforward and omitted here for brevity.

1. Initially, the robot's position $x$ is zero: $\mathit{init} \rightarrow x = 0$, where *init* abbreviates the formula $\neg\circleddash \mathit{true}$, which holds exactly in the initial points of runs.
2. Proposition $h$ is initially false and it is becomes true once the robot has halted. Halting is an irreversible action ($h \rightarrow \bigcirc h$) and means that the robot does not move anymore: $h \rightarrow x = \bigcirc x$.
3. Proposition $m$ is true iff the robot moves in the current step. Moving means that the robot's position is increased by one, otherwise it is unchanged: $(m \rightarrow x+1 = \bigcirc x) \wedge \neg m \rightarrow x = \bigcirc x$.
4. If the robot has not halted it should move eventually: $(\neg h)\,\mathsf{U}\,(h \vee m)$.
5. The robot's sensor reading is $s$ (an integer) and off by at most one from $x$, the actual position: $x - 1 \leq s \leq x + 1$.
6. Only the robot's basic action *halt* immediately halts the robot.

Let $\mathcal{S}$ be the set of pairs $(S, (\pi, I))$ such that $S$ and $\pi$ satisfy these constraints. In the full paper we introduce syntactic representation for such system constraints, give a formal semantics, and introduce valid refinement rules that exploit these constraints. These rules fall into two classes: assertion introduction rules and rules for specification implementation by basic actions. A typical assertion introduction rule is

$$[\mathit{init}]^X \{x = 0 \wedge \neg h\}_{\{X\}} \leq_\iota^{\mathcal{S}} \varepsilon \tag{1}$$

which allows to assert properties of initial states. For the halting action we would have

$$\mathit{halt} \leq_\iota^{\mathcal{S}} \mathbf{ss}[\mathit{true}, h \wedge x = \circleddash x] \quad .$$

For lack of space we simplified and pruned the set-up to the above. We refer to "use $\mathcal{S}$" instead of proper refinement rules at points of our derivation that suffer from these limitations.

In [4] a run-based specification of the system is given by a temporal logic formula equivalent to $\square(h \rightarrow g) \wedge \diamond h$, where $g$ abbreviates being in the goal region, i.e., $2 \leq x \leq 4$. The two conjuncts respectively formalize the safety and liveness property from above. The main problem in finding the robot's protocol is to derive a suitable local condition for halting.

We formally derive a protocol for the robot from an as abstract as possible specification of the protocol. The point of departure of our derivation below merely states that the robot must eventually halt in the goal region when started in an initial state.

$$[\mathit{init}, g \wedge h]^X$$

$\geq_{\iota}^{\mathcal{S}}$ (sequential composition [15])

$[init, g \wedge \neg h]^X * [g \wedge \neg h, g \wedge h]^X$

$\geq_{\iota}^{\mathcal{S}}$ (use $\mathcal{S}$ and **ss-imp** to establish $halt \leq_{\iota}^{\mathcal{S}} [g \wedge \neg h, g \wedge h]$)

$[init, g \wedge \neg h]^X * halt$

$\geq_{\iota}^{\mathcal{S}}$ (**i-lq** with local proposition $p$ not free in $[init, g \wedge \neg h]^X$)

$\exists_1 p\left([init, g \wedge \neg h]^X\right) * halt$

Next we introduce a coercion that enforces suitability of $p$ as a test for halting the robot in the goal region.

$\geq_{\iota}^{\mathcal{S}}$ (**i-coerc** with $st_p$ abbreviating $\mathsf{Nec}(\mathsf{first}(p) \rightarrow g)$, strengthen spec [15])

$\exists_1 p\left([st_p \wedge \Diamond p \wedge \neg h]^Y * [init, g \wedge \mathsf{first}(p) \wedge st_p \wedge \neg h]^X\right) * halt$

$\geq_{\iota}^{\mathcal{S}}$ (propagate coercion into precondition [15])

$\exists_1 p\left([st_p \wedge \Diamond p \wedge \neg h]^Y * [init \wedge st_p \wedge \neg h, g \wedge \mathsf{first}(p) \wedge st_p \wedge \neg h]^X\right) * halt$

$\geq_{\iota}^{\mathcal{S}}$ (introduce assertion from coercion [15])

$\exists_1 p\left([st_p \wedge \Diamond p \wedge \neg h]^Y * \{\Diamond p\}_{\{Y\}} * [init \wedge st_p \wedge \neg h, g \wedge \mathsf{first}(p) \wedge st_p \wedge \neg h]^X\right) * halt$

$\geq_{\iota}^{\mathcal{S}}$ (strengthen spec [15] with $\neg\ominus true \rightarrow \neg\ominus\diamondminus p$ valid)

$\exists_1 p\left([st_p \wedge \Diamond p \wedge \neg h]^Y * \{\Diamond p\}_{\{Y\}} * [\neg\ominus\diamondminus p \wedge st_p \wedge \neg h, g \wedge \mathsf{first}(p) \wedge st_p \wedge \neg h]^X\right) * halt$

$\geq_{[X \hookrightarrow \{Y\}]}^{\mathcal{S}}$ (**i-ss-loop**)

$\exists_1 p\begin{pmatrix}[st_p \wedge \Diamond p \wedge \neg h]^Y * \\ \textbf{while}^X\ \neg p\ \textbf{do ss}[\neg\ominus\diamondminus p \wedge st_p \wedge \neg h \wedge \neg p, \neg\ominus\diamondminus p \wedge st_p \wedge \neg h]^X\ \textbf{od}\end{pmatrix} *$
$halt$

$\geq_{\iota}^{\mathcal{S}}$ (use $\mathcal{S}$ establish $\Lambda \leq_{\iota}^{\mathcal{S}} \textbf{ss}[\ldots,\ldots]^X$)

$\exists_1 p\left([st_p \wedge \Diamond p \wedge \neg h]^Y\ \textbf{while}^X\ \neg p\ \textbf{do}\ \Lambda\ \textbf{od}\right) * halt$

$\geq_{\iota}^{\mathcal{S}}$ (**inst-lp *but how did we find*** $s > 2$**?**)

$[\exists_1 p\,(\mathsf{Nec}(p \equiv (s > 2)))]^Y * [st_{s>2} \wedge \Diamond(s > 2) \wedge \neg h]^Y *$
$\textbf{while}^X\ \neg(s > 2)\ \textbf{do}\ \Lambda\ \textbf{od} * halt$

$\geq_{\iota}^{\mathcal{S}}$ (use $\mathcal{S}$ and coercion elimination [15])

$[st_{s>2} \wedge \Diamond(s > 2) \wedge \neg h]^Y\ \textbf{while}^X\ s \leq 2\ \textbf{do}\ \Lambda\ \textbf{od} * halt$

To eliminate the coercion later we first assert $st_{s>2}$, which states that $s > 2$ is a sound test for exiting the loop and executing the *halt* action. We prove that $st_{s>2}$ holds in every single point of an element of $\mathcal{S}$. By definition of *st* and the semantics of Nec, it suffices to prove that $\mathsf{first}(s > 2) \rightarrow g$ follows from the description of $\mathcal{S}$. Let $(S, (\pi, I)) \in \mathcal{S}$ and suppose $\mathfrak{I}, (r, c) \models \mathsf{first}(s > 2)$ for some point $(r, c)$ in $I$. By definition of first, also $\mathfrak{I},(r,c) \models (s>2) \wedge \circleddash(s \leq 2)$ holds. The invariant guarantees that the sensor is off the current position by at most 1, thus, $\mathfrak{I},(r,c) \models (x \geq 2) \wedge \circleddash(x \leq 3)$. Moreover, if the position $x$ changes at all during a step, then it increases by one. Consequently $\mathfrak{I},(r,c) \models (x \leq 4)$.

$\geq_{\iota}^{\mathcal{S}}$ (use $\mathcal{S}$ according to the discussion above and weaken assertion)

$\{st_{s>2}\}_{\emptyset} * [st_{s>2} \wedge \Diamond(s>2) \wedge \neg h]^{Y}$ **while**$^{X}$ $s \leq 2$ **do** $\Lambda$ **od** $*$ *halt*

$\geq_{\iota}^{\mathcal{S}}$ (use coercion elimination [15] to get rid of one conjunct)

$[\Diamond(s>2) \wedge \neg h]^{Y}$ **while**$^{X}$ $s \leq 2$ **do** $\Lambda$ **od** $*$ *halt*

$\geq_{\iota}^{\mathcal{S}}$ (strengthen and split coercion [15])

$[init]^{Y} * [\Diamond(s>2)]^{Y}$ **while**$^{X}$ $s \leq 2$ **do** $\Lambda$ **od** $*$ *halt*

How can we eliminate the coercion $[\Diamond(s>2)]^{Y}$? It certainly does not follow from $\mathcal{S}$ alone, since that allows runs in which the robot halts too early, i.e., outside the goal region and without a sensor reading $s>2$ . Therefore the reasoning that allows to eliminate the coercion necessarily involves the program derived thus far.

From the initial state predicate it follows that the loop begins in a state satisfying $\neg h$. The only action executed in the loop is $\Lambda$, which, according to *sc* preserves the value of $h$. On termination of the loop the guard must be false, i.e., $s>2$. In (the purely hypothetical) case the loop diverges the run satisfies $\Box\neg h$, which together with point 4, $(\neg h)\,\mathsf{U}\,(h \vee m)$, allows us to conclude that the robot moves infinitely often. Actually, four times is enough, since that already guarantees a sensor reading of at least 3.

$\geq_{\iota}^{\mathcal{S}}$ (use $\mathcal{S}$ and the loop)

$[init]^{Y} * [\Diamond(s>2)]^{Y} * \{\Diamond(s>2)\}_{X}$ **while**$^{X}$ $s \leq 2$ **do** $\Lambda$ **od** $*$ *halt*

$\geq_{\iota}^{\mathcal{S}}$ (eliminate coercion)

$[init]^{Y}$ **while**$^{X}$ $s \leq 2$ **do** $\Lambda$ **od** $*$ *halt*

Finally, the rule

$$\frac{[\phi]\, P \leq_{\eta}^{\mathcal{S}} [\phi,\psi]^{X}}{P \leq_{\eta}^{\mathcal{S}} [\phi,\psi]^{X}} \qquad \textbf{coerc-elim1}$$

proves that **while**$^{X}$ $s \leq 2$ **do** $\Lambda$ **od** $*$ *halt* $\leq_{X \hookrightarrow \{Y\}}^{\mathcal{S}} [init, g \wedge h]^{X}$.

## 6 Conclusion and Future Work

*Keywords:* add labels to model knowledge-based programs with their fixed-point characterization of knowledge, asynchronous case, multi-agent case, tool support.

## References


1. R.-J. Back and J. von Wright. *Refinement Calculus: A Systematic Introduction*. Graduate Texts in Computer Science. Springer-Verlag, 1998.
2. R. I. Brafman, J.-C. Latombe, Y. Moses, and Y. Shoham. Applications of a logic of knowledge to motion planning under uncertainty. *Journal of the ACM*, 44(5):633–668, Sept. 1997.
3. K. Engelhardt, R. van der Meyden, and Y. Moses. Knowledge and the logic of local propositions. In I. Gilboa, editor, *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Seventh Conference (TARK 1998)*, pages 29–41. Morgan Kaufmann, July 1998.
4. R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge*. MIT-Press, 1995.
5. R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. Knowledge-based programs. *Distributed Computing*, 10(4):199–225, 1997.
6. J. Y. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, July 1990.
7. I. Hayes. Separating timing and calculation in real-time refinement. In J. Grundy, M. Schwenke, and T. Vickers, editors, *International Refinement Workshop and Formal Methods Pacific 1998*, Discrete Mathematics and Theoretical Computer Science, pages ??–?? Springer-Verlag, 1998.
8. Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Specification*. Springer-Verlag, 1992.
9. C. C. Morgan. *Programming from Specifications*. Prentice Hall, 1990.
10. J. M. Morris. A theoretical basis for stepwise refinement and the programming calculus. *Science of Computer Programming*, 9(3):287–306, Dec. 1987.
11. Y. Moses and O. Kislev. Knowledge-oriented programming. In *Proceeding of the 12th Annual ACM Symposium on Principles of Distributed Computing (PODC 93)*, pages 261–270, New York, USA, Aug. 1993. ACM Press.
12. Y. Moses and M. R. Tuttle. Programming simultaneous actions using common knowledge. *Algorithmica*, 3:121–169, 1988.
13. M. Utting and C. Fidge. A real-time refinement calculus that changes only time. In H. J., J. Cooke, and P. Wallis, editors, *BCS-FACS Seventh Refinement Workshop*. Springer-Verlag, 1996.
14. R. van der Meyden. Knowledge based programs: On the complexity of perfect recall in finite environments. In Y. Shoham, editor, *Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge*, pages 31–50. Morgan Kaufmann, Mar. 17–20 1996.
15. R. van der Meyden and Y. Moses. On refinement and temporal annotations. Submitted to TACAS 2000. A progress report on this subject appeared as [16].
16. R. van der Meyden and Y. Moses. Issues for a temporal refinement calculus. In J. Grundy, M. Schwenke, and T. Vickers, editors, *International Refinement Workshop and Formal Methods Pacific 1998*, Discrete Mathematics and Theoretical Computer Science, pages ??–?? Springer-Verlag, 1998.
17. R. van der Meyden and Y. Moses. Top-down considerations on distributed systems. In *Proceedings 12th International Symposium on Distributed Computing, DISC'98*, volume 1499 of *LNCS*, pages 16–19, Sept. 1998. Springer-Verlag.
18. M. Y. Vardi. Implementing knowledge-basd programs. In Y. Shoham, editor, *Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge*, pages 15–30. Morgan Kaufmann, Mar. 17–20 1996.