The Temporal Logic of Actions

LESLIE LAMPORT
Digital Equipment Corporation

The temporal logic of actions (TLA) is a logic for specifying and reasoning about concurrent systems. Systems and their properties are represented in the same logic, so the assertion that a system meets its specification and the assertion that one system implements another are both expressed by logical implication. TLA is very simple; its syntax and complete formal semantics are summarized in about a page. Yet, TLA is not just a logician's toy; it is extremely powerful, both in principle and in practice. This report introduces TLA and describes how it is used to specify and verify concurrent algorithms. The use of TLA to specify and reason about open systems will be described elsewhere.

Categories and Subject Descriptors: D.2.4 [**Software Engineering**]: Program Verification—*correctness proofs*; F.3.1 [**Logics and Meanings of Programs**]: Specifying and Verifying and Reasoning about Programs—*Specification techniques*

General terms: Theory, Verification

Additional Key Words and Phrases: Concurrent programming, liveness properties, safety properties

1. LOGIC VERSUS PROGRAMMING

A concurrent algorithm is usually specified with a program. Correctness of the algorithm means that the program satisfies a desired property. We propose a simpler approach in which both the algorithm and the property are specified by formulas in a single logic. Correctness of the algorithm means that the formula specifying the algorithm implies the formula specifying the property, where *implies* is ordinary logical implication.

We are motivated not by an abstract ideal of elegance, but by the practical problem of reasoning about real algorithms. Rigorous reasoning is the only way to avoid subtle errors in concurrent algorithms, and we want to make reasoning as simple as possible by making the underlying formalism simple.

How can we abandon conventional programming languages in favor of logic if the algorithm must be coded as a program to be executed? The answer is that we almost always reason about an abstract algorithm, not about a concurrent program that is actually executed. A typical example is the distributed spanning-tree algorithm used in the Autonet local area network [Schroeder et al. 1990]. The algorithm can be described in about one page of pseudo-code, but its implementation required

Author's address: Systems Research Center, Digital Equipment Corporation, 130 Lytton Avenue, Palo Alto, CA 94301.

about 5000 lines of C code and 500 lines of assembly code.[1] Reasoning about 5000 lines of C would be a herculean task, but we can reason about a one-page abstract algorithm. By starting from a correct algorithm, we can avoid the timing-dependent synchronization errors that are the bane of concurrent programming. If the algorithms we reason about are not real, compilable programs, then they do not have to be written in a programming language.

But, why replace a programming language by logic? Aren't programs simpler than logical formulas? The answer is no. Logic is the formalization of everyday mathematics, and everyday mathematics is simpler than programs. Consider the Pascal statement $y := x + 1$. Using the convention that y' denotes the new value of y, we can rewrite this statement as the mathematical formula $y' = x + 1$. Many readers will think that the Pascal statement and the formula are equally simple. They are wrong. The formula is much simpler than the Pascal statement. Equality is a simple concept that five-year-old children understand. Assignment (:=) is a complicated concept that university students find difficult. Equality obeys simple algebraic laws; assignment doesn't. If we assume that all variables are integer-valued, we can subtract y' from both sides of the formula to obtain the equivalent formula $0 = x + 1 - y'$. Trying this with the Pascal statement yields the absurdity $0 := x + 1 - y$.

A programming language may use mathematical terms like *function*, but the constructs they represent are not as simple as the corresponding mathematical concepts. Mathematical functions are simple; children in the United States learn about them at the age of twelve. Pascal functions are complicated, involving concepts like call by reference, call by value, and aliasing; it is unlikely that many university students understand them well. Advocates of so-called functional programming languages often claim that they just use ordinary mathematical functions, but try explaining to a twelve-year-old how evaluating a mathematical function can display a character on her computer screen.

Since real languages like Pascal are so complicated, methods for reasoning about algorithms are usually based on toy languages. Although simpler than real programming languages, toy languages are more complicated than simple logic. Moreover, their resemblance to real languages can be dangerously misleading. In toy languages, the Hoare triple $\{x = 0\}\ y := x + 1\ \{y = x + 1\}$ is valid, which means that executing $y := x + 1$ in a state in which x equals 0 produces a state in which y equals $x + 1$. However, in Pascal, the program fragment

$$x := 0;\ y := x + 1;\ \mathrm{write}(y, x + 1)$$

can print two different values when used in certain contexts, even if x and y are variables of type **integer**. The programmer who tries using toy-language rules to reason about real Pascal programs is in for a rude surprise.

We do not mean to belittle programming languages. They are complicated because they have a difficult job to do. Mathematics can be based on simple concepts like functions. Programming languages cannot, because they must allow reasonably simple compilers to translate programs into reasonably efficient code for complex computers. Real languages must embrace difficult concepts like the distinction between values and locations, which leads to call-by-reference arguments and

[1] Assembly code was needed because C has no primitives for sending messages across wires.

aliasing—complications that have no counterpart in simple mathematics. Programming languages are necessary for writing real programs; but mathematics offers a simpler alternative for reasoning about concurrent algorithms.

To offer a practical alternative to programming languages, a logic must be both simple and expressive. There is no point trading a programming language for a logic that is just as complicated and hard to understand. Furthermore, a logic that achieves simplicity at the expense of needed expressiveness will be impractical because the formulas describing real algorithms will be too long and complicated to understand.

The logic that we propose for reasoning about concurrent algorithms is the temporal logic of actions, abbreviated as TLA. All TLA formulas can be expressed in terms of familiar mathematical operators (such as $\wedge$) plus three new ones: $'$ (prime), $\Box$, and $\exists\!\!\exists$. TLA is simple enough that its syntax and complete formal semantics can be written in about a page. Almost all that is needed to specify and reason about algorithms in TLA—its syntax, formal semantics, derived notation, and axioms and proof rules—appears in Figures 4 and 5 of Section 5.6 and Figure 9 of Section 8.2. (Missing from those figures are the rules for adding auxiliary variables, mentioned in Section 8.3.2.)

Logic is a tool. Its true test comes with use. Although TLA and its proof rules can be described formally in a couple of pages, such a description would tell you nothing about how TLA is used. In this article, we develop TLA as a method of describing and reasoning about concurrent algorithms. We limit ourselves to simple examples, so we can only hint at how TLA works with real algorithms.

TLA combines two logics: a logic of actions, described in Section 2, and a standard temporal logic, described in Section 3. TLA is easiest to explain in terms of a logic called RTLA, which is defined in Section 4. We describe TLA itself and illustrate its use in Sections 5–8. Section 9 mentions further applications and discusses what TLA can and cannot do, and Section 10 relates TLA to other formalisms.

2. THE LOGIC OF ACTIONS

2.1 Values, Variables, and States

Algorithms manipulate data. We assume a collection Val of *values*, where a value is a data item. The collection Val includes numbers such as 1, 7, and -14, strings like "abc", and sets like the set Nat of natural numbers. We don't bother to define Val precisely, but simply assume that it contains all the values needed for our examples (Note[2] 1). We also assume the booleans true and false, which for technical reasons are not considered to be values.

We think of algorithms as assigning values to variables. We assume an infinite set **Var** of variable names. We won't describe a precise syntax for generating variable names, but will simply use names like x and sem.

A logic consists of a set of rules for manipulating formulas. To understand what the formulas and their manipulation mean, we need a semantics. A semantics is given by assigning a semantic meaning $[\![F]\!]$ to each syntactic object F in the logic.

The semantics of our logic is defined in terms of *states*. A state is an assignment of values to variables—that is, a mapping from the set **Var** of variable names to

[2]Notes appear at the end of the article.

the collection Val of values. Thus a state s assigns a value $s(x)$ to a variable x. The collection of all possible states is denoted $\mathbf{St}$.

We write $s[\![x]\!]$ to denote $s(x)$. Thus, we consider the meaning $[\![x]\!]$ of the variable x to be a mapping from states to values, using a postfix notation for function application. States and values are purely semantic concepts; they do not appear explicitly in formulas.

2.2 State Functions and Predicates

A *state function* is a nonboolean expression built from variables and constant symbols—for example, $x^2+y-\mathsf{3}$ (Note 2). The meaning $[\![f]\!]$ of a state function f is a mapping from the collection $\mathbf{St}$ of states to the collection Val of values. For example, $[\![x^2+y-\mathsf{3}]\!]$ is the mapping that assigns to a state s the value $(s[\![x]\!])^2+s[\![y]\!]-3$, where $\mathsf{2}$ and $\mathsf{3}$ are constant symbols, and 2 and 3 are the values that they represent. We will not bother distinguishing between constant symbols and their values. We use a postfix functional notation, letting $s[\![f]\!]$ denote the value that $[\![f]\!]$ assigns to state s. The semantic definition is

$$s[\![f]\!] \;\triangleq\; f(\forall\, `v' : s[\![v]\!]/v) \tag{1}$$

where $f(\forall\, `v' : s[\![v]\!]/v)$ denotes the value obtained from f by substituting $s[\![v]\!]$ for v, for all variables v. (The symbol $\triangleq$ means *equals by definition.*)

A variable x is a state function—the state function that assigns the value $s[\![x]\!]$ to the state s. The definition of $[\![f]\!]$ for a state function f therefore extends the definition of $[\![x]\!]$ for a variable x.

A *state predicate*, called a *predicate* for short, is a boolean expression built from variables and constant symbols—for example, $x^2 = y-3$ and $x \in \mathsf{Nat}$. The meaning $[\![P]\!]$ of a predicate P is a mapping from states to booleans, so $s[\![P]\!]$ equals true or false for every state s. We say that a state s *satisfies* a predicate P iff (if and only if) $s[\![P]\!]$ equals true.

State functions correspond both to expressions in ordinary programming languages and to subexpressions of the assertions used in ordinary program verification. Predicates correspond both to boolean-valued expressions in programming languages and to assertions.

2.3 Actions

An *action* is a boolean-valued expression formed from variables, primed variables, and constant symbols—for example, $x' + 1 = y$ and $x - 1 \notin z'$ are actions, where x, y, and z are variables.

An action represents a relation between old states and new states, where the unprimed variables refer to the old state and the primed variables refer to the new state. Thus, $y = x' + 1$ is the relation asserting that the value of y in the old state is one greater than the value of x in the new state. An atomic operation of a concurrent program will be represented in TLA by an action.

Formally, the meaning $[\![\mathcal{A}]\!]$ of an action $\mathcal{A}$ is a relation between states—a function that assigns a boolean $s[\![\mathcal{A}]\!]t$ to a pair of states s, t. We define $s[\![\mathcal{A}]\!]t$ by considering s to be the "old state" and t the "new state", so $s[\![\mathcal{A}]\!]t$ is obtained from $\mathcal{A}$ by replacing each unprimed variable v by $s[\![v]\!]$ and each primed variable v' by $t[\![v]\!]$:

$$s[\![\mathcal{A}]\!]t \;\triangleq\; \mathcal{A}(\forall\, `v' : s[\![v]\!]/v, t[\![v]\!]/v') \tag{2}$$

Thus, $s[\![y = x' + 1]\!]t$ equals the boolean $s[\![y]\!] = t[\![x]\!] + 1$.

The pair of states s, t is called an "$\mathcal{A}$ step" iff $s[\![\mathcal{A}]\!]t$ equals true. If action $\mathcal{A}$ represents an atomic operation of a program, then s, t is an $\mathcal{A}$ step iff executing the operation in state s can produce state t.

2.4 Predicates as Actions

We have defined a predicate P to be a boolean-valued expression built from variables and constant symbols, so $s[\![P]\!]$ is a boolean, for any state s. We can also view P as an action that contains no primed variables. Thus, $s[\![P]\!]t$ is a boolean, which equals $s[\![P]\!]$, for any states s and t. A pair of states s, t is a P step iff s satisfies P.

Both state functions and predicates are expressions built from variables and constant symbols. For any state function or predicate F, we define F' to be the expression obtained by replacing each variable v in F by the primed variable v':

$$F' \;\triangleq\; F(\forall \text{`}v\text{'} : v'/v) \tag{3}$$

If P is a predicate, then P' is an action, and $s[\![P']\!]t$ equals $t[\![P]\!]$ for any states s and t.

2.5 Validity and Provability

An action $\mathcal{A}$ is said to be *valid*, written $\models \mathcal{A}$, iff every step is an $\mathcal{A}$ step. Formally,

$$\models \mathcal{A} \;\triangleq\; \forall s, t \in \mathbf{St} : s[\![\mathcal{A}]\!]t$$

As a special case of this definition, if P is a predicate, then

$$\models P \;\triangleq\; \forall s \in \mathbf{St} : s[\![P]\!]$$

A valid action is one that is true regardless of what values one substitutes for the primed and unprimed variables. For example, the action

$$(x' + y \in \mathsf{Nat}) \;\Rightarrow\; (2(x' + y) \geq x' + y) \tag{4}$$

is valid. The validity of an action thus expresses a theorem about values.

A logic contains rules for proving formulas. It is customary to write $\vdash F$ to denote that formula F is provable by the rules of the logic. Soundness of the logic means that every provable formula is valid—in other words, that $\vdash F$ implies $\models F$. The validity of an action such as (4) is proved by ordinary mathematical reasoning (Note 3). How this reasoning is formalized does not concern us here, so we will not bother to define a logic for proving the validity of actions. But, this omission does not mean such reasoning is unimportant. When verifying the validity of TLA formulas, most of the work goes into proving the validity of actions (and of predicates, a special class of actions). The practical success of any TLA verification will depend primarily on how good the verifier is at ordinary mathematical reasoning.

2.6 Rigid Variables and Quantifiers

Consider a program that is described in terms of a parameter n—for example, an n-process mutual exclusion algorithm. An action representing an atomic operation of that program may contain the symbol n. This symbol does not represent a known value like 1 or "abc". But unlike the variables we have considered so far, the value of n does not change; it must be the same in the old and new state.

The symbol n denotes some fixed but unknown value. A programmer would call it a *constant* because its value doesn't change during execution of the program, while a mathematician would call it a *variable* because it is an "unknown". We call such a symbol n a *rigid variable*. The variables introduced above will be called *flexible variables*, or simply *variables*.

An expression like $\mathsf{n}+1$, built from rigid variables and constant symbols, is called a *constant expression*. We generalize state functions and actions to allow arbitrary constant expressions instead of just constant symbols, and to allow quantification over rigid variables. For example, if x is a (flexible) variable and m and n are rigid variables, then $\exists\, \mathsf{m} \in \mathsf{Nat} : \mathsf{m}x' = \mathsf{n} + x$ is the action asserting that there exists some natural number m such that m times the value of x in the new state equals n plus its value in the old state (Note 4):

$$s\,[\![\exists\, \mathsf{m} \in \mathsf{Nat} : \mathsf{m}x' = \mathsf{n} + x]\!]\,t \;\;\stackrel{\Delta}{=}\;\; \exists\, \mathsf{m} \in \mathsf{Nat} : \mathsf{m}(t[\![x]\!]) = \mathsf{n} + s[\![x]\!]$$

Thus, the semantics of state functions and actions is no longer given in terms only of values, but of first-order formulas containing free rigid variables and values. However, a state is still an assignment of values to flexible variables.

An action $\mathcal{A}$ is valid iff $s[\![\mathcal{A}]\!]t$ equals true for all states s and t and all possible values of its free rigid variables—for example:

$$\models\; (x' + y + \mathsf{m} \in \mathsf{Nat}) \;\;\Rightarrow\;\; \forall \mathsf{n} \in \mathsf{Nat} : \mathsf{n}(x' + y + \mathsf{m}) \geq (x' + y + \mathsf{m})$$

We do not permit quantification over flexible variables in state functions and actions.

2.7 The *Enabled* Predicate

For any action $\mathcal{A}$, we define *Enabled* $\mathcal{A}$ to be the predicate that is true for a state iff it is possible to take an $\mathcal{A}$ step starting in that state. Semantically, *Enabled* $\mathcal{A}$ is defined by

$$s[\![\mathit{Enabled}\ \mathcal{A}]\!] \;\;\stackrel{\Delta}{=}\;\; \exists\, t \in \mathbf{St} : s[\![\mathcal{A}]\!]t \tag{5}$$

for any state s. The predicate *Enabled* $\mathcal{A}$ can be defined syntactically as follows. If $v_1, \ldots, v_n$ are all the (flexible) variables that occur in $\mathcal{A}$, then

$$\mathit{Enabled}\ \mathcal{A} \;\;\stackrel{\Delta}{=}\;\; \exists\, \mathsf{c}_1, \ldots, \mathsf{c}_n : \mathcal{A}(\mathsf{c}_1/v_1', \ldots, \mathsf{c}_n/v_n')$$

where $\mathcal{A}(\mathsf{c}_1/v_1', \ldots, \mathsf{c}_n/v_n')$ denotes the formula obtained by substituting new rigid variables c_i for all occurrences of the v_i' in $\mathcal{A}$. For example,

$$\mathit{Enabled}\ (y = (x')^2 + \mathsf{n}) \;\;\equiv\;\; \exists\, \mathsf{c} : y = \mathsf{c}^2 + \mathsf{n}$$

If action $\mathcal{A}$ represents an atomic operation of a program, then *Enabled* $\mathcal{A}$ is true for those states in which it is possible to perform the operation.

3. SIMPLE TEMPORAL LOGIC

An execution of an algorithm is often thought of as a sequence of steps, each producing a new state by changing the values of one or more variables. We will consider an execution to be the resulting sequence of states, and will take the semantic meaning of an algorithm to be the collection of all its possible executions. Reasoning about algorithms will therefore require reasoning about sequences of states. Such reasoning is the province of temporal logic.

3.1 Temporal Formulas

A temporal formula is built from elementary formulas using boolean operators and the unary operator $\Box$ (read *always*). For example, if E_1 and E_2 are elementary formulas, then $\neg E_1 \wedge \Box(\neg E_2)$ and $\Box(E_1 \Rightarrow \Box(E_1 \vee E_2))$ are temporal formulas. We define simple temporal logic for an arbitrary class of elementary formulas. TLA will be defined later as a special case of simple temporal logic by specifying its elementary formulas.

The semantics of temporal logic is based on *behaviors*, where a behavior is an infinite sequence of states. Think of a behavior as the sequence of states that a computing device might go through when executing an algorithm. (It might seem that a terminating execution would be represented by a finite sequence of states, but we will see in Section 5.5 why infinite sequences are enough.)

We will define the meaning of a temporal formula in terms of the meanings of the elementary formulas it contains. Since an arbitrary temporal formula is built up from elementary formulas using boolean operators and the $\Box$ operator, and all the boolean operators can be defined in terms of $\wedge$ and $\neg$, it suffices to define $[\![F \wedge G]\!]$, $[\![\neg F]\!]$, and $[\![\Box F]\!]$ in terms of $[\![F]\!]$ and $[\![G]\!]$.

We interpret a temporal formula as an assertion about behaviors. Formally, the meaning $[\![F]\!]$ of a formula F is a boolean-valued function on behaviors. We let $\sigma[\![F]\!]$ denote the boolean value that formula F assigns to behavior σ, and we say that σ *satisfies* F iff $\sigma[\![F]\!]$ equals true.

The definitions of $[\![F \wedge G]\!]$ and $[\![\neg F]\!]$ are simple:

$$\begin{aligned} \sigma[\![F \wedge G]\!] &\triangleq \sigma[\![F]\!] \wedge \sigma[\![G]\!] \\ \sigma[\![\neg F]\!] &\triangleq \neg \sigma[\![F]\!] \end{aligned}$$

In other words, a behavior satisfies $F \wedge G$ iff it satisfies both F and G; and a behavior satisfies $\neg F$ iff it does not satisfy F. One can derive similar formulas for the other boolean operators. For example, since $F \Rightarrow G$ equals $\neg(F \wedge \neg G)$, a straightforward calculation proves that $\sigma[\![F \Rightarrow G]\!]$ equals $\sigma[\![F]\!] \Rightarrow \sigma[\![G]\!]$.

We now define $[\![\Box F]\!]$ in terms of $[\![F]\!]$. Let $\langle s_0, s_1, s_2, \ldots \rangle$ denote the behavior whose first state is s_0, second state is s_1, and so on. Then

$$\langle s_0, s_1, s_2, \ldots \rangle [\![\Box F]\!] \;\triangleq\; \forall n \in \mathsf{Nat} : \langle s_n, s_{n+1}, s_{n+2}, \ldots \rangle [\![F]\!] \tag{6}$$

Think of the behavior $\langle s_0, \ldots \rangle$ as representing the evolution of the universe, where s_n is the state of the universe at "time" n. The formula $\langle s_0, \ldots \rangle [\![F]\!]$ asserts that F is true at time 0 of this behavior, and $\langle s_n, \ldots \rangle [\![F]\!]$ asserts that it is true at time n. Thus, $\langle s_0, \ldots \rangle [\![\Box F]\!]$ asserts that F is true at all times during the behavior $\langle s_0, \ldots \rangle$. In other words, $\Box F$ asserts that F is *always* true.

3.2 Some Useful Temporal Formulas

3.2.1 *Eventually.* For any temporal formula F, let $\Diamond F$ be defined by

$$\Diamond F \;\triangleq\; \neg \Box \neg F \tag{7}$$

This formula asserts that it is not the case that F is always false. In other words, $\Diamond F$ asserts that F is *eventually* true. Since $\neg \forall \neg$ is the same as $\exists$, we have

$$\langle s_0, s_1, s_2, \ldots \rangle [\![\Diamond F]\!] \;\equiv\; \exists n \in \mathsf{Nat} : \langle s_n, s_{n+1}, s_{n+2}, \ldots \rangle [\![F]\!]$$

for any behavior $\langle s_0, s_1, \ldots \rangle$. Therefore, a behavior satisfies $\Diamond F$ iff F is true at some time during the behavior.

3.2.2 *Infinitely Often and Eventually Always.* The formula $\Box\Diamond F$ is true for a behavior iff $\Diamond F$ is true at all times n during that behavior, and $\Diamond F$ is true at time n iff F is true at some time m greater than or equal to n. Formally,

$$\langle s_0, s_1, \ldots \rangle[\![\Box\Diamond F]\!] \;\equiv\; \forall n \in \mathsf{Nat} : \exists m \in \mathsf{Nat} : \langle s_{n+m}, s_{n+m+1}, \ldots \rangle[\![F]\!]$$

A formula of the form $\forall n : \exists m : g(n+m)$ asserts that $g(i)$ is true for infinitely many values of i. Thus, a behavior satisfies $\Box\Diamond F$ iff F is true at infinitely many times during the behavior. In other words, $\Box\Diamond F$ asserts that F is true *infinitely often*.

The formula $\Diamond\Box F$ asserts that eventually F is always true. Thus, a behavior satisfies $\Diamond\Box F$ iff there is some time such that F is true from that time on.

3.2.3 *Leads To.* For any temporal formulas F and G, we define $F \leadsto G$ to equal $\Box(F \Rightarrow \Diamond G)$. This formula asserts that any time F is true, G is true then or at some later time. The operator $\leadsto$ (read *leads to*) is transitive, meaning that any behavior satisfying $F \leadsto G$ and $G \leadsto H$ also satisfies $F \leadsto H$. We suggest that readers convince themselves both that $\leadsto$ is transitive, and that it would not be had $F \leadsto G$ been defined to equal $F \Rightarrow \Diamond G$.

3.3 Validity and Provability

A temporal formula F is said to be *valid*, written $\models F$, iff it is satisfied by all possible behaviors. More precisely,

$$\models F \;\triangleq\; \forall \sigma \in \mathbf{St}^\infty : \sigma[\![F]\!] \tag{8}$$

where $\mathbf{St}^\infty$ denotes the collection of all behaviors (infinite sequences of elements of $\mathbf{St}$).

We will represent both algorithms and properties as temporal formulas. An algorithm is represented by a temporal formula F such that $\sigma[\![F]\!]$ equals true iff σ represents a possible execution of the algorithm. If G is a temporal formula, then $F \Rightarrow G$ is valid iff $\sigma[\![F \Rightarrow G]\!]$ equals true for every behavior σ. Since $\sigma[\![F \Rightarrow G]\!]$ equals $\sigma[\![F]\!] \Rightarrow \sigma[\![G]\!]$, validity of $F \Rightarrow G$ means that every behavior representing a possible execution of the algorithm satisfies G. In other words, $\models F \Rightarrow G$ asserts that the algorithm represented by F satisfies property G.

In Section 5.6, we give rules for proving temporal formulas. As usual, soundness of the rules means that every provable formula is valid—that is, $\vdash F$ implies $\models F$ for any temporal formula F.

4. THE RAW LOGIC

4.1 Actions as Temporal Formulas

The *Raw Temporal Logic of Actions*, or RTLA, is obtained by letting the elementary temporal formulas be actions. To define the semantics of RTLA formulas, we must define what it means for an action to be true on a behavior.

In Section 2.3, we defined the meaning $[\![\mathcal{A}]\!]$ of an action $\mathcal{A}$ to be a boolean-valued function that assigns the value $s[\![\mathcal{A}]\!]t$ to the pair of states s, t. We defined s, t to be an $\mathcal{A}$ step iff $s[\![\mathcal{A}]\!]t$ equals true. We now define $[\![\mathcal{A}]\!]$ to be true for a behavior iff the first pair of states in the behavior is an $\mathcal{A}$ step (Note 5). Formally,

$$\langle s_0, s_1, s_2, \ldots \rangle [\![\mathcal{A}]\!] \;\triangleq\; s_0 [\![\mathcal{A}]\!] s_1 \tag{9}$$

RTLA formulas are built up from actions using logical operators and the temporal operator $\Box$. Thus, if $\mathcal{A}$ is an action, then $\Box\mathcal{A}$ is an RTLA formula. Its meaning is computed as follows.

$$\begin{aligned}
&\langle s_0, s_1, s_2, \ldots \rangle [\![\Box\mathcal{A}]\!] \\
&\quad \equiv \forall n \in \mathsf{Nat} : \langle s_n, s_{n+1}, s_{n+2}, \ldots \rangle [\![\mathcal{A}]\!] && \text{by (6)} \\
&\quad \equiv \forall n \in \mathsf{Nat} : s_n [\![\mathcal{A}]\!] s_{n+1} && \text{by (9)}
\end{aligned}$$

In other words, a behavior satisfies $\Box\mathcal{A}$ iff every step of the behavior is an $\mathcal{A}$ step.

In Section 2.4, we observed that if P is a predicate, then $s[\![P]\!]t$ equals $s[\![P]\!]$. Therefore,

$$\begin{aligned}
\langle s_0, s_1, \ldots \rangle [\![P]\!] &\equiv s_0 [\![P]\!] \\
\langle s_0, s_1, \ldots \rangle [\![\Box P]\!] &\equiv \forall n \in \mathsf{Nat} : s_n [\![P]\!]
\end{aligned}$$

In other words, a behavior satisfies a predicate P iff the first state of the behavior satisfies P. A behavior satisfies $\Box P$ iff all states in the behavior satisfy P.

We will see that the raw logic RTLA is too powerful; it allows one to make assertions about behaviors that should not be assertable. We will define the formulas of TLA to be a subset of RTLA formulas.

4.2 Describing Programs with RTLA Formulas

We have defined the syntax and semantics of RTLA formulas, but have given no idea what RTLA is good for. We illustrate how RTLA can be used, by describing the simple Program 1 of Figure 1 as an RTLA formula. This program is written in a conventional language, using Dijkstra's **do** construct [Dijkstra 1976], with angle brackets enclosing operations that are assumed to be atomic. An execution of this program begins with x and y both zero, and repeatedly increments either x or y (in a single operation), choosing nondeterministically between them. We now define an RTLA formula Φ that represents this program, meaning that $\sigma[\![\Phi]\!]$ equals true iff the behavior σ represents a possible execution of Program 1.

The formula Φ is defined in Figure 2. The predicate $Init_\Phi$ asserts the initial condition, that x and y are both zero. The semantic meaning of action $\mathcal{M}_1$ is a relation between states asserting that the value of x in the new state is one greater than its value in the old state, and the value of y is the same in the old and new states. Thus, an $\mathcal{M}_1$ step represents an execution of the program's atomic operation of incrementing x. Similarly, an $\mathcal{M}_2$ step represents an execution of the program's other atomic operation, which increments y. The action $\mathcal{M}$ is defined to be the disjunction of $\mathcal{M}_1$ and $\mathcal{M}_2$, so an $\mathcal{M}$ step represents an execution of one program operation. Formula Φ is true of a behavior iff $Init_\Phi$ is true of the first state and every step is an $\mathcal{M}$ step. In other words, Φ asserts that the initial condition is true

```
var natural x, y = 0 ;
do ⟨ true → x := x + 1 ⟩
  []
   ⟨ true → y := y + 1 ⟩     od
```

Fig. 1. Program 1—a simple program, written in a conventional language.

$$\begin{array}{l} Init_\Phi \;\stackrel{\Delta}{=}\; (x = 0) \;\wedge\; (y = 0) \\ \mathcal{M}_1 \;\stackrel{\Delta}{=}\; (x' = x + 1) \;\wedge\; (y' = y) \qquad \mathcal{M}_2 \;\stackrel{\Delta}{=}\; (y' = y + 1) \;\wedge\; (x' = x) \\ \mathcal{M} \;\stackrel{\Delta}{=}\; \mathcal{M}_1 \vee \mathcal{M}_2 \\ \Phi \;\stackrel{\Delta}{=}\; Init_\Phi \;\wedge\; \Box\mathcal{M} \end{array}$$

Fig. 2. An RTLA formula Φ describing Program 1.

initially, and that every step of the behavior represents the execution of an atomic operation of the program. Clearly, a behavior satisfies Φ iff it represents a possible execution of Program 1 (Note 6).

There is nothing special about our choice of names, or in the particular way of writing Φ. There are many ways of writing equivalent logical formulas. Here are a couple of formulas that are equivalent to Φ.

$$(x = 0) \;\wedge\; \Box(\mathcal{M}_1 \vee \mathcal{M}_2) \;\wedge\; (y = 0)$$
$$Init_\Phi \;\wedge\; \Box((x' = x + 1) \;\vee\; (y' = y + 1)) \;\wedge\; \Box((x' = x) \;\vee\; (y' = y))$$

The particular way of defining Φ in Figure 2 was chosen to make the correspondence with Figure 1 obvious.

5. TLA

5.1 Adding Stuttering Steps

Formula Φ of Figure 2 is very simple. Unfortunately, it is too simple. In addition to steps in which x or y is incremented, a formula describing Program 1 should allow "stuttering" steps that leave both x and y unchanged.

To understand why stuttering steps are needed, consider a clock that displays hours and minutes. It is specified by a formula Π with two variables: h representing the hours display and m representing the minutes display. Now consider a clock that displays hours, minutes, and seconds; it is represented by a formula Ψ with three variables: h, m, and another variable s representing the seconds display. A clock that displays hours, minutes, and seconds should satisfy the specification Π of a clock that displays hours and minutes. (If we don't want the seconds display, we can always cover it up.) Hence, any behavior satisfying Ψ should satisfy Π. Behaviors satisfying Ψ contain sequences of 59 consecutive steps in which h and m do not change, so Π must allow such steps. From the point of view of a clock displaying only hours and minutes, steps in which h and m do not change are stuttering steps. In general, a specification Π should be invariant under stuttering, meaning that adding or removing stuttering steps from a behavior does not affect whether the behavior satisfies Π.

It is easy to modify formula Φ of Figure 2 so it asserts that every step is either an $\mathcal{M}$ step or a step that leaves x and y unchanged; the new definition is

$$\Phi \;\stackrel{\Delta}{=}\; Init_\Phi \;\wedge\; \Box(\mathcal{M} \;\vee\; ((x' = x) \wedge (y' = y))) \tag{10}$$

We now introduce notation that makes it easy to ensure that a formula allows stuttering steps. Two ordered pairs are equal iff their components are equal, so the conjunction $(x' = x) \wedge (y' = y)$ is equivalent to the single equality $\langle x', y'\rangle = \langle x, y\rangle$. The definition of priming a state function (formula (3)) allows us to write $\langle x', y'\rangle$

as $\langle x, y\rangle'$ (Note 7). For any action $\mathcal{A}$ and state function f, we let

$$[\mathcal{A}]_f \ \triangleq\ \mathcal{A} \vee (f' = f) \tag{11}$$

(The action $[\mathcal{A}]_f$ is read *square* $\mathcal{A}$ *sub* f.) Then

$$\begin{aligned}[\mathcal{M}]_{\langle x,\, y\rangle} &\equiv \mathcal{M} \vee (\langle x,\, y\rangle' = \langle x,\, y\rangle) \\ &\equiv \mathcal{M} \vee ((x' = x) \wedge (y' = y))\end{aligned}$$

and we can rewrite (10) as

$$\Phi \ \triangleq\ \mathit{Init}_\Phi \wedge \Box[\mathcal{M}]_{\langle x,\, y\rangle} \tag{12}$$

We define TLA to be the temporal logic whose elementary formulas are predicates and formulas of the form $\Box[\mathcal{A}]_f$, where $\mathcal{A}$ is an action and f a state function. Since these formulas are RTLA formulas, we have already defined their semantic meanings.

5.2 Adding Liveness

The formula Φ defined by (12) allows behaviors that start with Init_Φ true (x and y both zero) and never change x or y. Such behaviors do not represent acceptable executions of Program 1, so we must strengthen Φ to disallow them.

Formula Φ of (12) asserts that a behavior may not start in any state other than one satisfying Init_Φ and may never take any step other than a $[\mathcal{M}]_{\langle x,\, y\rangle}$ step. An assertion that something may never happen is called a *safety* property. An assertion that something eventually does happen is called a *liveness* property. (Safety and liveness have been defined formally by Alpern and Schneider [1985].) The formula $\mathit{Init}_\Phi \wedge \Box[\mathcal{M}]_{\langle x,\, y\rangle}$ is a safety property. To complete the description of Program 1, we need an additional liveness property asserting that the program keeps going.

By Dijkstra's semantics for his **do** construct, the liveness property for Program 1 should assert only that the program never terminates. In other words, Dijkstra would require that a behavior must contain infinitely many steps that increment x or y. This property is expressed by the RTLA formula $\Box\Diamond\mathcal{M}$, which asserts that there are infinitely many $\mathcal{M}$ steps. Dijkstra would have us define Φ by

$$\Phi \ \triangleq\ \mathit{Init}_\Phi \ \wedge\ \Box[\mathcal{M}]_{\langle x,\, y\rangle} \ \wedge\ \Box\Diamond\mathcal{M} \tag{13}$$

However, the example becomes more interesting if we add the fairness requirement that both x and y must be incremented infinitely often. (Dijkstra's definition would allow an execution in which one variable is incremented infinitely often while the other is incremented only a finite number of times.) Since we are not fettered by the dictates of conventional programming languages, we will adopt this stronger liveness requirement. The formula Φ representing the program with this fairness requirement is

$$\Phi \ \triangleq\ \mathit{Init}_\Phi \ \wedge\ \Box[\mathcal{M}]_{\langle x,\, y\rangle} \ \wedge\ \Box\Diamond\mathcal{M}_1 \ \wedge\ \Box\Diamond\mathcal{M}_2 \tag{14}$$

Formulas (13) and (14) are RTLA formulas, but not TLA formulas. An action $\mathcal{A}$ can appear in a TLA formula only in the form $\Box[\mathcal{A}]_f$ (unless $\mathcal{A}$ is a predicate), so $\Box\Diamond\mathcal{M}_1$ and $\Box\Diamond\mathcal{M}_2$ are not TLA formulas. We now rewrite them as TLA formulas.

Let $\mathcal{A}$ be any action and f any state function. Then $\neg\mathcal{A}$ is also an action, so $\neg\Box[\neg\mathcal{A}]_f$ is a TLA formula. Applying our definitions gives

$$\begin{aligned}\neg\Box[\neg\mathcal{A}]_f &\equiv \Diamond\neg[\neg\mathcal{A}]_f && \text{by (7), which implies } \neg\Box\ldots\equiv\Diamond\neg\ldots\\ &\equiv \Diamond\neg(\neg\mathcal{A}\lor(f'=f)) && \text{by (11)}\\ &\equiv \Diamond(\mathcal{A}\land(f'\neq f)) && \text{by simple logic}\end{aligned}$$

We define the action $\langle\mathcal{A}\rangle_f$ (read *angle* $\mathcal{A}$ *sub* f) by

$$\langle\mathcal{A}\rangle_f \triangleq \mathcal{A}\land(f'\neq f) \tag{15}$$

The calculation above shows that $\Diamond\langle\mathcal{A}\rangle_f$ equals $\neg\Box[\neg\mathcal{A}]_f$, so it is a TLA formula (Note 8).

Since incrementing a variable changes its value, both $\mathcal{M}_1$ and $\mathcal{M}_2$ imply $\langle x, y\rangle' \neq \langle x, y\rangle$ (Note 9). Hence, $\mathcal{M}_1$ is equivalent to $\langle\mathcal{M}_1\rangle_{\langle x, y\rangle}$, and $\mathcal{M}_2$ is equivalent to $\langle\mathcal{M}_2\rangle_{\langle x, y\rangle}$. We can therefore rewrite Φ as a TLA formula as follows.

$$\Phi \triangleq \mathit{Init}_\Phi \;\land\; \Box[\mathcal{M}]_{\langle x, y\rangle} \;\land\; \Box\Diamond\langle\mathcal{M}_1\rangle_{\langle x, y\rangle} \;\land\; \Box\Diamond\langle\mathcal{M}_2\rangle_{\langle x, y\rangle} \tag{16}$$

5.3 Fairness

Using arbitrary liveness properties like $\Box\Diamond\langle\mathcal{M}_1\rangle_{\langle x, y\rangle}$ to express fairness requirements is dangerous because it can add unexpected safety properties. For example, conjoining the liveness property $\Box\Diamond(x = 0)$, which asserts that x infinitely often equals 0, to $\mathit{Init}_\Phi \land \Box[\mathcal{M}]_{\langle x, y\rangle}$ implies the additional safety property that the value of x never changes. Accidentally adding safety properties in this way is a common source of errors in temporal logic specifications. We avoid such errors by expressing liveness in terms of fairness.

Fairness means that if a certain operation is possible, then the program must eventually execute it. The fairness requirements for concurrent algorithms can be expressed in terms of *weak fairness* and *strong fairness* conditions. We first define weak and strong fairness informally, then translate the informal definitions into TLA formulas.

Weak fairness asserts that an operation must be executed if it remains possible to do so for a long enough time. "Long enough" means until the operation is executed, so weak fairness asserts that eventually the operation must either be executed or become impossible to execute—perhaps only briefly. A naive temporal logic translation is

weak fairness: $(\Diamond \text{ executed}) \lor (\Diamond \text{ impossible})$

Strong fairness asserts that the operation must be executed if it is often enough possible to do so. Interpreting "often enough" to mean infinitely often, strong fairness asserts that either the operation is eventually executed, or its execution is not infinitely often possible. Not infinitely often possible is the same as eventually always impossible (because (7) implies $\neg\Box\Diamond\ldots\equiv\Diamond\Box\neg\ldots$), so we get

strong fairness: $(\Diamond \text{ executed}) \lor (\Diamond\Box \text{ impossible})$

These two temporal formulas assert fairness at "time zero", but we want fairness to hold at all times. The correct formulas are therefore

weak fairness: $\Box((\Diamond \text{ executed}) \lor (\Diamond \text{ impossible}))$
strong fairness: $\Box((\Diamond \text{ executed}) \lor (\Diamond\Box \text{ impossible}))$

Temporal logic reasoning, using either the axioms in Section 5.6 or the semantic definitions of $\Box$ and $\Diamond$, shows that these conditions are equivalent to

$$\begin{array}{ll} \textit{weak fairness}: & (\Box\Diamond \text{ executed}) \lor (\Box\Diamond \text{ impossible}) \\ \textit{strong fairness}: & (\Box\Diamond \text{ executed}) \lor (\Diamond\Box \text{ impossible}) \end{array}$$

To formalize these definitions, we must define "executed" and "impossible".

In Program 1, execution of the operation $x := x + 1$ corresponds to an $\mathcal{M}_1$ step in the behavior. To obtain a TLA formula, the "$\Diamond$ executed" for this operation must be expressed as $\Diamond\langle\mathcal{M}_1\rangle_{\langle x,\, y\rangle}$. In general, "$\Diamond$ executed" will be expressed as $\Diamond\langle\mathcal{A}\rangle_f$, where $\mathcal{A}$ is the action that corresponds to an execution of the operation, and f is an n-tuple of relevant variables. Recall that an $\langle\mathcal{A}\rangle_f$ step is an $\mathcal{A}$ step that changes the value of f. Steps that do not change the values of any relevant variables might as well not have occurred, so there is no need to consider them as representing operation executions.

We now define "impossible". Executing an operation means taking an $\langle\mathcal{A}\rangle_f$ step for some action $\mathcal{A}$ and state function f. It is possible to take such a step iff *Enabled* $\langle\mathcal{A}\rangle_f$ is true. Thus, *Enabled* $\langle\mathcal{A}\rangle_f$ asserts that it is possible to execute the operation represented by the action $\langle\mathcal{A}\rangle_f$, so "impossible" is $\neg\textit{Enabled}\ \langle\mathcal{A}\rangle_f$. Weak fairness and strong fairness are therefore expressed by the two formulas

$$\mathrm{WF}_f(\mathcal{A}) \triangleq (\Box\Diamond\langle\mathcal{A}\rangle_f) \lor (\Box\Diamond\neg\textit{Enabled}\ \langle\mathcal{A}\rangle_f) \tag{17}$$

$$\mathrm{SF}_f(\mathcal{A}) \triangleq (\Box\Diamond\langle\mathcal{A}\rangle_f) \lor (\Diamond\Box\neg\textit{Enabled}\ \langle\mathcal{A}\rangle_f) \tag{18}$$

Since $\Diamond\Box F$ implies $\Box\Diamond F$ for any F, the strong fairness condition $\mathrm{SF}_f(\mathcal{A})$ implies the weak fairness condition $\mathrm{WF}_f(\mathcal{A})$.

The pair of formulas $\textit{Init} \land \Box[\mathcal{N}]_v$, F is said to be *machine closed* if conjoining F to $\textit{Init} \land \Box[\mathcal{N}]_v$ introduces no additional safety properties. (In this case, we often say that $\textit{Init} \land \Box[\mathcal{N}]_v \land F$ is machine closed.) We avoid accidentally adding safety properties by writing machine-closed specifications. It can be shown that if F is the conjunction of fairness conditions of the form $\mathrm{WF}_f(\mathcal{A})$ and/or $\mathrm{SF}_f(\mathcal{A})$, where each $\langle\mathcal{A}\rangle_f$ implies $\mathcal{N}$, then $\textit{Init} \land \Box[\mathcal{N}]_v \land F$ is machine closed [Abadi and Lamport 1992].

5.4 Rewriting the Fairness Requirement

We now rewrite the property $\Box\Diamond\langle\mathcal{M}_1\rangle_{\langle x,\, y\rangle} \land \Box\Diamond\langle\mathcal{M}_2\rangle_{\langle x,\, y\rangle}$ in terms of fairness conditions. An $\langle\mathcal{M}_1\rangle_{\langle x,\, y\rangle}$ step is one that increments x by one, leaves y unchanged, and changes the value of $\langle x,\, y\rangle$. It is always possible to take a step that adds one to x and leaves y unchanged, and adding one to a number changes it. Hence, *Enabled* $\langle\mathcal{M}_1\rangle_{\langle x,\, y\rangle}$ equals true throughout any execution of Program 1 (Note 10). Since $\Box\neg\mathsf{true}$ equals false, both $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$ and $\mathrm{SF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$ equal $\Box\Diamond\langle\mathcal{M}_1\rangle_{\langle x,\, y\rangle}$. Similarly, $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_2)$ and $\mathrm{SF}_{\langle x,\, y\rangle}(\mathcal{M}_2)$ both equal $\Box\Diamond\langle\mathcal{M}_2\rangle_{\langle x,\, y\rangle}$. We can therefore rewrite the definition (16) of Φ as shown in Figure 3.

Suppose we wanted the weaker liveness condition that execution never terminates, so the program is described by the RTLA formula (13). The same argument as for $\mathcal{M}_1$ and $\mathcal{M}_2$ shows that $\Box\Diamond\langle\mathcal{M}\rangle_{\langle x,\, y\rangle}$ equals $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M})$. Therefore, Program 1 with this weaker liveness condition is described by the TLA formula $\textit{Init}_\Phi \land \Box[\mathcal{M}]_{\langle x,\, y\rangle} \land \mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M})$.

The actions $\langle\mathcal{M}_1\rangle_{\langle x,\, y\rangle}$, $\langle\mathcal{M}_2\rangle_{\langle x,\, y\rangle}$, and $\langle\mathcal{M}\rangle_{\langle x,\, y\rangle}$ all imply $\mathcal{M}$. Hence neither of the liveness conditions $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_1) \land \mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_2)$ and $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$ add safety properties to $Init_{\Phi} \wedge \Box[\mathcal{M}]_{\langle x, y \rangle}$.

5.5 Examining Formula Φ

TLA formulas that represent programs can always be written in the same form as Φ of Figure 3—that is, as a conjunction $Init \wedge \Box[\mathcal{N}]_f \wedge F$, where

- *Init* is a predicate specifying the initial values of variables.
- $\mathcal{N}$ is the program's *next-state relation*, the action whose steps represent executions of individual atomic operations.
- f is the n-tuple of all flexible variables.
- F is the conjunction of formulas of the form $\mathrm{WF}_f(\mathcal{A})$ and/or $\mathrm{SF}_f(\mathcal{A})$, where $\mathcal{A}$ is an action representing some subset of the program's atomic operations.

We now examine the behaviors that satisfy formula Φ. Let

$$\langle\!\langle x \triangleq 7,\ y \triangleq -10,\ z \triangleq \text{“}\mathsf{abc}\text{”},\ \ldots \rangle\!\rangle$$

denote a state s such that $s[\![x]\!] = 7$, $s[\![y]\!] = -10$, and $s[\![z]\!] = \text{“}\mathsf{abc}\text{”}$. (The “…” indicates that the value of $s[\![v]\!]$ is left unspecified for all other variables v.) A behavior that satisfies Φ begins in a state satisfying $Init_{\Phi}$, and consists of a sequence of $[\mathcal{M}]_{\langle x, y \rangle}$ steps—ones that are either $\mathcal{M}$ steps or else leave x and y unchanged. One such behavior is

$$\begin{array}{l}
\langle\!\langle\ x \triangleq 0,\ y \triangleq 0,\ z \triangleq \text{“}\mathsf{abc}\text{”} \ \ldots\ \rangle\!\rangle \\
\langle\!\langle\ x \triangleq 1,\ y \triangleq 0,\ z \triangleq 14 \quad \ldots\ \rangle\!\rangle \\
\langle\!\langle\ x \triangleq 2,\ y \triangleq 0,\ z \triangleq \mathsf{Nat} \quad \ldots\ \rangle\!\rangle \\
\langle\!\langle\ x \triangleq 2,\ y \triangleq 0,\ z \triangleq -20 \quad \ldots\ \rangle\!\rangle \\
\langle\!\langle\ x \triangleq 2,\ y \triangleq 1,\ z \triangleq \sqrt{2} \quad \ldots\ \rangle\!\rangle \\
\quad \vdots
\end{array}$$

Observe that Φ constrains only the values of x and y; it allows all other variables such as z to assume completely arbitrary values. Suppose Ψ is a formula describing a program that has no variables in common with Φ. Then a behavior satisfies $\Phi \wedge \Psi$ iff it represents an execution of both programs—that is, iff it describes a universe in which both Φ and Ψ are executed concurrently. Thus, $\Phi \wedge \Psi$ is the TLA formula representing the parallel composition of the two programs.

In general, parallel composition is represented in TLA by conjunction. For example, let

$$\begin{array}{lcl}
\Phi_1 & \triangleq & (x = 0) \ \wedge\ \Box[\mathcal{M}_1]_x \ \wedge\ \mathrm{WF}_x(\mathcal{M}_1) \\
\Phi_2 & \triangleq & (y = 0) \ \wedge\ \Box[\mathcal{M}_2]_y \ \wedge\ \mathrm{WF}_y(\mathcal{M}_2)
\end{array}$$

$$\begin{array}{l}
Init_{\Phi} \ \triangleq\ (x = 0) \ \wedge\ (y = 0) \\
\mathcal{M}_1 \ \triangleq\ (x' = x + 1) \ \wedge\ (y' = y) \qquad \mathcal{M}_2 \ \triangleq\ (y' = y + 1) \ \wedge\ (x' = x) \\
\mathcal{M} \ \triangleq\ \mathcal{M}_1 \vee \mathcal{M}_2 \\
\Phi \ \triangleq\ Init_{\Phi} \ \wedge\ \Box[\mathcal{M}]_{\langle x, y \rangle} \ \wedge\ \mathrm{WF}_{\langle x, y \rangle}(\mathcal{M}_1) \ \wedge\ \mathrm{WF}_{\langle x, y \rangle}(\mathcal{M}_2)
\end{array}$$

Fig. 3. The TLA formula Φ describing Program 1.

A straightforward calculation shows that $[\mathcal{M}_1]_x \wedge [\mathcal{M}_2]_y$ is equivalent to $[\mathcal{M}]_{\langle x, y \rangle}$, and temporal logic reasoning (using the axioms of Section 5.6) then shows that Φ is equivalent to $\Phi_1 \wedge \Phi_2$. Formulas Φ_1 and Φ_2 are the specifications of the two processes forming Program 1. This example illustrates a general method for decomposing the specification of a multiprocess program as the conjunction of the specifications of its processes [Abadi and Lamport 1993].

The observation that a single behavior can represent an execution of two or more noninteracting programs explains why we represent terminating as well as nonterminating executions by infinite behaviors. Termination of a program means that it has stopped; it does not mean that the entire universe has come to a halt. A terminating execution is represented by a behavior in which eventually all of the program's variables stop changing.

It is unusual in computer science for the semantics of a formula describing a program with variables x and y to involve other variables such as z that appear nowhere in the program. One of the keys to TLA's simplicity is that its semantics rests on a single, infinite set of variables—not on a different set of variables for each program. Thus, in TLA as in elementary logic, we can take the conjunction $F \wedge G$ of any formulas F and G—not just of formulas with properly matching variable declarations.

5.6 Simple TLA

We now complete the definition of *Simple TLA* by adding one more bit of notation. (The full logic, containing quantification, is introduced in Section 8.) It is convenient to define the action *Unchanged* f, for f a state function, by

$$\mathit{Unchanged}\ f \ \triangleq\ f' = f$$

Thus, an *Unchanged* f step is one in which the value of f does not change.

The syntax and semantics of Simple TLA, along with the additional notation we use to write TLA formulas, are all summarized in Figure 4. This figure explains all you need to know to understand TLA formulas such as formula Φ of Figure 3.

A logic contains not only syntax and semantics, but also rules for proving theorems. Figure 5 lists all the axioms and proof rules we need for proving simple TLA formulas.[3]

The rules of simple temporal logic are used to derive temporal tautologies—formulas that are true regardless of the meanings of their elementary formulas. Rule STL1 encompasses the rules of ordinary logic, such as *modus ponens* (Note 11). The Lattice Rule assumes a (possibly infinite) set S and a mapping that assigns a TLA formula H_{c} to each element c of S. A partial order $\succ$ on S is *well-founded* iff there exists no infinite descending chain $\mathsf{c}_1 \succ \mathsf{c}_2 \succ \ldots$ with all the c_i in S. This rule permits the formalization of counting-down arguments, such as the ones traditionally used to prove termination of sequential programs.

Rules STL1–STL6, the Lattice Rule, and the basic rules TLA1 and TLA2 form a relatively complete proof system for reasoning about algorithms in TLA. Roughly speaking, this means that every valid TLA formula that we must prove to verify properties of algorithms would be provable from these rules if we could prove all

[3] A proof rule $\frac{F, G}{H}$ asserts that $\vdash F$ and $\vdash G$ imply $\vdash H$. We use the term "rule" for both axioms and proof rules, since an axiom may be viewed as a proof rule with no hypotheses.

Syntax

$$\begin{array}{lcl}
\langle formula \rangle & \triangleq & \langle predicate \rangle \;\mid\; \Box[\langle action \rangle]_{\langle state\ function \rangle} \;\mid\; \neg\langle formula \rangle \\
& & \quad \mid\; \langle formula \rangle \wedge \langle formula \rangle \;\mid\; \Box\, \langle formula \rangle \\
\langle action \rangle & \triangleq & \text{boolean-valued expression containing constant symbols,} \\
& & \quad \text{variables, and primed variables} \\
\langle predicate \rangle & \triangleq & \langle action \rangle \text{ with no primed variables } \mid\; \mathit{Enabled}\ \langle action \rangle \\
\langle state\ function \rangle & \triangleq & \text{nonboolean expression containing constant symbols and variables}
\end{array}$$

Semantics

$$\begin{array}{lcllcl}
s[\![f]\!] & \triangleq & f(\forall\ `v\text{'} : s[\![v]\!]/v) & \sigma[\![F \wedge G]\!] & \triangleq & \sigma[\![F]\!] \wedge \sigma[\![G]\!] \\
s[\![\mathcal{A}]\!]t & \triangleq & \mathcal{A}(\forall\ `v\text{'} : s[\![v]\!]/v,\ t[\![v]\!]/v') & \sigma[\![\neg F]\!] & \triangleq & \neg\sigma[\![F]\!] \\
\models \mathcal{A} & \triangleq & \forall s, t \in \mathbf{St} : s[\![\mathcal{A}]\!]t & \models F & \triangleq & \forall \sigma \in \mathbf{St}^{\infty} : \sigma[\![F]\!]
\end{array}$$

$$\begin{array}{lcl}
s[\![\mathit{Enabled}\ \mathcal{A}]\!] & \triangleq & \exists t \in \mathbf{St} : s[\![\mathcal{A}]\!]t \\
\langle s_0, s_1, \ldots \rangle[\![\Box F]\!] & \triangleq & \forall n \in \mathsf{Nat} : \langle s_n, s_{n+1}, \ldots \rangle[\![F]\!] \\
\langle s_0, s_1, \ldots \rangle[\![\mathcal{A}]\!] & \triangleq & s_0[\![\mathcal{A}]\!]s_1
\end{array}$$

Additional notation

$$\begin{array}{lcllcl}
p' & \triangleq & p(\forall\ `v\text{'} : v'/v) & \Diamond F & \triangleq & \neg\Box\neg F \\
[\mathcal{A}]_f & \triangleq & \mathcal{A} \vee (f' = f) & F \leadsto G & \triangleq & \Box(F \Rightarrow \Diamond G) \\
\langle \mathcal{A} \rangle_f & \triangleq & \mathcal{A} \wedge (f' \neq f) & \mathrm{WF}_f(\mathcal{A}) & \triangleq & \Box\Diamond\langle \mathcal{A} \rangle_f \vee \Box\Diamond\neg \mathit{Enabled}\ \langle \mathcal{A} \rangle_f \\
\mathit{Unchanged}\ f & \triangleq & f' = f & \mathrm{SF}_f(\mathcal{A}) & \triangleq & \Box\Diamond\langle \mathcal{A} \rangle_f \vee \Diamond\Box\neg \mathit{Enabled}\ \langle \mathcal{A} \rangle_f
\end{array}$$

where f is a $\langle state\ function \rangle$ | s, s_0, s_1, $\ldots$ are states
$\mathcal{A}$ is an $\langle action \rangle$ | σ is a behavior
F and G are $\langle formula \rangle$s | $(\forall\ `v\text{'} : \ldots/v,\ \ldots/v')$ denotes substitution
p is a $\langle state\ function \rangle$ or $\langle predicate \rangle$ | for all variables v

Fig. 4. Simple TLA.

valid action formulas. (This is analogous to the traditional relative completeness results for program verification, which assume provability of all valid predicates [Apt 1981].) A more precise statement of this result is given in Section 8.3.2 below (Note 12).

A complete proof system is not necessarily a convenient one. For practical reasoning, STL1–STL6 should be augmented with some useful temporal tautologies like

$$\vdash (\Box F) \wedge (\Diamond G) \;\Rightarrow\; \Diamond(F \wedge G)$$

With practice, such simple tautologies become as obvious as the ordinary laws of propositional logic. They are usually taken for granted in hand proofs, and practical decision procedures exist for checking them mechanically [Burch et al. 1992]. The temporal operators $\Box$, $\Diamond$, and $\leadsto$ are standard [Manna and Pnueli 1991], so we will not discuss the rules of simple temporal logic.

Assuming simple temporal reasoning, we have found that TLA2 and the "additional rules" INV1–SF2 of Figure 5 provide a convenient system for all the proofs that arise in reasoning about programs with TLA. The overbars in rules WF2 and SF2 are explained in Section 8.3.3; for now, the reader can pretend that they are not there, obtaining special cases of the rules.

The Rules of Simple Temporal Logic

STL1. $$\frac{F \text{ provable by propositional logic}}{\Box F}$$

STL2. $\vdash \Box F \Rightarrow F$

STL3. $\vdash \Box\Box F \equiv \Box F$

STL4. $$\frac{F \Rightarrow G}{\Box F \Rightarrow \Box G}$$

STL5. $\vdash \Box(F \land G) \equiv (\Box F) \land (\Box G)$

STL6. $\vdash (\Diamond\Box F) \land (\Diamond\Box G) \equiv \Diamond\Box(F \land G)$

LATTICE. $\succ$ a well-founded partial order on a set S
$$\frac{F \land (\mathsf{c} \in \mathsf{S}) \Rightarrow (H_{\mathsf{c}} \leadsto (G \lor \exists \mathsf{d} \in \mathsf{S} : (\mathsf{c} \succ \mathsf{d}) \land H_{\mathsf{d}}))}{F \Rightarrow ((\exists \mathsf{c} \in \mathsf{S} : H_{\mathsf{c}}) \leadsto G)}$$

The Basic Rules of TLA

TLA1. $$\frac{P \land (f' = f) \Rightarrow P'}{\Box P \equiv P \land \Box[P \Rightarrow P']_f}$$

TLA2. $$\frac{P \land [\mathcal{A}]_f \Rightarrow Q \land [\mathcal{B}]_g}{\Box P \land \Box[\mathcal{A}]_f \Rightarrow \Box Q \land \Box[\mathcal{B}]_g}$$

Additional Rules

INV1. $$\frac{I \land [\mathcal{N}]_f \Rightarrow I'}{I \land \Box[\mathcal{N}]_f \Rightarrow \Box I}$$

INV2. $\vdash \Box I \Rightarrow (\Box[\mathcal{N}]_f \equiv \Box[\mathcal{N} \land I \land I']_f)$

WF1.
$$\frac{\begin{array}{l} P \land [\mathcal{N}]_f \Rightarrow (P' \lor Q') \\ P \land \langle \mathcal{N} \land \mathcal{A} \rangle_f \Rightarrow Q' \\ P \Rightarrow \mathit{Enabled}\ \langle \mathcal{A} \rangle_f \end{array}}{\Box[\mathcal{N}]_f \land \mathrm{WF}_f(\mathcal{A}) \Rightarrow (P \leadsto Q)}$$

WF2.
$$\frac{\begin{array}{l} \langle \mathcal{N} \land \mathcal{B} \rangle_f \Rightarrow \langle \overline{\mathcal{M}} \rangle_{\overline{g}} \\ P \land P' \land \langle \mathcal{N} \land \mathcal{A} \rangle_f \land \overline{\mathit{Enabled}\ \langle \mathcal{M} \rangle_g} \Rightarrow \mathcal{B} \\ P \land \overline{\mathit{Enabled}\ \langle \mathcal{M} \rangle_g} \Rightarrow \mathit{Enabled}\ \langle \mathcal{A} \rangle_f \\ \Box[\mathcal{N} \land \neg\mathcal{B}]_f \land \mathrm{WF}_f(\mathcal{A}) \land \Box F \\ \quad \land \Diamond\Box\overline{\mathit{Enabled}\ \langle \mathcal{M} \rangle_g} \Rightarrow \Diamond\Box P \end{array}}{\Box[\mathcal{N}]_f \land \mathrm{WF}_f(\mathcal{A}) \land \Box F \Rightarrow \overline{\mathrm{WF}_g(\mathcal{M})}}$$

SF1.
$$\frac{\begin{array}{l} P \land [\mathcal{N}]_f \Rightarrow (P' \lor Q') \\ P \land \langle \mathcal{N} \land \mathcal{A} \rangle_f \Rightarrow Q' \\ \Box P \land \Box[\mathcal{N}]_f \land \Box F \Rightarrow \Diamond \mathit{Enabled}\ \langle \mathcal{A} \rangle_f \end{array}}{\Box[\mathcal{N}]_f \land \mathrm{SF}_f(\mathcal{A}) \land \Box F \Rightarrow (P \leadsto Q)}$$

SF2.
$$\frac{\begin{array}{l} \langle \mathcal{N} \land \mathcal{B} \rangle_f \Rightarrow \langle \overline{\mathcal{M}} \rangle_{\overline{g}} \\ P \land P' \land \langle \mathcal{N} \land \mathcal{A} \rangle_f \Rightarrow \mathcal{B} \\ P \land \overline{\mathit{Enabled}\ \langle \mathcal{M} \rangle_g} \Rightarrow \mathit{Enabled}\ \langle \mathcal{A} \rangle_f \\ \Box[\mathcal{N} \land \neg\mathcal{B}]_f \land \mathrm{SF}_f(\mathcal{A}) \land \Box F \\ \quad \land \Box\Diamond\overline{\mathit{Enabled}\ \langle \mathcal{M} \rangle_g} \Rightarrow \Diamond\Box P \end{array}}{\Box[\mathcal{N}]_f \land \mathrm{SF}_f(\mathcal{A}) \land \Box F \Rightarrow \overline{\mathrm{SF}_g(\mathcal{M})}}$$

where F, G, H_{c} are TLA formulas; P, Q, I are predicates; $\mathcal{A}$, $\mathcal{B}$, $\mathcal{N}$, $\mathcal{M}$ are actions; f, g are state functions

Fig. 5. The axioms and proof rules of Simple TLA.

The validity of these rules can be proved rigorously with the raw logic RTLA. Rules STL1–STL6 are valid when F and G are arbitrary RTLA formulas, not just TLA formulas. The validity of TLA1–SF2 can be proved using STL1–STL6 and the RTLA rule $\Box P \equiv P \wedge \Box(P \Rightarrow P')$. The validity of this rule follows easily from the semantic definitions of $\Box$ and $'$ (prime). We leave the rigorous proofs as an exercise for the reader; instead, we give informal justifications. Sections 6 and 7 illustrate how the rules are used.

Rule TLA1 provides an induction principle for proving the formula $\Box P$. It asserts the obvious fact that a predicate P is always true iff P holds initially and every step starting with P true leaves P true.

Rule TLA2 follows immediately from the validity of STL4 and STL5 for RTLA formulas. (We have to use RTLA because $\Box(P \wedge [\mathcal{A}]_f)$ is not a TLA formula.)

Rule INV1 is used to prove that a program satisfies an invariance property $\Box I$. The hypothesis asserts that a $[\mathcal{N}]_f$ step cannot falsify I. The conclusion asserts the obvious consequence that if I is true initially and every step is a $[\mathcal{N}]_f$ step, then I is always true.

Rule WF1 is used to deduce a leads-to property $P \leadsto Q$ from a weak fairness condition $\mathrm{WF}_f(\mathcal{A})$. It can be applied when an $\mathcal{A}$ step that starts with P true makes Q true. To prove the validity of the conclusion, we assume that every step is a $[\mathcal{N}]_f$ step and that $\mathrm{WF}_f(\mathcal{A})$ holds, and we prove $P \leadsto Q$. It suffices to derive a contradiction by assuming that P is true at some time n and Q is false then and at all later times. Since every step is a $[\mathcal{N}]_f$ step and Q is false from time n on, the first hypothesis implies that P is true from time n on. The third hypothesis then implies that *Enabled* $\langle\mathcal{A}\rangle_f$ is true from time n on. Hence, $\mathrm{WF}_f(\mathcal{A})$ implies that infinitely many $\mathcal{A}$ steps occur. Any such step occurring after time n starts with P true, and the second hypothesis implies that the step makes Q true. This contradicts the assumption that Q remains false after time n, proving the validity of the rule.

Rule WF2 is used to deduce one weak fairness condition from another. We deduce $\mathrm{WF}_g(\mathcal{M})$ from $\mathrm{WF}_f(\mathcal{A})$ by finding an action $\mathcal{B}$ such that every $\mathcal{B}$ step is an $\mathcal{M}$ step and, if $\mathcal{M}$ remains forever enabled, then eventually every $\mathcal{A}$ step is a $\mathcal{B}$ step. (We ignore the overbars.) To prove the validity of WF2, we first observe that since $\mathrm{WF}_g(\mathcal{M})$ equals $\Box\Diamond\neg \mathit{Enabled}\ \langle\mathcal{M}\rangle_g \vee \Box\Diamond\langle\mathcal{M}\rangle_g$ and $\Diamond$ equals $\neg\Box\neg$, we can rewrite the conclusion as follows.

$$\Box[\mathcal{N}]_f \wedge \mathrm{WF}_f(\mathcal{A}) \wedge \Box F \wedge \Diamond\Box \mathit{Enabled}\ \langle\mathcal{M}\rangle_g \;\Rightarrow\; \Box\Diamond\langle\mathcal{M}\rangle_g$$

It therefore suffices to obtain a contradiction by assuming that $\Box[\mathcal{N}]_f \wedge \mathrm{WF}_f(\mathcal{A}) \wedge \Box F \wedge \Diamond\Box \mathit{Enabled}\ \langle\mathcal{M}\rangle_g$ holds and only finitely many $\langle\mathcal{M}\rangle_g$ steps occur. Since $[\mathcal{N}]_f \wedge \langle\mathcal{B}\rangle_f$ equals $\langle\mathcal{N} \wedge \mathcal{B}\rangle_f$, it follows from the first hypothesis that only finitely many $\langle\mathcal{B}\rangle_f$ steps can occur. Hence, there must eventually be a time after which no more $\langle\mathcal{B}\rangle_f$ steps occur, so every further step is a $[\mathcal{N} \wedge \neg\mathcal{B}]_f$ step. By the fourth hypothesis, there must then be a time at which P becomes true forever. The third hypothesis then implies that *Enabled* $\langle\mathcal{A}\rangle_f$ eventually becomes true forever, so $\mathrm{WF}_f(\mathcal{A})$ implies that there are infinitely many $\langle\mathcal{A}\rangle_f$ steps. The second hypothesis then implies that there are infinitely many $\langle\mathcal{B}\rangle_f$ steps, which is the required contradiction.

Rules SF1 and SF2 are the analogs of WF1 and WF2 for strong fairness. We omit their justifications, which are similar to those of WF1 and WF2.

6. PROVING SIMPLE PROPERTIES OF PROGRAMS

Having expressed Program 1 of Figure 1 as the TLA formula Φ of Figure 3, we now consider how to express and prove properties of such a program. A property is expressed by a TLA formula F. The assertion "program Φ has property F" is expressed in TLA by the validity of the formula $\Phi \Rightarrow F$, which asserts that every behavior satisfying Φ satisfies F. We consider two popular classes of properties, invariance and eventuality.

6.1 Invariance Properties

6.1.1 *Definition.* An invariance property is expressed by a TLA formula $\Box P$, where P is a predicate. Examples of invariance properties include

partial correctness. P asserts that if the program has terminated, then the answer is correct.

deadlock freedom. P asserts that the program is not deadlocked.

mutual exclusion. P asserts that at most one process is in its critical section.

Invariance properties are proved with rule INV1 of Figure 5.

6.1.2 *An Example: Type Correctness.* One part of the program in Figure 1 does not correspond to anything in Figure 3—the type declaration of the variables x and y. Such a declaration is not needed because type-correctness is an invariance property of the program, asserting that x and y are always natural numbers. We illustrate invariance proofs by proving type correctness of Program 1. Type correctness is expressed formally as $\Phi \Rightarrow \Box T$, where

$$T \;\triangleq\; (x \in \mathsf{Nat}) \wedge (y \in \mathsf{Nat}) \tag{19}$$

Rule INV1 tells us that we must prove

$$\mathit{Init}_\Phi \Rightarrow T \tag{20}$$

$$T \wedge [\mathcal{M}]_{\langle x,\, y\rangle} \Rightarrow T' \tag{21}$$

from which we deduce $\Phi \Rightarrow \Box T$ as follows

$$\begin{array}{rll} \Phi & \Rightarrow \mathit{Init}_\Phi \wedge \Box[\mathcal{M}]_{\langle x,\, y\rangle} & \text{by definition of } \Phi \text{ (Figure 3)} \\ & \Rightarrow T \wedge \Box[\mathcal{M}]_{\langle x,\, y\rangle} & \text{by (20)} \\ & \Rightarrow \Box T & \text{by (21) and INV1} \end{array}$$

The proof of (20) is trivial. The proof of (21) is quite simple, but we will sketch it to show how the structure of the formulas leads to a natural decomposition of the proof. First, we expand the definition of $[\mathcal{M}]_{\langle x,\, y\rangle}$.

$$\begin{array}{rll} [\mathcal{M}]_{\langle x,\, y\rangle} & \equiv \mathcal{M} \vee (\langle x,\, y\rangle' = \langle x,\, y\rangle) & \text{by (11)} \\ & \equiv \mathcal{M}_1 \vee \mathcal{M}_2 \vee (\langle x,\, y\rangle' = \langle x,\, y\rangle) & \text{by definition of } \mathcal{M} \end{array}$$

Since $[\mathcal{M}]_{\langle x,\, y\rangle}$ is the disjunction of three actions, the proof of (21) decomposes into the proof of three simpler formulas:

$$T \wedge \mathcal{M}_1 \;\Rightarrow\; T' \tag{22}$$

$$T \wedge \mathcal{M}_2 \;\Rightarrow\; T' \tag{23}$$

$$T \wedge (\langle x,\, y\rangle' = \langle x,\, y\rangle) \;\Rightarrow\; T' \tag{24}$$

We consider the proof of (22); the others are equally simple. First, we expand the definition of T'.

$$\begin{aligned} T' &\equiv ((x \in \mathsf{Nat}) \wedge (y \in \mathsf{Nat}))' && \text{by (19)} \\ &\equiv (x' \in \mathsf{Nat}) \wedge (y' \in \mathsf{Nat}) && \text{by (3)} \end{aligned}$$

The structure of T' as the conjunction of two actions leads to the decomposition of the proof of (22) into the proof of the two simpler formulas

$$T \wedge \mathcal{M}_1 \;\Rightarrow\; x' \in \mathsf{Nat} \tag{25}$$

$$T \wedge \mathcal{M}_1 \;\Rightarrow\; y' \in \mathsf{Nat} \tag{26}$$

The proof of (25) is

$$\begin{aligned} T \wedge \mathcal{M}_1 &\Rightarrow (x \in \mathsf{Nat}) \wedge (x' = x + 1) && \text{by definition of } T \text{ and } \mathcal{M}_1 \\ &\Rightarrow x' \in \mathsf{Nat} && \text{by properties of natural numbers} \end{aligned}$$

and the proof of (26) is equally trivial.

The purpose of this exercise in simple mathematics is to illustrate how “mechanical” the proof of this invariance property is. Rule INV1 tells us we must prove (20) and (21), and the structure of $[\mathcal{M}]_{\langle x,\, y\rangle}$ and T leads to the decomposition of those proofs into the verification of simple facts about natural numbers, such as $(x \in \mathsf{Nat}) \Rightarrow (x + 1 \in \mathsf{Nat})$.

6.1.3 *General Invariance Proofs.* The proof of $\Phi \Rightarrow \Box T$ was simple because T is an *invariant* of the action $[\mathcal{M}]_{\langle x,\, y\rangle}$, meaning that $T \wedge [\mathcal{M}]_{\langle x,\, y\rangle}$ implies T'. Therefore, $\Phi \Rightarrow \Box T$ could be proved by simply substituting T for I in rule INV1. For invariance properties $\Box P$ other than simple type correctness, P is usually not an invariant. In general, one proves that $\Box P$ is an invariance property of the program represented by the TLA formula $\mathit{Init} \wedge \Box[\mathcal{N}]_f \wedge F$ by finding a predicate I (the invariant) satisfying the three conditions

$$\mathit{Init} \Rightarrow I \tag{27}$$

$$I \Rightarrow P \tag{28}$$

$$I \wedge [\mathcal{N}]_f \Rightarrow I' \tag{29}$$

Rule INV1 and some simple temporal reasoning shows that (27)–(29) imply $\mathit{Init} \wedge \Box[\mathcal{N}]_f \Rightarrow \Box P$.

Creative thought is needed to find the invariant I. Once I is found, verifying (27)–(29) is a matter of mechanically applying the definitions and using the structure of the formulas to decompose the proofs, just as in the proof of $\Phi \Rightarrow \Box T$ above. The formulas Init, I, and $\mathcal{N}$ will usually be much more complicated than in the example, but the principle is the same.

Formulas (27)–(29) are assertions about predicates and actions; they are not temporal formulas. All the work in proving an invariance property is done in the realm of predicates and actions—expressions involving variables and primed variables that can be manipulated by ordinary mathematics. Temporal reasoning is used only to deduce $\mathit{Init} \wedge \Box[\mathcal{N}]_f \Rightarrow \Box P$ from (27)–(29). TLA is practical because it minimizes temporal reasoning, relying on ordinary, nontemporal reasoning whenever possible.

6.1.4 *More About Invariance Proofs.* Over the years, many methods have been proposed for proving invariance properties of programs, including Floyd's method

[Floyd 1967], Hoare logic [Hoare 1969], and the Owicki-Gries method [Owicki 1975]. All of these methods are essentially the same—when applied to the same program, they involve the same proof steps, though perhaps in different orders and with different notation. These methods can be described formally in TLA as applications of rule INV1. The advantage of TLA is that the proof method arises directly from the logic, without the need for proof rules based on a particular programming language.

We illustrate the advantage of working in a simple logic by considering the use of one invariance property to prove another. We have just proved $\Phi \Rightarrow \Box T$, the assertion that the program satisfies the invariance property $\Box T$. How can we use this fact when proving that the program satisfies a second invariance property $\Box P$? Some methods have a special rule saying that if a program satisfies $\Box T$, then one can pretend that T is true when reasoning about the program. (The "substitution axiom" of Unity [Chandy and Misra 1988] is such a rule.) In TLA, we use Rule INV2 of Figure 5. This rule implies that having proved $\Phi \Rightarrow \Box T$, we can rewrite the definition of Φ in Figure 3 as

$$\Phi \triangleq Init_{\Phi} \wedge \Box[\mathcal{M} \wedge T \wedge T']_{\langle x, y\rangle} \wedge \mathrm{WF}_{\langle x, y\rangle}(\mathcal{M}_1) \wedge \mathrm{WF}_{\langle x, y\rangle}(\mathcal{M}_2)$$

It follows that in proving Φ implies $\Box P$, instead of assuming every step to be a $[\mathcal{M}]_{\langle x, y\rangle}$ step, we can make the stronger assumption that every step is a $[\mathcal{M} \wedge T \wedge T']_{\langle x, y\rangle}$ step. More precisely, we can substitute $\mathcal{M} \wedge T \wedge T'$ instead of $\mathcal{M}$ for $\mathcal{N}$ in rule INV1, giving a stronger proof rule. This stronger rule is tantamount to "pretending T is true". The validity of this pretense follows directly from the logic; it is not an ad hoc rule.

6.1.5 *More About Types.* In TLA, variables have no types. Any variable can assume any value. Type-correctness of a program is a provable property, not a syntactic requirement as in strongly-typed programming languages. This has some subtle consequences. Consider the action $x' = x + 1$. Its meaning is a boolean-valued function on pairs of states. Suppose s and t are states that assign the values "abc" and 17 to x, respectively—that is, so $s[\![x]\!]$ equals "abc" and $t[\![x]\!]$ equals 17. Then $s[\![x' = x + 1]\!]t$ equals $17 =$ "abc" $+ 1$. But what is "abc" $+ 1$? Does it equal 17?

We don't know the answers to these questions, and we don't care. All we know is that "abc" $+ 1$ is some value. Since that value is either equal to or unequal to 17, the expression $17 =$ "abc" $+ 1$ is equal to either true or false. More precisely, we assume that $\mathsf{m} + \mathsf{n}$ is a value, for any values m and n. However, we have no rules for deducing anything about the value of $\mathsf{m} + \mathsf{n}$ except when m and n are numbers. In general, we assume that all operators such as $+$ are *total*—they are defined on all possible values. What we usually think of as the domain of an operator is just the set of values for which we know how to evaluate the operator. We know how to evaluate $\mathsf{m} + \mathsf{n}$ only when m and n are numbers, but it is defined (in the mathematical sense of being a meaningful expression) for all values m and n.

Since we can't deduce anything about the value "abc" $+ 1$, whatever we prove about an algorithm is true regardless of what that value is. If we can prove that the program is correct, then either it will never add 1 to "abc" (as in the case of Program 1), or else correctness does not depend on the result of that addition.

This approach may seem strange to computer scientists used to types in programming languages, but it captures the way mathematicians have reasoned for thousands of years. For example, mathematicians would say that the formula

$$(\mathsf{n} \in \mathsf{Nat}) \Rightarrow (\mathsf{n} + 1 > \mathsf{n}) \tag{30}$$

is true for all n. Substituting "abc" for n yields

$$(\text{“abc”} \in \mathsf{Nat}) \Rightarrow (\text{“abc”} + 1 > \text{“abc”})$$

This formula is true regardless of what "abc" $+ 1$ equals, and whether or not that value is greater than "abc", because "abc" $\in \mathsf{Nat}$ is false (Note 13). The formula is not meaningless or "type-incorrect" just because we don't know the value of "abc" $+ 1$.

There is one subtle pitfall raised by the absence of types in TLA. It is tempting to think that we can replace $x' = x + 1$ by $x = x' - 1$ in the TLA formula Φ describing Program 1. However, these two expressions need not be equivalent unless x and x' are both numbers. For example, if x equals 16, then $x' = x + 1$ is true only if x' equals 17. However, we don't know what "abc" $- 1$ equals, so it might equal 16. Hence, the formula $16 = x' - 1$ might be true when x' equals "abc" as well as when x' equals 17. We could not prove the property $\Phi \Rightarrow \Box(x \in \mathsf{Nat})$ if we replaced $x' = x + 1$ by $x = x' - 1$ in the definition of Φ.

We can avoid this pitfall by writing actions as conjunctions and disjunctions of formulas of the form $v' = e$ and $v' \in e$, where v is a variable and e a state function. The absence of types then produces no surprises. Moreover, actions written in this form have the advantage of being easier to understand, since they express the new values of variables directly in terms of their old values.

We could define a typed version of TLA. Semantically, we just restrict the collection of states to ones that assign to each variable a value of the proper type. However, types add a great deal of complexity to a logic. For example, what is the type of the division operator? If it is $\mathit{Real} \times \mathit{NonzeroReal} \rightarrow \mathit{Real}$, then type correctness becomes undecidable, so we lose automatic type checking, arguably the major benefit of types. If the type is $\mathit{Real} \times \mathit{Real} \rightarrow \mathit{Real}$, then what is the value of $1/0$? If it is a real number, then we will be able to prove the correctness of algorithms that we would usually consider to be incorrect, such as an iterative algorithm that takes $1/0$ as an initial approximation. Letting $1/0$ be a special "undefined" value $\perp$ leads to a complicated logic with truth values true, false, and $\perp$.

The type of an operator like division is just one of many problems introduced by types. These problems are easily hidden in informal presentation such as ours and the ones in most articles and books—for example, [Manna and Pnueli 1991] and [Chandy and Misra 1988]. The problems cannot be avoided in a formal treatment, such as is necessary for true mechanical verification. (They can be hidden when mechanically checking hand-translated verification conditions). Although one can formalize a typed version of TLA, the result is not nearly so simple as the untyped version. Types may be good for programming languages, but we believe that the difficulties they add far outweigh their advantages in a logic for reasoning about algorithms (Note 14).

6.2 Eventuality Properties

The second class of properties we consider are eventuality properties—ones asserting that something eventually happens. Here are some traditional eventuality properties and their expressions in temporal logic:

termination. The program eventually terminates: $\Diamond$ *terminated*.

service. If a process has requested service, then it eventually is served: *requested* $\leadsto$ *served*.

message delivery. If a message is sent often enough, then it is eventually delivered: $(\Box\Diamond\, sent) \Rightarrow \Diamond\, delivered$.

Although eventuality properties are expressed by a variety of temporal formulas, their proofs can always be reduced to the proof of leads-to properties—formulas of the form $P \leadsto Q$. For example, suppose we want to prove that Program 1 increases the value of x without bound. The TLA formula to be proved is

$$\Phi \wedge (\mathsf{n} \in \mathsf{Nat}) \;\Rightarrow\; \Diamond(x > \mathsf{n}) \tag{31}$$

The Lattice Rule of Figure 5 together with some simple temporal reasoning shows that (31) follows from

$$\Phi \;\Rightarrow\; ((\mathsf{n} \in \mathsf{Nat} \,\wedge\, x = \mathsf{n}) \leadsto (x = \mathsf{n}+1)) \tag{32}$$

To illustrate the use of TLA in proving leads-to properties, we now sketch the proof of (32).

Since safety properties don't imply that anything ever happens, leads-to properties must be derived from the program's fairness condition. Examining Figure 5 leads us to try rule WF1, with the following substitutions:

$$\begin{array}{lll} P \leftarrow \mathsf{n} \in \mathsf{Nat} \,\wedge\, x = \mathsf{n} & \mathcal{N} \leftarrow \mathcal{M} & f \leftarrow \langle x, y\rangle \\ Q \leftarrow x = \mathsf{n}+1 & \mathcal{A} \leftarrow \mathcal{M}_1 & \end{array}$$

The rule's hypotheses become

$$\begin{array}{l} (\mathsf{n} \in \mathsf{Nat} \,\wedge\, x = \mathsf{n}) \wedge [\mathcal{M}]_{\langle x, y\rangle} \;\Rightarrow\; ((\mathsf{n} \in \mathsf{Nat} \,\wedge\, x' = \mathsf{n}) \vee (x' = \mathsf{n}+1)) \\ (\mathsf{n} \in \mathsf{Nat} \,\wedge\, x = \mathsf{n}) \wedge \langle \mathcal{M}_1\rangle_{\langle x, y\rangle} \;\Rightarrow\; (x' = \mathsf{n}+1) \\ (\mathsf{n} \in \mathsf{Nat} \,\wedge\, x = \mathsf{n}) \;\Rightarrow\; \mathit{Enabled}\ \langle \mathcal{M}_1\rangle_{\langle x, y\rangle} \end{array}$$

which follow easily from the definitions of $\mathcal{M}_1$ and $\mathcal{M}$ in Figure 3. The rule's conclusion becomes

$$\Box[\mathcal{M}]_{\langle x, y\rangle} \wedge \mathrm{WF}_{\langle x, y\rangle}(\mathcal{M}_1) \;\Rightarrow\; ((\mathsf{n} \in \mathsf{Nat} \,\wedge\, x = \mathsf{n}) \leadsto (x = \mathsf{n}+1))$$

which, by definition of Φ, implies (32).

6.3 Other Properties

We have seen how invariance properties and eventuality properties are expressed as TLA formulas and proved. But, what about more complicated properties? How would one state the following property as a TLA formula?

> A behavior begins with x and y both zero, and repeatedly increments either x or y (in a single operation), choosing nondeterministically between them, but choosing each infinitely many times.

The answer, of course, it that we already have expressed this property in TLA. It is formula Φ of Figure 3.

In TLA, there is no distinction between a program and a property. Instead of viewing Φ as a description of a program, we can just as well consider it to be a

$$
\begin{array}{l}
\textbf{var integer } x, y \quad = 0 \; ; \\
\quad\quad \textbf{semaphore } \mathit{sem} \; = 1 \; ; \\
\textbf{cobegin loop } \alpha_1\!: \langle\, P(\mathit{sem}) \,\rangle \; ; \\
\quad\quad\quad\quad\quad\quad \beta_1\!: \langle\, x := x + 1 \,\rangle \; ; \\
\quad\quad\quad\quad\quad\quad \gamma_1\!: \langle\, V(\mathit{sem}) \,\rangle \quad \textbf{endloop} \\
\quad \Box \\
\quad\quad\quad\quad \textbf{loop } \alpha_2\!: \langle\, P(\mathit{sem}) \,\rangle \; ; \\
\quad\quad\quad\quad\quad\quad \beta_2\!: \langle\, y := y + 1 \,\rangle \; ; \\
\quad\quad\quad\quad\quad\quad \gamma_2\!: \langle\, V(\mathit{sem}) \,\rangle \quad \textbf{endloop} \\
\textbf{coend}
\end{array}
$$

Fig. 6. Program 2—our second example program.

property that we want a program to satisfy. The formula Φ, like the program of Figure 1 that it represents, is so simple that we can regard it as a specification of how we want a program to behave. As our next example, we consider a program that implements property Φ. That is, we give a program represented by a TLA formula Ψ that implies Φ.

7. ANOTHER EXAMPLE

7.1 Program 2

Our next example is Program 2 of Figure 6, written in a language invented for this program. (Since its only purpose is to help us write the TLA formula, the programming-language description of the program can be written with any convenient notation.) The program consists of two processes, each repeatedly executing a loop that contains three atomic operations. The variable *sem* is an integer semaphore, and P and V are the standard semaphore operations [Dijkstra 1968]. Since Figure 6 is an informal description, it doesn't matter whether or not you understand it. The real definition of Program 2 is the TLA formula Ψ defined below.

Describing the execution of Program 2 as a sequence of states requires each state to specify not only the values of the variables x, y, and *sem*, but also the control state of each process. Control in process 1 can be at one of the three "control points" α_1, β_1, or γ_1. We introduce the variable pc_1 that will assume the values "a", "b", and "g", denoting that control is at α_1, β_1, and γ_1, respectively. A similar variable pc_2 denotes the control state of process 2.

The definition of the TLA formula Ψ that represents Program 2 is given in Figure 7 .[4] A vertically aligned list of formulas preceded by "$\wedge$"s or "$\vee$"s denotes the conjunction or disjunction of those formulas, and we use indentation to eliminate parentheses. (These notational conventions make large formulas much easier to read.) Thus, Figure 7 defines the predicate $Init_\Psi$ to be the conjunction of three formulas, the second of which is $(x = 0) \wedge (y = 0)$.

As we explained in Section 5.5, a program is represented by a formula $Init \wedge \Box[\mathcal{N}]_f \wedge F$. In this example, $Init$ and f are fairly obvious: $Init$ is the predicate $Init_\Psi$ that specifies the initial values of the variables, and f is the 5-tuple w consisting of all the program's variables. The next-state relation $\mathcal{N}$ and the fairness requirement F are less obvious and merit some discussion.

[4]Section 9.2 discusses why Figure 7 is longer and seems more complex than Figure 6.

$$
\begin{array}{lcl}
Init_{\Psi} & \triangleq & \wedge\ (pc_1 = \text{“a”}) \wedge (pc_2 = \text{“a”}) \\
 & & \wedge\ (x = 0) \wedge (y = 0) \\
 & & \wedge\ sem = 1
\end{array}
$$

$$
\begin{array}{lcl@{\qquad}lcl}
\alpha_1 & \triangleq & \wedge\ (pc_1 = \text{“a”}) \wedge (0 < sem) & \alpha_2 & \triangleq & \wedge\ (pc_2 = \text{“a”}) \wedge (0 < sem) \\
 & & \wedge\ pc_1' = \text{“b”} & & & \wedge\ pc_2' = \text{“b”} \\
 & & \wedge\ sem' = sem - 1 & & & \wedge\ sem' = sem - 1 \\
 & & \wedge\ Unchanged\ \langle x,\ y,\ pc_2 \rangle & & & \wedge\ Unchanged\ \langle x,\ y,\ pc_1 \rangle \\
\beta_1 & \triangleq & \wedge\ pc_1 = \text{“b”} & \beta_2 & \triangleq & \wedge\ pc_2 = \text{“b”} \\
 & & \wedge\ pc_1' = \text{“g”} & & & \wedge\ pc_2' = \text{“g”} \\
 & & \wedge\ x' = x + 1 & & & \wedge\ y' = y + 1 \\
 & & \wedge\ Unchanged\ \langle y,\ sem,\ pc_2 \rangle & & & \wedge\ Unchanged\ \langle x,\ sem,\ pc_1 \rangle \\
\gamma_1 & \triangleq & \wedge\ pc_1 = \text{“g”} & \gamma_2 & \triangleq & \wedge\ pc_2' = \text{“a”} \\
 & & \wedge\ pc_1' = \text{“a”} & & & \wedge\ pc_2 = \text{“g”} \\
 & & \wedge\ sem' = sem + 1 & & & \wedge\ sem' = sem + 1 \\
 & & \wedge\ Unchanged\ \langle x,\ y,\ pc_2 \rangle & & & \wedge\ Unchanged\ \langle x,\ y,\ pc_1 \rangle \\
\mathcal{N}_1 & \triangleq & \alpha_1 \vee \beta_1 \vee \gamma_1 & \mathcal{N}_2 & \triangleq & \alpha_2 \vee \beta_2 \vee \gamma_2
\end{array}
$$

$$
\begin{array}{lcl}
\mathcal{N} & \triangleq & \mathcal{N}_1 \vee \mathcal{N}_2 \\
w & \triangleq & \langle x, y, sem, pc_1, pc_2 \rangle \\
\Psi & \triangleq & Init_{\Psi} \wedge \Box[\mathcal{N}]_w \wedge \mathrm{SF}_w(\mathcal{N}_1) \wedge \mathrm{SF}_w(\mathcal{N}_2)
\end{array}
$$

Fig. 7. The formula Ψ describing Program 2.

7.1.1 *The Next-State Relation.* Corresponding to the six atomic operations in Figure 6 are the six actions $\alpha_1, \ldots, \gamma_2$ defined in Figure 7. The four conjuncts in the definition of α_1 assert that an α_1 step:

(1) Starts in a state with $pc_1 = \text{“a”}$ (control in the first process is at control point α_1) and $0 < sem$ (the semaphore is positive).
(2) Ends in a state with $pc_1 = \text{“b”}$ (control in the first process is at control point β_1).
(3) Decrements *sem*.
(4) Does not change the values of x, y, and pc_2

Thus, an α_1 step represents an execution of statement α_1 of Figure 6. Similarly, the other actions represent the other operations of the program in Figure 6.

An $\mathcal{N}_1$ step is either an α_1 step, a β_1 step, or a γ_1 step, so it represents an execution of an atomic operation by the first process. Similarly, an $\mathcal{N}_2$ step represents an execution of an atomic operation by the second process. An $\mathcal{N}$ step represents a step of either process, so every program step is an $\mathcal{N}$ step—in other words, $\mathcal{N}$ is the program's next-state relation. Thus, $\Box[\mathcal{N}]_w$ is true for a behavior iff every step of the behavior is either a program step or else leaves the variables x, y, sem, pc_1, and pc_2 unchanged.

7.1.2 *The Fairness Requirement.* We want program Ψ to implement program Φ. Hence, Ψ must guarantee that both x and y are incremented infinitely often. To guarantee that x is incremented infinitely often, we need some fairness requirement to ensure that infinitely many $\mathcal{N}_1$ steps occur. This requirement must rule out the following behavior, in which process 1 is never executed.

$$
\begin{array}{l}
\langle\!\langle x \triangleq 0,\ y \triangleq 0,\ sem \triangleq 1,\ pc_1 \triangleq \text{“a”},\ pc_2 \triangleq \text{“a”},\ \ldots \rangle\!\rangle \\
\langle\!\langle x \triangleq 0,\ y \triangleq 0,\ sem \triangleq 0,\ pc_1 \triangleq \text{“a”},\ pc_2 \triangleq \text{“b”},\ \ldots \rangle\!\rangle
\end{array}
$$

$$
\begin{array}{l}
\llparenthesis x \triangleq 0,\ y \triangleq 1,\ sem \triangleq 0,\ pc_1 \triangleq \text{“a”},\ pc_2 \triangleq \text{“g”},\ \ldots \rrparenthesis \\
\llparenthesis x \triangleq 0,\ y \triangleq 1,\ sem \triangleq 1,\ pc_1 \triangleq \text{“a”},\ pc_2 \triangleq \text{“a”},\ \ldots \rrparenthesis \\
\llparenthesis x \triangleq 0,\ y \triangleq 1,\ sem \triangleq 0,\ pc_1 \triangleq \text{“a”},\ pc_2 \triangleq \text{“b”},\ \ldots \rrparenthesis \\
\llparenthesis x \triangleq 0,\ y \triangleq 2,\ sem \triangleq 0,\ pc_1 \triangleq \text{“a”},\ pc_2 \triangleq \text{“g”},\ \ldots \rrparenthesis \\
\llparenthesis x \triangleq 0,\ y \triangleq 2,\ sem \triangleq 1,\ pc_1 \triangleq \text{“a”},\ pc_2 \triangleq \text{“a”},\ \ldots \rrparenthesis \\
\quad \vdots
\end{array}
$$

Observe that an α_1 step is possible iff pc_1 equals “a” and *sem* is positive, so *Enabled* α_1 equals $(pc_1 = \text{“a”}) \wedge (0 < sem)$. In this behavior, *Enabled* α_1 is true whenever pc_2 equals “a”, and false otherwise—both situations occurring infinitely often. An α_1 step is also an $\mathcal{N}_1$ step. Moreover, every α_1 step changes pc_1 and *sem*, so it changes w. Hence, any α_1 step is an $\langle \mathcal{N}_1 \rangle_w$ step, so $\langle \mathcal{N}_1 \rangle_w$ is enabled and disabled infinitely often in this behavior.

The weak fairness condition $\mathrm{WF}_w(\mathcal{N}_1)$ asserts that $\langle \mathcal{N}_1 \rangle_w$ is disabled infinitely often or infinitely many $\langle \mathcal{N}_1 \rangle_w$ steps occur. Since $\langle \mathcal{N}_1 \rangle_w$ is disabled infinitely often, $\mathrm{WF}_w(\mathcal{N}_1)$ does not rule out this behavior.

The strong fairness condition $\mathrm{SF}_w(\mathcal{N}_1)$ asserts that either $\langle \mathcal{N}_1 \rangle_w$ is eventually forever disabled or else infinitely many $\langle \mathcal{N}_1 \rangle_w$ steps occur. Neither assertion is true for this behavior, so the behavior does not satisfy $\mathrm{SF}_w(\mathcal{N}_1)$. This example indicates why we need the fairness condition $\mathrm{SF}_w(\mathcal{N}_1)$ to guarantee that x is incremented infinitely often.

There are other ways of writing this fairness condition. An equivalent definition of Ψ is obtained by replacing $\mathrm{SF}_w(\mathcal{N}_1)$ with $\mathrm{SF}_w(\alpha_1) \wedge \mathrm{SF}_w(\beta_1) \wedge \mathrm{SF}_w(\gamma_1)$ or with $\mathrm{SF}_w(\alpha_1) \wedge \mathrm{WF}_w(\beta_1) \wedge \mathrm{WF}_w(\gamma_1)$. Equivalence of these definitions follows from the formulas

$$Init_\Psi \wedge \Box[\mathcal{N}]_w \;\Rightarrow\; (\mathrm{SF}_w(\mathcal{N}_1) \;\equiv\; \mathrm{SF}_w(\alpha_1) \wedge \mathrm{SF}_w(\beta_1) \wedge \mathrm{SF}_w(\gamma_1)) \tag{33}$$

$$Init_\Psi \wedge \Box[\mathcal{N}]_w \;\Rightarrow\; (\mathrm{SF}_w(\beta_1) \equiv \mathrm{WF}_w(\beta_1)) \tag{34}$$

$$Init_\Psi \wedge \Box[\mathcal{N}]_w \;\Rightarrow\; (\mathrm{SF}_w(\gamma_1) \equiv \mathrm{WF}_w(\gamma_1)) \tag{35}$$

Intuitively, (33) holds because once control reaches α_1, β_1, or γ_1, it remains there until the corresponding action is executed; (34) holds because once control reaches β_1, action β_1 is enabled until it is executed; and (35) is similar to (34).

Corresponding reasoning about y and $\mathcal{N}_2$ leads to the fairness condition $\mathrm{SF}_w(\mathcal{N}_2)$ for the second process.

7.2 Proving Program 2 Implements Program 1

To show that Program 2 implements Program 1, we must prove the TLA formula $\Psi \Rightarrow \Phi$, where Ψ is defined in Figure 7 and Φ is defined in Figure 3. By these definitions, $\Psi \Rightarrow \Phi$ follows from the following three formulas.

$$Init_\Psi \;\Rightarrow\; Init_\Phi \tag{36}$$

$$\Box[\mathcal{N}]_w \;\Rightarrow\; \Box[\mathcal{M}]_{\langle x,\, y \rangle} \tag{37}$$

$$\Psi \;\Rightarrow\; \mathrm{WF}_{\langle x,\, y \rangle}(\mathcal{M}_1) \wedge \mathrm{WF}_{\langle x,\, y \rangle}(\mathcal{M}_2) \tag{38}$$

Formula (36) asserts that the initial condition of Ψ implies the initial condition of Φ. It follows easily from the definitions of $Init_\Psi$ and $Init_\Phi$.

Roughly speaking, formula (37) asserts that every $\mathcal{N}$ step simulates an $\mathcal{M}$ step, and (38) asserts that Program 2 implements Program 1's fairness conditions. We now sketch the proofs of these two formulas.

7.2.1 *Proof of Step-Simulation.* Applying rule TLA2 of Figure 5 with true substituted for P and Q shows that (37) follows from

$$[\mathcal{N}]_w \;\Rightarrow\; [\mathcal{M}]_{\langle x,\, y\rangle} \tag{39}$$

By definition, $[\mathcal{N}]_w$ equals $\alpha_1 \vee \ldots \vee \gamma_2 \vee (w' = w)$ and $[\mathcal{M}]_{\langle x,\, y\rangle}$ equals $\mathcal{M}_1 \vee \mathcal{M}_2 \vee (\langle x,\, y\rangle' = \langle x,\, y\rangle)$. Formula (37) therefore follows from

$$\begin{array}{ll} \alpha_1 \;\Rightarrow\; \langle x,\, y\rangle' = \langle x,\, y\rangle & \alpha_2 \;\Rightarrow\; \langle x,\, y\rangle' = \langle x,\, y\rangle \\ \beta_1 \;\Rightarrow\; \mathcal{M}_1 & \beta_2 \;\Rightarrow\; \mathcal{M}_2 \\ \gamma_1 \;\Rightarrow\; \langle x,\, y\rangle' = \langle x,\, y\rangle & \gamma_2 \;\Rightarrow\; \langle x,\, y\rangle' = \langle x,\, y\rangle \\ \multicolumn{2}{c}{(w' = w) \;\Rightarrow\; \langle x,\, y\rangle' = \langle x,\, y\rangle} \end{array} \tag{40}$$

These implications are all trivial consequences of the definitions.

7.2.2 *Proof of Fairness.* For the fairness requirement (38), we sketch the proof that Ψ implies $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$. The proof that it implies $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_2)$ is similar.

Strong fairness of Program 2 is necessary to insure that x is incremented infinitely often, so Figure 5 suggests applying SF2 (without the overbars). At first glance, SF2 doesn't seem to work because its conclusion implies a strong fairness condition, and we want to prove $\Psi \Rightarrow \mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$. However, if x is a natural number, so $x' \neq x + 1$, then $\mathit{Enabled}\ \langle \mathcal{M}_1\rangle_{\langle x,\, y\rangle}$ equals true. A simple invariance argument proves $\Psi \Rightarrow \Box(x \in \mathsf{Nat})$, so $\Psi \Rightarrow \Box(\mathit{Enabled}\ \langle \mathcal{M}_1\rangle_{\langle x,\, y\rangle})$. Hence, Ψ implies that $\mathrm{SF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$ and $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$ are equivalent—both being equal to $\Box\Diamond\langle \mathcal{M}_1\rangle_{\langle x,\, y\rangle}$. We can thus expect to prove $\Psi \Rightarrow \mathrm{SF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$, which implies $\Psi \Rightarrow \mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M}_1)$.

Comparing the conclusion of rule SF2 with the formula we are trying to prove apparently leads to the following substitutions in the rule.

$$\mathcal{N} \leftarrow \mathcal{N} \qquad \mathcal{M} \leftarrow \mathcal{M}_1 \qquad f \leftarrow w \qquad g \leftarrow \langle x,\, y\rangle$$

However, it turns out that we need to strengthen $\mathcal{N}$ by the use of an invariant. We must find a predicate I (an invariant) that satisfies

$$\mathit{Init}_\Psi \wedge \Box[\mathcal{N}]_w \;\Rightarrow\; \Box I \tag{41}$$

By rule INV2, we can then rewrite Ψ as

$$\mathit{Init}_\Psi \;\wedge\; \Box[\mathcal{N} \wedge I \wedge I']_w \;\wedge\; \mathrm{SF}_w(\mathcal{N}_1) \;\wedge\; \mathrm{SF}_w(\mathcal{N}_2)$$

and substitute $\mathcal{N} \wedge I \wedge I'$ for $\mathcal{N}$. We will discover the invariant I in the course of the proof.

The first hypothesis of the rule and (40) suggest substituting β_1 for $\mathcal{B}$. The conclusion and the second hypothesis leads to the substitution of $\mathcal{N}_1$ for $\mathcal{A}$ and $\mathrm{SF}_w(\mathcal{N}_2)$ for $\Box F$, using the temporal tautology $\mathrm{SF}_w(\mathcal{N}_2) \equiv \Box\mathrm{SF}_w(\mathcal{N}_2)$. The second and fourth hypotheses lead to the substitution of $pc_1 =$ "b" for P. With these substitutions, the proof rule becomes

$$\frac{\begin{array}{l} \langle \mathcal{N} \wedge I \wedge I' \wedge \beta_1\rangle_w \;\Rightarrow\; \langle \mathcal{M}_1\rangle_{\langle x,\, y\rangle} \\ (pc_1 = \text{“b”}) \wedge (pc_1' = \text{“b”}) \wedge \langle \mathcal{N} \wedge I \wedge I' \wedge \mathcal{N}_1\rangle_w \;\Rightarrow\; \beta_1 \\ (pc_1 = \text{“b”}) \wedge \mathit{Enabled}\ \langle \mathcal{M}_1\rangle_{\langle x,\, y\rangle} \;\Rightarrow\; \mathit{Enabled}\ \langle \mathcal{N}_1\rangle_w \\ \Box[\mathcal{N} \wedge I \wedge I' \wedge \neg\beta_1]_w \wedge \mathrm{SF}_w(\mathcal{N}_1) \wedge \mathrm{SF}_w(\mathcal{N}_2) \wedge \Box\Diamond \mathit{Enabled}\ \langle \mathcal{M}_1\rangle_{\langle x,\, y\rangle} \\ \qquad \Rightarrow\; \Diamond\Box(pc_1 = \text{“b”}) \end{array}}{\Box[\mathcal{N} \wedge I \wedge I']_w \wedge \mathrm{SF}_w(N_1) \wedge \mathrm{SF}_w(\mathcal{N}_2) \;\Rightarrow\; \mathrm{SF}_{\langle x,\, y\rangle}(\mathcal{M}_1)}$$

The first three hypotheses are simple action formulas. The second and third follow easily from the definitions of $\mathcal{N}_1$, β_1 and $\mathcal{M}_1$. To prove the first hypothesis, we must show that $\mathcal{N} \wedge I \wedge I' \wedge \beta_1 \wedge (w' \neq w)$ implies $\mathcal{M}_1 \wedge (\langle x, y\rangle' \neq \langle x, y\rangle)$. As we observed in (40), β_1 implies $\mathcal{M}_1$. Since β_1 also implies $x' = x+1$, which implies $x' \neq x$ if x is a natural number, the first hypothesis holds if the invariant I implies $x \in \mathsf{Nat}$.

The fourth hypothesis is a temporal formula, which we now examine. To simplify the intuitive reasoning, let us ignore steps that don't change w. The fourth hypothesis then asserts that if every step is an $\mathcal{N} \wedge I \wedge I'$ step that is not a β_1 step, and the fairness conditions hold, then eventually control reaches β_1 and remains there forever. From the informal description of the program in Figure 6, this seems valid. No matter where control starts in process 1, fairness implies that eventually it must reach β_1, where it must remain forever if no β_1 step is performed.

Unfortunately, this intuitive reasoning is wrong. The fourth hypothesis is not a valid TLA formula. For example, consider a behavior that starts in a state with $pc_1 = pc_2 =$ "a" and $sem = 0$, and that remains in this state forever. In such a behavior, the left-side of the implication in the fourth hypothesis is true, but pc_1 never becomes equal to "b". Thus, the hypothesis is not satisfied by these behaviors.

The fourth hypothesis is invalid for behaviors starting in "bad" states—ones that are not reachable by executing the program from an initial state satisfying $Init_\Psi$. Such states have to be ruled out by the invariant I. We must substitute $\mathrm{SF}_w(\mathcal{N}_2) \wedge \Box I$ for $\Box F$ (using the tautology $\mathrm{SF}_w(\mathcal{N}_2) \wedge \Box I \equiv \Box(\mathrm{SF}_w(\mathcal{N}_2) \wedge I)$) and $(pc_1 = \text{"b"}) \wedge I$ for P in Rule SF2, obtaining the following as the fourth hypothesis.

$$
\begin{array}{l}
G \Rightarrow \Diamond\Box((pc_1 = \text{"b"}) \wedge I) \\
\text{where } G \triangleq \Box[\mathcal{N} \wedge I \wedge I' \wedge \neg\beta_1]_w \wedge \mathrm{SF}_w(\mathcal{N}_1) \wedge \mathrm{SF}_w(\mathcal{N}_2) \wedge \Box I \\
\qquad\qquad \wedge \Box\Diamond \mathit{Enabled}\ \langle \mathcal{M}_1\rangle_{\langle x,\, y\rangle}
\end{array}
\tag{42}
$$

Remembering that I must imply $x \in \mathsf{Nat}$, the reader with experience reasoning about concurrent programs will discover that the appropriate invariant is

$$
\begin{array}{lll}
I & \triangleq & \wedge\ x \in \mathsf{Nat} \\
 & & \wedge\ \vee\ (sem = 1) \wedge (pc_1 = pc_2 = \text{"a"}) \\
 & & \quad\ \vee\ (sem = 0) \wedge \vee\ (pc_1 = \text{"a"}) \wedge (pc_2 \in \{\text{"b"}, \text{"g"}\}) \\
 & & \qquad\qquad\qquad\quad\ \vee\ (pc_2 = \text{"a"}) \wedge (pc_1 \in \{\text{"b"}, \text{"g"}\})
\end{array}
$$

With this definition of I, the invariance property (41) follows easily from Rule INV1.

Having deduced that we need to prove (42), we must understand why it is true. A little thought reveals that (42) holds because control in process 1 must eventually reach β_1, and $\Box[\mathcal{N} \ldots \wedge \neg\beta_1]_w$, which asserts that a β_1 action is never executed, implies that control must then remain there forever. This reasoning is formalized by applying simple temporal reasoning based on the Lattice Rule to derive (42) from:

$$G \Rightarrow (\,(pc_1 = \text{"g"}) \wedge I \leadsto (pc_1 = \text{"a"}) \wedge I\,) \tag{43}$$

$$G \Rightarrow (\,(pc_1 = \text{"a"}) \wedge I \leadsto (pc_1 = \text{"b"}) \wedge I\,) \tag{44}$$

$$G \Rightarrow (\,(pc_1 = \text{"b"}) \wedge I \Rightarrow \Box((pc_1 = \text{"b"}) \wedge I)\,) \tag{45}$$

To see how to prove these formulas, we once again use simple pattern matching against the proof rules of Figure 5. We find that (43) and (44) should be proved by Rule SF1 with $\mathcal{N}_1$ substituted for $\mathcal{A}$ and $\mathrm{SF}_w(\mathcal{N}_2)$ substituted for $\Box F$, and that (45) should be proved by rule INV1 with $(pc_1 = \text{"b"}) \wedge I$ substituted for I. The proofs of (43) and (45) are simple. The proof of (44) is not so easy, the hard part being the proof of the third hypothesis:

$$H \;\Rightarrow\; \Diamond \mathit{Enabled}\ \langle \mathcal{N}_1 \rangle_w \qquad (46)$$
$$\text{where } H \;\triangleq\; \Box((pc_1 = \text{"a"}) \wedge I) \;\wedge\; \Box[\mathcal{N} \wedge I \wedge I' \wedge \neg\beta_1]_w \;\wedge\; \mathrm{SF}_w(\mathcal{N}_2)$$

Once again, we have reached a point where blind application of rules fails; we must understand why (46) is true. If pc_1 equals "a", then action $\mathcal{N}_1$ is enabled when control in process 2 is at α_2, and strong fairness for $\mathcal{N}_2$ implies that control must eventually reach α_2. This intuitive reasoning leads us to deduce (46) by temporal reasoning from

$$(pc_1 = \text{"a"}) \wedge I \;\;\Rightarrow\;\; (\mathit{Enabled}\ \langle \mathcal{N}_1 \rangle_w \equiv (pc_2 = \text{"a"}))$$
$$H \;\;\Rightarrow\;\; ((pc_2 = \text{"b"}) \leadsto (pc_2 = \text{"g"}))$$
$$H \;\;\Rightarrow\;\; ((pc_2 = \text{"g"}) \leadsto (pc_2 = \text{"a"}))$$

The first formula follows from the observation that

$$\mathit{Enabled}\ \langle \mathcal{N}_1 \rangle_w \;\equiv\; ((pc_1 = \text{"a"}) \wedge (0 < \mathit{sem})) \vee (pc_1 = \text{"b"}) \vee (pc_1 = \text{"g"})$$

Pattern matching against the proof rules leads to simple proofs of the remaining two formulas by substituting $\mathcal{N}_2$ for $\mathcal{A}$ and true for F in SF1.

7.3 Comments on the Proof

This example illustrates the general method of proving that a lower-level program Ψ implements a higher-level program Φ. There are three things to prove: (i) the initial predicate of Ψ implies the initial predicate of Φ, (ii) a step of Ψ simulates a step of Φ, and (iii) Ψ implies the fairness requirement of Φ.

As in the example, proving the initial condition is generally straightforward. Of course, in more realistic examples there will be more details to check.

Because our example was so simple, the proof of step-simulation was atypical. Usually, a step of the lower-level program starting in a completely arbitrary state does not simulate a step of the higher-level program. We must first find the proper invariant, and then apply Rule INV2 to prove step-simulation. Once the invariant is found, the proof is a straightforward exercise in showing that one action implies another. The structure of the formulas tells us how to decompose a large proof into a number of smaller ones.

Our proof of fairness was quite typical in its alternation of blind application of proof rules with the need to understand why a property holds. As in this proof, an invariant is almost always required. Usually, it is the same invariant as in the proof of step-simulation. Of course, the proofs of real algorithms will be more complicated.

Our proof may already have seemed rather complicated for such a simple example, but the example is a bit more subtle than it appears. The reader who attempts a rigorous informal proof will discover that each step in the TLA proof mirrors a step in the informal proof. The more rigorous the informal proof, the more it will resemble the TLA proof. Rules SF1 and SF2 conveniently encapsulate reasoning that occurs over and over again in informal proofs. We believe that temporal logic provides an ideal formalism for translating intuitive understanding of why a liveness property holds into a formal proof.

Were we to choose the weaker fairness requirement $\mathrm{WF}_{\langle x,\, y\rangle}(\mathcal{M})$ for Program 1, then Program 2's fairness requirement could be weakened to $\mathrm{WF}_w(\mathcal{N})$. The proof of $\Psi \Rightarrow \Phi$, using WF2 instead of SF2, would then be simpler. Writing out this proof is a good exercise in applying TLA.

7.4 Stuttering and Refinement

Program 2 is *finer grained* than Program 1, in the sense that the three atomic operations of each process's loop in Program 2 correspond to a single atomic operation of Program 1. Besides the steps that increment x or y, Program 2 takes steps that modify *sem* and pc_1 or pc_2, but leave x and y unchanged. Program 2 implements Program 1—that is, the formula $\Psi \Rightarrow \Phi$ is valid—only because Φ allows stuttering steps that do not change x and y. Program 2 can in turn be implemented by a still finer-grained program because Ψ allows steps that do not change any of its variables. Allowing stuttering steps is the key to refining the grain of atomicity.

8. HIDING VARIABLES

8.1 A Memory Specification

We now consider another example: a simple processor/memory interface. The processor issues *read* and *write* operations that are executed by the memory. The interface consists of three registers, represented by the following three variables.

- *op* Set by the processor to indicate the desired operation, and reset by the memory after executing the operation.
- *adr* Set by the processor to indicate the address of the memory location to be read or written.
- *val* Set by the processor to indicate the value to be written by a *write*, and set by the memory to return the result of a *read*.

Here is a typical behavior, where "—" indicates that the value is irrelevant, and memory location 432 happens to have the initial value 777.

$$
\begin{array}{l}
\langle\!\langle\, op \triangleq \text{“ready”},\ adr \triangleq \text{—},\ val \triangleq \text{—},\ \ldots \,\rangle\!\rangle \\
\langle\!\langle\, op \triangleq \text{“read”},\ adr \triangleq 432,\ val \triangleq \text{—},\ \ldots \,\rangle\!\rangle \\
\langle\!\langle\, op \triangleq \text{“ready”},\ adr \triangleq \text{—},\ val \triangleq 777,\ \ldots \,\rangle\!\rangle \\
\langle\!\langle\, op \triangleq \text{“write”},\ adr \triangleq 196,\ val \triangleq 0,\ \ldots \,\rangle\!\rangle \\
\langle\!\langle\, op \triangleq \text{“ready”},\ adr \triangleq \text{—},\ val \triangleq \text{—},\ \ldots \,\rangle\!\rangle \\
\quad \vdots
\end{array}
$$

It is easy to specify this interface if we introduce an additional variable *memory* to denote the contents of memory, so $memory(n)$ is the current value of memory location n. The property Φ describing the behaviors is shown in Figure 8, where Address is the set of legal addresses, and MemVal is the set of possible memory values. Action $\mathcal{S}(m, v)$ represents the assignment $memory(m) := v$ (Note 15). Actions $\mathcal{R}_{proc}$ and $\mathcal{W}_{proc}$ represent the processor's *read*- and *write*-request operations; actions $\mathcal{R}_{mem}$ and $\mathcal{W}_{mem}$ represent the memory's responses to those requests.

$$
\begin{aligned}
Init_{\Phi} &\triangleq \land\ op = \text{“ready”} \\
&\ \land\ \forall \mathsf{n} \in \mathsf{Address} : memory(\mathsf{n}) \in \mathsf{MemVal} \\
\mathcal{S}(m, v) &\triangleq \forall \mathsf{n} \in \mathsf{Address} : \land\ (\mathsf{n} = m) \Rightarrow (memory(\mathsf{n})' = v) \\
&\phantom{\triangleq \forall \mathsf{n} \in \mathsf{Address} :}\ \land\ (\mathsf{n} \neq m) \Rightarrow (memory(\mathsf{n})' = memory(\mathsf{n}))
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{R}_{proc} &\triangleq \land\ op = \text{“ready”} & \mathcal{R}_{mem} &\triangleq \land\ op = \text{“read”} \\
&\ \land\ op' = \text{“read”} & &\ \land\ op' = \text{“ready”} \\
&\ \land\ adr' \in \mathsf{Address} & &\ \land\ val' = memory(adr) \\
&\ \land\ memory' = memory & &\ \land\ memory' = memory \\
\mathcal{W}_{proc} &\triangleq \land\ op = \text{“ready”} & \mathcal{W}_{mem} &\triangleq \land\ op = \text{“write”} \\
&\ \land\ op' = \text{“write”} & &\ \land\ op' = \text{“ready”} \\
&\ \land\ adr' \in \mathsf{Address} & &\ \land\ \mathcal{S}(adr, val) \\
&\ \land\ val' \in \mathsf{MemVal} & & \\
&\ \land\ memory' = memory & &
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{N}_{mem} &\triangleq \mathcal{R}_{mem} \lor \mathcal{W}_{mem} \\
\mathcal{N} &\triangleq \mathcal{N}_{mem} \lor \mathcal{R}_{proc} \lor \mathcal{W}_{proc} \\
w &\triangleq \langle op, adr, val, memory \rangle \\
\Phi &\triangleq Init_{\Phi} \land \Box[\mathcal{N}]_w \land \mathrm{WF}_w(\mathcal{N}_{mem})
\end{aligned}
$$

Fig. 8. “Internal” specification of a processor/memory interface.

Action $\mathcal{N}_{mem}$ denotes the memory's next-state relation. The fairness condition $\mathrm{WF}_w(\mathcal{N}_{mem})$ asserts that the memory eventually responds to each request; there is no requirement that the processor ever issues requests.

Observe that the action $\mathcal{S}(m, v)$ is used only to define $\mathcal{W}_{mem}$; it was introduced just to keep the definition of $\mathcal{W}_{mem}$ from running off the page. There is no formal significance to our choice of names such as $\mathcal{R}_{proc}$. Our decision to define $\mathcal{N}_{mem}$ as the disjunction of two simpler actions was completely arbitrary; we could just as well have defined it all at once, or as the disjunction of more than two actions. There are countless ways of writing logically equivalent formulas Φ.

The formula Φ specifies the right behavior for the interface variables *op*, *adr*, and *val*. However, it also specifies the value of the variable *memory*, which we did not want to specify. We want to specify only how the three interface variables change; we do not care how any other variables such as x, *sem*, or *memory* change. We therefore want a formula asserting that *op*, *adr*, and *val* behave as described by Φ, but that it doesn't matter what values *memory* assumes. Such a formula is sometimes described as Φ with the variable *memory* “hidden”. This formula is written $\boldsymbol{\exists}\ memory : \Phi$.

The precise meaning of the formula $\boldsymbol{\exists}\ memory : \Phi$ is defined below. Here, we simply want to observe that the free (flexible) variables of this formula are *op*, *adr*, and *val*. Since x, *sem*, and *memory* do not occur free, the formula does not constrain them in any way.

8.2 Quantification over Flexible Variables

We now define $\boldsymbol{\exists}\ x : F$, where x is a (flexible) variable and F a temporal formula. Intuitively, $\boldsymbol{\exists}\ x : F$ asserts that it doesn't matter what the actual values of x are, but that there are some values x can assume for which F holds. For example, $\boldsymbol{\exists}\ x : \Box[y = x']_{\langle x,\, y \rangle}$ is satisfied by the behavior

$$\langle\!\langle x \triangleq \text{“a”},\ \ y \triangleq 0,\ \ z \triangleq \text{“uvw”}, \ldots \rangle\!\rangle$$

$$\begin{array}{l}
(\!|x \triangleq \text{“b”},\ \ y \triangleq 1,\ \ z \triangleq -13,\ \ \ldots|\!) \\
(\!|x \triangleq \text{“c”},\ \ y \triangleq 1,\ \ z \triangleq -13,\ \ \ldots|\!) \\
(\!|x \triangleq 77,\ \ \ \ y \triangleq 2,\ \ z \triangleq \text{“vw”},\ldots|\!) \\
\quad\vdots
\end{array}$$

because by changing only the values of x, we get the following behavior that satisfies $\Box[y = x']_{\langle x,\, y\rangle}$.

$$\begin{array}{l}
(\!|x \triangleq \text{“a”},\ \ y \triangleq 0,\ z \triangleq \text{“uvw”},..|\!) \\
(\!|x \triangleq 0,\ \ \ \ \ y \triangleq 1,\ z \triangleq -13,\ \ \ldots|\!) \\
(\!|x \triangleq 1,\ \ \ \ \ y \triangleq 1,\ z \triangleq -13,\ \ \ldots|\!) \\
(\!|x \triangleq 1,\ \ \ \ \ y \triangleq 2,\ z \triangleq \text{“vw”},\ \ldots|\!) \\
\quad\vdots
\end{array}$$

In fact, every behavior satisfies $\exists\, x : \Box[y = x']_{\langle x,\, y\rangle}$.

To define $\exists\, x : F$ formally, we need some auxiliary definitions. For any variable x and states s and t, let $s =_x t$ mean that s and t assign the same values to all variables other than x. More precisely,

$$s =_x t \ \ \triangleq\ \ \forall\, `v\text{'} \neq `x\text{'} : s[\![v]\!] = t[\![v]\!]$$

We extend the relation $=_x$ to behaviors in the obvious way:

$$\langle s_0, s_1, \ldots \rangle =_x \langle t_0, t_1, \ldots \rangle \ \ \triangleq\ \ \forall\, \mathsf{n} \in \mathsf{Nat} : s_{\mathsf{n}} =_x t_{\mathsf{n}}$$

The obvious next step is to define

$$\sigma[\![\exists\, x : F]\!] \ \ \triangleq\ \ \exists\, \tau \in \mathbf{St}^{\infty} : (\sigma =_x \tau) \wedge \tau[\![F]\!] \tag{47}$$

for any behavior σ. (Recall that $\mathbf{St}^{\infty}$ is the collection of all behaviors.) However, this definition is not quite right, because the formula it defines is not necessarily invariant under stuttering. For example, suppose F is satisfied only by behaviors in which x changes before y does, including the behavior

$$\begin{array}{l}
(\!|x \triangleq 1,\ y \triangleq \text{“a”},\ z \triangleq 7,\ \ldots|\!) \\
(\!|x \triangleq 2,\ y \triangleq \text{“a”},\ z \triangleq 7,\ \ldots|\!) \\
(\!|x \triangleq 2,\ y \triangleq \text{“b”},\ z \triangleq 14,\ \ldots|\!) \\
\quad\vdots
\end{array}$$

Then definition (47) implies that the behavior

$$\begin{array}{l}
(\!|x \triangleq 999,\ y \triangleq \text{“a”},\ z \triangleq 7,\ \ldots|\!) \\
(\!|x \triangleq 999,\ y \triangleq \text{“a”},\ z \triangleq 7,\ \ldots|\!) \\
(\!|x \triangleq 999,\ y \triangleq \text{“b”},\ z \triangleq 14,\ \ldots|\!) \\
\quad\vdots
\end{array}$$

satisfies $\exists\, x : F$ (because we can produce a behavior satisfying F by changing only the values of x). However, the behavior

$$\begin{array}{l}
(\!|x \triangleq 999,\ y \triangleq \text{“a”},\ z \triangleq 7,\ \ldots|\!) \\
(\!|x \triangleq 999,\ y \triangleq \text{“b”},\ z \triangleq 14,\ \ldots|\!) \\
\quad\vdots
\end{array}$$

does not satisfy $\exists\!\!\!\exists\, x : F$ (because of the assumption that F requires x to change before y does). With appropriate values for all other variables, these two behaviors differ only by stuttering steps. Hence, with definition (47), $\exists\!\!\!\exists\, x : F$ is not necessarily invariant under stuttering even though F is.

To obtain invariance under stuttering, we must define $\exists\!\!\!\exists\, x : F$ to be satisfied by a behavior σ iff we can obtain a behavior that satisfies F by first adding stuttering and then changing the values of x. We define $\natural\sigma$ to be the behavior obtained from the behavior σ by removing all stuttering steps—except that if σ ends with infinite stuttering, then those final stuttering steps are kept. The precise definition is:

$$\natural\langle s_0, s_1, s_2, \ldots\rangle \;\stackrel{\Delta}{=}\; \begin{array}[t]{l}\textbf{if } \forall \mathsf{n} \in \mathsf{Nat} : s_{\mathsf{n}} = s_0 \\ \quad \textbf{then } \langle s_0, s_0, s_0, \ldots\rangle \\ \quad \textbf{else } \begin{array}[t]{l} \textbf{if } s_1 = s_0 \textbf{ then } \natural\langle s_1, s_2, s_3, \ldots\rangle \\ \phantom{\textbf{if } s_1 = s_0} \textbf{ else } \langle s_0\rangle \circ \natural\langle s_1, s_2, \ldots\rangle \end{array}\end{array} \tag{48}$$

where $\circ$ denotes concatenation of sequences. We then define $\exists\!\!\!\exists$ by

$$\sigma[\![\exists\!\!\!\exists\, x : F]\!] \;\stackrel{\Delta}{=}\; \exists\, \rho, \tau \in \mathbf{St}^{\infty} : (\natural\sigma = \natural\rho) \wedge (\rho =_x \tau) \wedge \tau[\![F]\!] \tag{49}$$

The operator $\exists\!\!\!\exists\, x$ differs from ordinary existential quantification because it asserts the existence not of a single value to be substituted for x, but of an infinite sequence of values. However, it really is existential quantification because it obeys the ordinary laws of existential quantification. In particular, the usual rules E1 and E2 of Figure 9 are sound. From these rules, one can deduce the expected properties of existential quantification, such as

$$(\exists\!\!\!\exists\, x : F \vee G) \;\equiv\; (\exists\!\!\!\exists\, x : F) \vee (\exists\!\!\!\exists\, x : G)$$

We can extend TLA to allow quantification over rigid as well as flexible variables. Since the value of a rigid variable is constant throughout a behavior, quantification over rigid variables is much simpler than quantification over flexible variables (Note 16). However, it is of less use. The semantics of quantification over rigid variables is defined in Figure 9.

General TLA formulas consist of all formulas obtained from simple TLA formulas by logical operators and quantification over program and rigid variables. The syntax and semantics of quantification are summarized in Figure 9, which together with Figure 4 gives the complete definition of TLA. It is easy to check that TLA formulas are invariant under stuttering, which means formally that $\natural\sigma = \natural\tau$ implies $\sigma[\![F]\!] = \tau[\![F]\!]$ for all TLA formulas F and behaviors σ and τ.

8.3 Refinement Mappings

8.3.1 *Implementing The Memory Specification.* We now give a simple implementation of the processor/memory interface specified by the formula $\exists\!\!\!\exists\, \mathit{memory} : \Phi$, where Φ is defined in Figure 8. The implementation uses a main memory and a cache, represented by variables *main* and *cache*. The value of $\mathit{cache}(m)$ represents the cache's value for memory location m, the special value $\perp$ (assumed not to be in MemVal) denoting that this memory location is not in the cache. The processor's read and write requests are serviced from the cache, and separate internal actions (not visible from the interface) move values between the cache and main memory. When the processor reads a value not in the cache, the value is first moved into the cache and then put in *val*.

Syntax

$$
\begin{array}{lcl}
\langle \textit{general formula} \rangle & \triangleq & \langle \textit{formula} \rangle \mid \boldsymbol{\exists}\, \langle \textit{variable} \rangle : \langle \textit{general formula} \rangle \\
& & \mid \exists\, \langle \textit{rigid variable} \rangle : \langle \textit{general formula} \rangle \\
& & \mid \langle \textit{general formula} \rangle \wedge \langle \textit{general formula} \rangle \\
& & \mid \neg \langle \textit{general formula} \rangle \\
\langle \textit{formula} \rangle & \triangleq & \text{a simple TLA formula (see Figure 4)}
\end{array}
$$

Semantics

$$\langle s_0, s_1, \ldots \rangle =_x \langle t_0, t_1, \ldots \rangle \;\triangleq\; \forall\, \mathsf{n} \in \mathsf{Nat} : \forall\, `v\text{'} \neq `x\text{'} : s_{\mathsf{n}}[\![v]\!] = t_{\mathsf{n}}[\![v]\!]$$

$$
\begin{array}{lll}
\natural\langle s_0, s_1, s_2, \ldots \rangle \;\triangleq & \textbf{if} & \forall \mathsf{n} \in \mathsf{Nat} : s_{\mathsf{n}} = s_0 \\
& \textbf{then} & \langle s_0, s_0, s_0, \ldots \rangle \\
& \textbf{else} & \textbf{if}\ \ s_1 = s_0\ \ \textbf{then}\ \ \natural\langle s_1, s_2, s_3, \ldots \rangle \\
& & \qquad\qquad\quad\ \textbf{else}\ \ \langle s_0 \rangle \circ \natural\langle s_1, s_2, \ldots \rangle
\end{array}
$$

$$\sigma[\![\boldsymbol{\exists}\, x : F]\!] \;\triangleq\; \exists\, \rho, \tau \in \mathbf{St}^{\infty} : (\natural\sigma = \natural\rho) \wedge (\rho =_x \tau) \wedge \tau[\![F]\!]$$

$$\sigma[\![\exists\, \mathsf{c} : F]\!] \;\triangleq\; \exists \mathsf{c} \in \mathsf{Val} : \sigma[\![F]\!]$$

Proof Rules

$E1.\ \ \vdash F(f/x) \Rightarrow \boldsymbol{\exists}\, x : F$

$E2.\ \ \dfrac{\begin{array}{l} F \Rightarrow G \\ x \text{ does not occur free in } G \end{array}}{(\boldsymbol{\exists}\, x : F) \Rightarrow G}$

$F1.\ \ \vdash F(\mathsf{e}/\mathsf{c}) \Rightarrow \exists\, \mathsf{c} : F$

$F2.\ \ \dfrac{\begin{array}{l} F \Rightarrow G \\ \mathsf{c} \text{ does not occur free in } G \end{array}}{(\exists\, \mathsf{c} : F) \Rightarrow G}$

where x is a $\langle \textit{variable} \rangle$ — F, G are $\langle \textit{general formula} \rangle$s
f is a state function — s, s_0, t_0, s_1, t_1, … are states
c is a $\langle \textit{rigid variable} \rangle$ — σ is a behavior
e is a constant expression — $\circ$ denotes concatenation of sequences

Fig. 9. Quantification in TLA.

The "internal" description, in which *main* and *cache* are free variables, is the formula Ψ of Figure 10. The actions defined in the figure have the following interpretations.

$\mathcal{T}(a, m, v)$. Represents the assignment $a(m) := v$. This action is introduced only to simplify the definitions of other actions.

$\mathcal{R}_{pro}$, $\mathcal{W}_{pro}$. The processor's *read*- and *write*-request operations.

$\mathcal{R}_{cch}$, $\mathcal{W}_{cch}$. The memory's responses to processor requests, the value being read from or written to the cache. An $\mathcal{R}_{cch}$ action can be executed only if the value to be read is in the cache.

$\mathcal{C}_{get}(m)$, $\mathcal{C}_{fl}(m)$. The internal actions of moving a value from main memory to the cache, and of flushing a value from the cache to main memory. The second conjunct of $\mathcal{C}_{fl}(m)$ prevents a value from being flushed while it is being read. This is the only constraint on when values can be moved into or out of the cache; no particular cache maintenance policy is specified.

$\mathcal{P}$. The next-state relation, which is the disjunction of all possible actions of the processor and the memory.

$\mathcal{F}$. The disjunction of all the memory actions that must be performed to respond to a processor request. The third disjunct represents the action of moving the value

$$
\begin{array}{lcl}
\mathit{Init}_\Psi & \triangleq & \land\ op = \text{"ready"} \\
 & & \land\ \forall \mathsf{n} \in \mathsf{Address} : (\mathit{main}(\mathsf{n}) \in \mathsf{MemVal})\ \land\ (\mathit{cache}(\mathsf{n}) = \bot)
\end{array}
$$

$$
\mathcal{T}(a, m, v) \ \triangleq\ \forall \mathsf{n} \in \mathsf{Address} : \begin{array}{l} \land\ (\mathsf{n} = m) \Rightarrow (a'(\mathsf{n}) = v) \\ \land\ (\mathsf{n} \neq m) \Rightarrow (a'(\mathsf{n}) = a(\mathsf{n})) \end{array}
$$

$$
\mathcal{R}_{pro} \triangleq \begin{array}{l} \land\ op = \text{"ready"} \\ \land\ op' = \text{"read"} \\ \land\ adr' \in \mathsf{Address} \\ \land\ \mathit{Unchanged}\ \langle main, cache \rangle \end{array}
\qquad
\mathcal{R}_{cch} \triangleq \begin{array}{l} \land\ op = \text{"read"} \\ \land\ cache(adr) \neq \bot \\ \land\ op' = \text{"ready"} \\ \land\ val' = cache(adr) \\ \land\ \mathit{Unchanged}\ \langle main, cache \rangle \end{array}
$$

$$
\mathcal{W}_{pro} \triangleq \begin{array}{l} \land\ op = \text{"ready"} \\ \land\ op' = \text{"write"} \\ \land\ adr' \in \mathsf{Address} \\ \land\ val' \in \mathsf{MemVal} \\ \land\ \mathit{Unchanged}\ \langle main, cache \rangle \end{array}
\qquad
\mathcal{W}_{cch} \triangleq \begin{array}{l} \land\ op = \text{"write"} \\ \land\ op' = \text{"ready"} \\ \land\ \mathcal{T}(cache, adr, val) \\ \land\ \mathit{Unchanged}\ main \end{array}
$$

$$
\mathcal{C}_{get}(m) \triangleq \begin{array}{l} \land\ cache(m) = \bot \\ \land\ \mathcal{T}(cache, m, main(m)) \\ \land\ \mathit{Unchanged}\ \langle op, adr, \\ \qquad\qquad\qquad val, main \rangle \end{array}
\qquad
\mathcal{C}_{fl}(m) \triangleq \begin{array}{l} \land\ cache(m) \neq \bot \\ \land\ \lor\ op \neq \text{"read"} \\ \quad\ \lor\ m \neq adr \\ \land\ \mathcal{T}(main, m, cache(m)) \\ \land\ \mathcal{T}(cache, m, \bot) \\ \land\ \mathit{Unchanged}\ \langle op, adr, val \rangle \end{array}
$$

$$
\begin{array}{lcl}
\mathcal{P} & \triangleq & \mathcal{R}_{pro}\ \lor\ \mathcal{W}_{pro}\ \lor\ \mathcal{R}_{cch}\ \lor\ \mathcal{W}_{cch}\ \lor\ (\exists\, \mathsf{m} \in \mathsf{Address} : \mathcal{C}_{get}(\mathsf{m}) \lor \mathcal{C}_{fl}(\mathsf{m})) \\
\mathcal{F} & \triangleq & \mathcal{R}_{cch}\ \lor\ \mathcal{W}_{cch}\ \lor\ (\mathcal{C}_{get}(adr) \land (op = \text{"read"})) \\
u & \triangleq & \langle op, adr, val, main, cache \rangle \\
\Psi & \triangleq & \mathit{Init}_\Psi\ \land\ \Box[\mathcal{P}]_u\ \land\ \mathrm{WF}_u(\mathcal{F})
\end{array}
$$

Fig. 10. A simple cached memory.

for a *read* request from main memory into the cache. (It is enabled only if the value is not already in the cache.)

If we consider *main* and *cache* to be internal variables, then the cached memory is described by the TLA formula[5] $\exists\, main, cache : \Psi$. The assertion that the cached memory correctly implements the processor/memory interface is expressed by the formula

$$(\exists\, main, cache : \Psi) \ \Rightarrow\ (\exists\, memory : \Phi) \tag{50}$$

To prove (50), we define the state function $\overline{memory}$ by

$$\overline{memory}(m) \ \triangleq\ \textbf{if } cache(m) = \bot \textbf{ then } main(m) \\ \qquad\qquad\qquad\qquad\qquad\qquad \textbf{else } cache(m)$$

and then prove $\Psi \Rightarrow \overline{\Phi}$, where $\overline{\Phi}$ denotes the formula $\Phi(\overline{memory}/memory)$ obtained by substituting $\overline{memory}$ for all free occurrences of *memory* in Φ. Applying rule E1 of Figure 9, substituting $\overline{memory}$ for f and *memory* for x, we obtain $\Psi \Rightarrow \exists\, memory : \Phi$. Rule E2 then yields (50).

[5]As usual in logic, we write $\exists\, x, y : F$ as an abbreviation for $\exists\, x : \exists\, y : F$, which by E1 and E2 of Figure 9 is equivalent to $\exists\, y : \exists\, x : F$.

The formula $\Psi \Rightarrow \overline{\Phi}$ asserts that any sequence of values for the variables *op*, *adr*, and *val*, and for the state function $\overline{memory}$, that is allowed by Ψ is a sequence of values that Φ allows for the variables *op*, *adr*, *val*, and *memory*. We can regard $\overline{memory}$ as the "concrete" state function with which Ψ implements the "abstract" variable *memory*.

How do we prove that Ψ implies $\overline{\Phi}$? To find the answer, we examine the structure of $\overline{\Phi}$. For any formula F, let $\overline{F}$ denote the formula $F(\overline{memory}/memory)$ obtained by substituting $\overline{memory}$ for all free occurrences of *memory* in F. For example, $\overline{w}$ is the state function $\langle op, adr, val, \overline{memory}\rangle$. Then $\overline{\Phi}$ equals $\overline{Init_\Phi} \wedge \Box[\overline{\mathcal{N}}]_{\overline{w}} \wedge \overline{\mathrm{WF}_w(\mathcal{N}_{mem})}$. The formula $\overline{\Phi}$ therefore looks much like an ordinary TLA formula representing a program, with initial condition $\overline{Init_\Phi}$ and next-state relation $\overline{\mathcal{N}}$. The only difference is that instead of an ordinary weak fairness condition, $\overline{\Phi}$ has as a conjunct the "barred" fairness condition $\overline{\mathrm{WF}_w(\mathcal{N}_{mem})}$.

The proof of $\Psi \Rightarrow \overline{\Phi}$ is similar to the proof in Section 7.2 that Program 2 implements Program 1. We first prove that $Init_\Psi$ implies $\overline{Init_\Phi}$. We next prove that Ψ implies $\Box[\overline{\mathcal{N}}]_{\overline{w}}$ (step-simulation) by applying rule TLA2 of Figure 5 with the substitutions

$$\mathcal{A} \leftarrow \mathcal{P} \qquad \mathcal{B} \leftarrow \overline{\mathcal{N}} \qquad f \leftarrow u \qquad g \leftarrow \overline{w} \qquad P \leftarrow \mathsf{true} \qquad Q \leftarrow \mathsf{true}$$

Finally, we prove that Ψ implies $\overline{\mathrm{WF}_w(\mathcal{N}_{mem})}$ (fairness) by applying WF2 with the substitutions

$$\begin{array}{lll} \mathcal{M} \leftarrow \mathcal{N}_{mem} & \mathcal{A} \leftarrow \mathcal{F} & f \leftarrow u \\ \mathcal{N} \leftarrow \mathcal{P} & \mathcal{B} \leftarrow \mathcal{R}_{cch} \vee \mathcal{W}_{cch} & g \leftarrow w \\ \multicolumn{3}{l}{P \leftarrow (op = \text{“write”}) \;\vee\; (op = \text{“read”} \;\wedge\; cache(adr) \neq \bot)} \end{array}$$

(Observe that Rule WF2 has the appropriate "bars" to prove the desired conclusion.) As in our previous example, the proofs consist of straightforward calculations punctuated by the occasional need for insight into why what we are trying to prove is true.

This cached memory is quite abstract; it allows any policy for deciding when to move values between the cache and main memory. Given a particular caching algorithm, we would prove that it implements the simple cached memory—meaning that the TLA formula representing the algorithm implies $\exists\, main, cache : \Psi$. By the transitivity of implication, this proves that the algorithm implements the memory/processor interface.

8.3.2 *Refinement Mappings.* It is clear how to generalize the example above to the problem of proving

$$(\exists\, x_1, \ldots, x_m : \Psi) \;\Rightarrow\; (\exists\, y_1, \ldots, y_n : \Phi) \tag{51}$$

for arbitrary Ψ and Φ. We must define state functions $\overline{y_1}, \ldots, \overline{y_n}$ in terms of the variables that occur in Ψ and prove $\Psi \Rightarrow \overline{\Phi}$, where for any formula F, we let $\overline{F}$ denote the formula $F(\overline{y_1}/y_1, \ldots, \overline{y_n}/y_n)$ obtained by substituting $\overline{y_i}$ for the free occurrences of y_i in F, for all i. We then infer (51) from rules E1 and E2.

The collection of state functions $\overline{y_1}, \ldots, \overline{y_n}$ is called a *refinement mapping*. The "barred variable" $\overline{y_i}$ is the state function with which Ψ implements the variable y_i of Φ.

To prove (51), one must find a refinement mapping such that $\Psi \Rightarrow \overline{\Phi}$ is valid, and use the rules of Figure 5 to prove its validity. But can the requisite refinement mapping always be found? Does the validity of (51) imply the existence of a refinement mapping such that $\Psi \Rightarrow \overline{\Phi}$ is valid?

The answer is no; a refinement mapping need not exist. As an example, we return to Programs 1 and 2, represented by formulas Φ of Figure 3 and Ψ of Figure 7. Program 2 permits precisely the same sequences of values for x and y as does Program 1. Therefore, the formula $\exists\, sem, pc_1, pc_2 : \Psi$, which describes only the sequences of values for x and y allowed by Program 2, is equivalent to Φ. Can we prove this equivalence?

We already sketched the proof of $\Psi \Rightarrow \Phi$, which by Rule E2 implies

$$(\exists\, sem, pc_1, pc_2 : \Psi) \Rightarrow \Phi$$

In this case, Φ has no internal variables, so the refinement mapping is the trivial one consisting of the empty set of barred variables. Now consider the converse,

$$\Phi \;\Rightarrow\; (\exists\, sem, pc_1, pc_2 : \Psi) \tag{52}$$

Can we define the requisite state functions $\overline{sem}$, $\overline{pc_1}$, and $\overline{pc_2}$ in terms of x and y (the only variables that occur in Φ) so that Program 1 allows them to assume only those sequences of values that Program 2 allows the corresponding variables to assume? Clearly not. There is no way to infer from the values of x and y what the values of sem, pc_1, and pc_2 should be. Thus, there does not exist a refinement mapping for which Φ implies $\overline{\Psi}$.

To prove (52), one must modify Φ by adding *auxiliary variables*. Intuitively, an auxiliary variable is one that is added to a program without affecting the program's behavior. Formally, adding an auxiliary variable d to a formula Π means finding a formula Π^d such that $\exists\, d : \Pi^d$ is equivalent to Π. (The variable d is assumed not to occur free in Π.) Formula (52) can be proved by adding two auxiliary variables h and p to Φ. That is, we can construct a formula Φ^{hp} such that $\exists\, h, p : \Phi^{hp}$ is equivalent to Φ, and can then prove

$$(\exists\, h, p : \Phi^{hp}) \;\Rightarrow\; (\exists\, sem, pc_1, pc_2 : \Psi)$$

by constructing a refinement mapping such that Φ^{hp} implies $\overline{\Psi}$. The refinement mapping can be found because the state functions $\overline{sem}$, $\overline{pc_1}$, and $\overline{pc_2}$ are allowed to depend upon h and p as well as x and y.

In general, refinement mappings can be found if we add the right auxiliary variables. The completeness theorem of Abadi and Lamport [1991] shows that, under certain reasonable assumptions about Φ and Ψ, if (51) is valid, then one can in principle add auxiliary variables to Φ to obtain the formula Φ^{hp} and find the requisite refinement mapping such that $\Phi^{hp} \Rightarrow \overline{\Psi}$ is valid. Relative completeness of STL1–STL6, the Lattice Rule, TLA1, and TLA2 for simple TLA means that this implication is provable from those rules if Φ and Ψ have the form $Init \wedge \Box[\mathcal{N}]_v \wedge F$, where F is the conjunction of weak and strong fairness formulas, and Φ is machine closed. We thus have a relative completeness result for TLA formulas of the form (51).

8.3.3 *"Barring" Fairness.* The "barring" operator denotes substitution of state functions $\overline{y_i}$ for variables y_i. Barring distributes over most of our operators; for example, $\overline{\Box(F \vee G)}$ equals $\Box(\overline{F} \vee \overline{G})$, for any formulas F and G. Thus, when Φ has the canonical form $Init \wedge \Box[\mathcal{N}]_f \wedge F$, the formula $\overline{\Phi}$ equals $\overline{Init} \wedge \Box[\overline{\mathcal{N}}]_{\overline{f}} \wedge \overline{F}$. If F

is the conjunction of fairness conditions of the form $\mathrm{WF}_g(\mathcal{M})$ and $\mathrm{SF}_g(\mathcal{M})$, then $\overline{F}$ is the conjunction of barred fairness conditions $\overline{\mathrm{WF}_g(\mathcal{M})}$ and $\overline{\mathrm{SF}_g(\mathcal{M})}$.

We might expect that $\overline{\mathrm{WF}_g(\mathcal{M})}$ would be equivalent to $\mathrm{WF}_{\overline{g}}(\overline{\mathcal{M}})$ and $\overline{\mathrm{SF}_g(\mathcal{M})}$ equivalent to $\mathrm{SF}_{\overline{g}}(\overline{\mathcal{M}})$, but that need not be the case. It is true that

$$\begin{aligned} \overline{\mathrm{WF}_g(\mathcal{M})} &\equiv \Box\Diamond\neg\overline{\mathit{Enabled}\ \langle\mathcal{M}\rangle_g} \ \lor\ \Box\Diamond\langle\overline{\mathcal{M}}\rangle_{\overline{g}} \\ \overline{\mathrm{SF}_g(\mathcal{M})} &\equiv \Box\Diamond\neg\overline{\mathit{Enabled}\ \langle\mathcal{M}\rangle_g} \ \lor\ \Diamond\Box\langle\overline{\mathcal{M}}\rangle_{\overline{g}} \end{aligned} \tag{53}$$

However, $\overline{\mathit{Enabled}\ \langle\mathcal{M}\rangle_g}$ is not necessarily equivalent to $\mathit{Enabled}\ \langle\overline{\mathcal{M}}\rangle_{\overline{g}}$. For example, let $\mathcal{M}$ be the action $(x' = x) \land (y' \neq y)$, let g equal $\langle x, y\rangle$, and let the refinement mapping be defined by $\overline{x} = z$ and $\overline{y} = z$. Then $\mathit{Enabled}\ \langle\mathcal{M}\rangle_g$ equals $\exists\, \mathsf{c}, \mathsf{d} : (\mathsf{c} = x) \land (\mathsf{d} \neq y)$, which equals true. Hence $\overline{\mathit{Enabled}\ \langle\mathcal{M}\rangle_g}$, the formula obtained by substituting $\overline{x}$ for x and $\overline{y}$ for y in $\mathit{Enabled}\ \langle\mathcal{M}\rangle_g$, equals true. But

$$\begin{aligned} &\mathit{Enabled}\ \langle\overline{\mathcal{M}}\rangle_{\overline{g}} \\ &\quad\equiv \mathit{Enabled}\ \langle\overline{(x' = x) \land (y' \neq y)}\rangle_{\overline{\langle x,\, y\rangle}} && \text{by definition of } \mathcal{M} \text{ and } g \\ &\quad\equiv \mathit{Enabled}\ \langle(\overline{x}' = \overline{x}) \land (\overline{y}' \neq \overline{y})\rangle_{\langle\overline{x},\overline{y}\rangle} && \text{by definition of } \overline{\ldots} \\ &\quad\equiv \mathit{Enabled}\ \langle(z' = z) \land (z' \neq z)\rangle_{\langle z,z\rangle} && \text{by definition of } \overline{x} \text{ and } \overline{y} \\ &\quad\equiv \mathit{Enabled}\ \mathsf{false} && \text{by definition of } \langle\ldots\rangle_{\ldots} \\ &\quad\equiv \mathsf{false} && \text{by definition of } \mathit{Enabled} \end{aligned}$$

Thus, $\overline{\mathit{Enabled}\ \langle\mathcal{M}\rangle_g}$ is not equivalent to $\mathit{Enabled}\ \langle\overline{\mathcal{M}}\rangle_{\overline{g}}$. In general, the primed variables in the action $\langle\mathcal{M}\rangle_g$ are not free variables of the expression $\mathit{Enabled}\ \langle\mathcal{M}\rangle_g$, so we can't obtain $\overline{\mathit{Enabled}\ \langle\mathcal{M}\rangle_g}$ from $\mathit{Enabled}\ \langle\mathcal{M}\rangle_g$ by blindly barring all variables.

In rules WF2 and SF2, the formulas $\overline{\mathrm{WF}_g(\mathcal{M})}$ and $\overline{\mathrm{SF}_g(\mathcal{M})}$ are defined by (53). The rules are sound when $\overline{\mathcal{M}}$ is any action, $\overline{g}$ any state function, and $\overline{\mathit{Enabled}\ \langle\mathcal{M}\rangle_g}$ any predicate—assuming that $\overline{\mathrm{WF}_g(\mathcal{M})}$ and $\overline{\mathrm{SF}_g(\mathcal{M})}$ are defined by (53). In practice, the barred formulas will be obtained from unbarred ones by substituting barred variables (state functions) for variables, as in our example.

9. FURTHER COMMENTS

9.1 Mechanical Verification

Because it is a simple logic, TLA is ideally suited for mechanization. Urban Engberg and Peter Grønning have been working on the mechanical verification of TLA, using LP—an "off-the-shelf" verification system based on rewriting [Garland and Guttag 1989]. Although initial experiments showed that LP can be used directly, Engberg and Grønning decided to develop a system called TLP to translate TLA definitions and proofs into LP input [Engberg et al. 1992] (Note 17). In addition to allowing more readable specifications, TLP allows separate LP proofs for action formulas and temporal formulas, using simpler encodings of the formulas than would be possible with a single proof. Since most reasoning in a TLA proof is about actions, a simple encoding of action formulas is important. They also hope to use verification systems other than LP to check parts of the proof.

The proof in Section 7.2, that the formula Ψ describing Program 2 implies the formula Φ describing Program 1, has been checked with TLP. Figure 11 shows the definitions of Φ and Ψ in the actual TLP input. (For simplicity, we are omitting some declarations and TLP directives.) Observe that these definitions are almost

```
InitPhi ==  (x = 0) /\ (y = 0)
M1      ==  (x' = x + 1) /\ (y' = y)
M2      ==  (y' = y + 1) /\ (x' = x)
M       ==  M1 \/ M2
v       ==  (x * y)
Phi     ==  InitPhi /\ [][M]_v /\ WF(v,M1) /\ WF(v,M2)

InitPsi  ==  /\ (pc1 = a) /\ (pc2 = a)
             /\ (x = 0) /\ (y = 0)
             /\ sem = 1
alpha1   ==  /\ (pc1 = a) /\ (0 << sem)
             /\ pc1' = b
             /\ sem' = sem - 1
             /\ Unchanged(x * y * pc2)

          ⋮

gamma2   ==  /\ pc2 = g
             /\ pc2' = a
             /\ sem' = sem + 1
             /\ Unchanged(x * y * pc1)
N1       ==  alpha1 \/ beta1 \/ gamma1
N2       ==  alpha2 \/ beta2 \/ gamma2
N        ==  N1 \/ N2
w        ==  (x * y * pc1 * pc2 * sem)
Psi      ==  InitPsi /\ [][N]_w /\ SF(w,N1) /\ SF(w,N2)
```

Fig. 11: The representation of the formulas Φ and Ψ of Figures 3 and 7 for the mechanical verification of the theorem $\Psi \Rightarrow \Phi$.

perfect transliterations of the ones in Figures 3 and 7. The major differences are the use of "*" to represent tuples and "<<" instead of "<"—differences introduced because comma and "<" have other meanings.

Following these definitions is the statement

```
Theorem  Psi  =>  Phi
```

that asserts the validity of the temporal formula $\Psi \Rightarrow \Phi$. The rest of the TLP input is a hierarchically structured proof of this theorem. TLP translates the definitions and proof into a form that can be checked by LP. The only part of the proof not checked by LP is the computation of the *Enabled* predicates. Although algorithmically simple, these computations are awkward to do in LP. We hope that a future version of TLP will compute *Enabled* predicates.

The work on mechanically verifying TLA formulas is preliminary. So far, only simple examples have been completed. The verification described above was done with an early version of the preprocessor, whose implementation required about two man-months of effort. The goal of the project is to assess the feasibility of implementing a verification system that will be useful for real problems. Recent work has concentrated on developing a convenient user interface for managing proofs, which we feel is a prerequisite for a practical system.

9.2 TLA versus Programming Languages

Let us compare Figure 6, the description of Program 2 in a conventional programming language, with Figure 7, its representation as a TLA formula. At first glance, the program looks simpler than the TLA formula. However, the program seems simple only because you are already familiar with its notation. To understand what the program means, you need to understand the meaning of the **var** declarations, the **cobegin**, **loop**, ";", and ":=" constructs, and the P and V operations. In contrast, everything needed to understand the TLA formula appears in Figure 4. It is easy to make something seem simple by omitting the complicated definitions needed to understand it.

One reason for the conventional program's apparent simplicity is that it does not specify liveness properties. Nothing in Figure 6 told us that the fairness requirement for Ψ should be strong fairness $(\mathrm{SF}_w(\mathcal{M}_1) \wedge \mathrm{SF}_w(\mathcal{M}_2))$ rather than weak fairness $(\mathrm{WF}_w(\mathcal{M}))$. To allow either fairness requirement, a programming language should provide different flavors of **cobegin** and semaphore operations. If the language provides only one kind of fairness, specifying a different fairness requirement needs a complicated encoding with additional variables—if it is even possible.

The TLA formula can be made shorter and easier to read by introducing some simple definitions. The representation of program control can be encapsulated by defining $Go(i, d, e)$ to mean that control in process i goes from d to e (Note 18):

$$Go(i, d, e) \;\triangleq\; (pc_i = d) \wedge (pc'_i = e) \wedge (pc'_{3-i} = pc_{3-i}) \tag{54}$$

The semaphore operations can be expressed more compactly by defining $P(sem)$ to equal $(0 < sem) \wedge (sem' = sem - 1)$ and $V(sem)$ to equal $sem' = sem + 1$. The definition of Ψ then appears much simpler; for example α_1 is defined by

$$\alpha_1 \;\triangleq\; P(sem) \wedge Go(1, \text{"a"}, \text{"b"}) \wedge Unchanged\ \langle x, y\rangle$$

The ability to use definitions to simplify formulas makes TLA practical for large specifications.

There are just two basic reasons why a TLA formula is longer than the corresponding conventional program: (i) what remains unchanged is implicit in a program statement, but must be stated explicitly in an action definition; and (ii) how the control state changes is implicit in the program, but is described explicitly in the formula. We now discuss these two sources of length.

The explicit *Unchanged* clauses add only about 10% to the length of the definition of Ψ in Figure 7. In larger examples they add less; about 1% of one 700-line TLA specification consists of *Unchanged* clauses. Still, why pay that price? An obvious way of simplifying the formulas is to let the omission of a variable from an action mean that the variable is left unchanged. Thus, $x' = x + 1$ would be equivalent to $(x' = x+1) \wedge (y' = y)$. However, this "simplification" would in fact make TLA much more complicated. For example, it would mean that the obviously true formula $y' = y'$ is not equivalent to true, since the formulas $(x' = x + 1) \wedge (y' = y')$ and $x' = x+1$ would not be equivalent—the first would allow y to change and the second would not. Like ordinary mathematics, TLA is simple because a formula constrains only the variables that it explicitly mentions. This is what makes $x' = x + 1$ so much simpler than $x := x + 1$. Writing *Unchanged* clauses is a small price to pay for the simplicity of ordinary mathematics.

The control structures of ordinary programming languages provide a convenient method of specifying that operations are to be performed in a particular order. Specifying this in TLA requires the use of explicit control variables, which are a source of complexity. However, remember that we are interested in reasoning about abstract descriptions of algorithms, not C code. An abstract algorithm usually has few separate actions, so its control structure is simple; if control variables are needed, they usually add little complexity.

In abstract algorithms, it is just as common to specify that actions can occur in any order as it is to specify that they occur in some particular order. Conventional languages make it awkward to allow operations to occur in any order. Dijkstra's guarded commands provide a simple mechanism for allowing nondeterminism, but they lack a convenient way to specify the grain of atomicity in the evaluation of guards. We urge the reader to code the cache example of Figure 10 in his favorite programming language. The precise liveness condition will probably be very difficult or even impossible to express within the language. Even ignoring the liveness condition, we expect that the TLA formula will be simpler than the program.

Any language will be better than TLA at representing a program written especially for that language. Furthermore, a familiar notation, no matter how cumbersome, invariably seems simpler than an unfamiliar one. Our experience suggests that after one gets used to its notation, the TLA description of a "randomly chosen" algorithm is likely to seem simpler than its representation in a conventional programming language, though it may be longer. (If brevity were synonymous with simplicity, APL would be easier to read than Pascal.)

9.3 Reduction

An algorithm must ultimately be translated into a computer program. One develops a program through a series of refinements, starting from a high-level algorithm and eventually reaching a low-level program. Just as we went from Program 1 to the finer-grained Program 2, and from the simple processor/memory interface to the more complicated cached memory, the entire process from specification to C code could in principle be carried out in TLA. "All" we would need is a precise semantics of C, which would allow the translation of any C program into a TLA formula.

In practice, the refinement will be carried out in TLA until it becomes obvious how to hand-translate the TLA formula into a program in a real programming language—one with a compiler that produces satisfactory code. But what does it mean for the translation to be obvious? From the point of view of concurrency, the translation from the TLA formula to the program is obvious when any step of the next-state relation corresponds to an atomic operation of the program. In this sense, the translation from an action $(sem' = sem + 1) \wedge \mathit{Unchanged}\ \langle \ldots \rangle$ to an atomic $V(sem)$ program statement is obvious.

Real programming languages usually guarantee only an extremely fine grain of atomicity. When executing the statement $x := x + 1$, the read and write of x might each consist of several atomic operations. It would be impractical to describe such a fine-grained program with a TLA formula. Instead, one refines the TLA formula to the point where each step of the next-state relation either corresponds to an atomic program operation like $V(sem)$, or else can be implemented with any grain of atomicity—for example, because it occurs inside an appropriate critical section.

When can an atomic operation be implemented with any grain of atomicity? To answer this, we must first ask: when does a fine-grained program implement a coarser-grained one? There have been a number of partial answers to this question. Some lie in folk theorems—for example, that if shared variables are accessed only in critical sections, then an entire critical section is equivalent to a single atomic operation. Other answers lie in precise results stating that certain classes of properties are satisfied by a fine-grained program if they are satisfied by a coarser-grained version [Lipton 1975].

The question of when a fine-grained program implements a coarser-grained one is answered in TLA by a "reduction" theorem. This theorem seems to include all prior answers as special cases—both the folk theorems and the precise results. The precise statement of the theorem is somewhat complicated, and will be given elsewhere. Here, we give only a rough description of what it says. The theorem's conclusion is approximately

$$\Phi \;\Rightarrow\; \exists\, w_1, \ldots, w_n : \Phi_{red}(w_1/v_1, \ldots, w_n/v_n) \;\wedge\; \Box R \tag{55}$$

where

Φ is the simple TLA formula (with no hidden variables) describing the original program.

Φ_{red} is the coarser-grained "reduced" version of the program.

$v_1, \ldots, v_n$ are all the variables that occur in Φ and Φ_{red}.

R is a predicate containing the variables w_i and v_i.

Think of the v_i as "real" variables and the w_i as "pretend" variables. Formula (55) asserts that there exist pretend variables such that the original program operating on the real variables implements the reduced program operating on the pretend variables, and the relation R always holds between the real and the pretend variables.

In applying the reduction theorem to critical sections, the reduced formula Φ_{red} is obtained from the original formula Φ by changing the next-state relation to turn an entire execution of a critical section into a single step. The relation R asserts that the real and the pretend variables are equal when no process is in its critical section.

In practice, one reasons about the reduced formula Φ_{red} and checks that (55) implies the correctness of the fine-grained formula Φ. For example, in the critical-section application, if a property does not depend on the values assumed by variables while processes are in their critical sections, then Φ satisfies the property if Φ_{red} does. One must then verify that the formula Φ representing the actual program satisfies the hypotheses of the Reduction Theorem, without actually writing Φ. For complicated languages like C, which lack a reasonable formal semantics, this verification must be informal. Whether formal verification is practical, with either a new language or a useful subset of an existing one, is a topic for research.

9.4 What is TLA Good For?

TLA, like any useful formal system, has a limited domain of applicability. A formalism that encompasses everything is good for nothing. We believe that TLA is useful for specifying and verifying safety and liveness properties of discrete systems.

Intuitively, a safety property asserts that something bad does not happen, and a liveness property asserts that something good does eventually happen.

We feel that the most significant limitation of TLA is that TLA properties are true or false for an individual behavior. Thus, one cannot express statistical properties of sets of behaviors—for example, that the program has probability greater than .99 of terminating. The only way we know of verifying such properties is to construct a formal model of the system, use TLA to verify that the system correctly implements the model, and then apply other techniques such as Markov analysis to verify that the model has the desired property.

The limited expressiveness of TLA is not always a disadvantage. As we have seen, TLA allows fine-grained implementations of coarser-grained specifications because it can express only properties that are invariant under stuttering. A formalism that distinguished between doing nothing and taking a step that produces no change would seem to have a tenuous relation to reality. Another class of properties whose inexpressibility in TLA causes us no concern are possibility properties. We have never found it useful to be able to assert that it is possible for a system to produce the right answer. We want to assert that, under certain assumptions, the system *must* produce the right answer.

TLA can be used to reason about a discrete system even if its behavior depends upon continuous physical values. A particularly important physical value is time. Best- and worst-case time bounds on algorithms can be expressed as safety properties and proved with TLA. For example, the assertion that an algorithm always terminates within 15 seconds is a safety property, where time having advanced 15 seconds without the algorithm having terminated is the "something bad" that does not happen. A real-time algorithm can be specified by conjoining timing constraints to the TLA specification of the untimed version of the algorithm. A description of how TLA is used to reason about real time appears in [Abadi and Lamport 1992]. The use of TLA for hybrid systems is described in [Lamport 1993].

9.5 What We Have Omitted

9.5.1 *Program Derivation.* Derivation of a program by a rigorous procedure that guarantees its correctness is preferable to post hoc verification. Concurrent algorithms are derived by refining higher-level, coarser-grained algorithms to lower-level, finer-grained ones. Refinement is the same as implementation—Ψ refines Π means that Ψ implements Π. Any method for proving that one program implements another can be used as the basis for program derivation.

Formalisms based on a programming language can usually prove only that a program satisfies certain properties, not that one program implements another. In such formalisms, refinement is either done informally or with special-purpose rules. In TLA, refinement is implication, and we have shown how one can prove that a finer-grained algorithm refines a coarser-grained one. Moreover, some traditional refinement steps can be performed in TLA by applying standard mathematical laws to rewrite formulas. We believe that TLA should be at least as good as any other formalism for deriving concurrent algorithms. Unfortunately, we know of no concurrent algorithm used in a real system that was systematically derived, not simply justified by a post hoc derivation. The derivation of concurrent algorithms is still in the realm of research.

9.5.2 *"True" Concurrency.* TLA is based on an interleaving model of concurrency, in which we assume that an execution of the system consists of a sequence of atomic events. It seems paradoxical to represent concurrent systems with a formalism in which events are never concurrent. We will not attempt to justify the philosophical correctness of interleaving models for reasoning about concurrent algorithms. Instead, we have tried to demonstrate the best reason we know for using TLA: it is a practical formalism for specifying and verifying safety and liveness properties.

9.5.3 *Open Systems.* We have discussed only closed systems. A closed system is one that is completely self-contained—in contrast to an open system, which interacts with its environment. Any real system is open; it does not eternally contemplate its navel, oblivious to the outside world. But for most purposes, one can model the actual system together with its environment as a single closed system—as we did for the memory example of Section 8. Such an approach is generally adequate for reasoning about algorithms. However, some problems can be studied only in the context of open systems. For example, composing component systems to form one large system makes sense only for components that are open systems.

TLA can be used to describe and reason about open as well as closed systems. But closed systems are simpler, and they provide a necessary foundation for the study of open systems. Here, we have developed TLA and applied it to closed systems. Open systems are discussed elsewhere [Abadi and Lamport 1993].

9.5.4 *System Specifications.* Most readers would expect a two-page Pascal program to be simple and a two-page mathematical formula to be too complicated to understand. Yet, since the semantics of TLA is simpler than the semantics of Pascal, a TLA formula should be simpler than a Pascal program of the same length. The main reason mathematical formulas seem to get very complicated when they get large is that mathematicians have not developed notations for structuring large formulas. We have introduced some simple conventions that make the TLA formulas describing abstract algorithms as easy to read as the corresponding programs.

TLA can be used not just to describe abstract algorithms, but also to specify complex systems. System specifications can be dozens or even hundreds of pages long. Managing the complexity of large specifications requires additional notation for modular structuring. We have added such notation to TLA to form a language called TLA$^+$, a purely syntactic extension to TLA with nothing new semantically. TLA$^+$ will be described elsewhere.

10. CONCLUSIONS

10.1 Historical Note

TLA is in the tradition of assertional methods for reasoning about programs. These methods go back to Floyd [1967], who first proved partial correctness and termination of sequential programs. Hoare [1969] recast partial correctness reasoning into a logical framework. The first practical assertional method for reasoning about concurrent programs was proposed by Ashcroft [1975]. Ashcroft's work was followed by a number of variations on the same theme [Flon and Suzuki 1978; Keller 1976; Lamport 1977]; but the one that became popular is the Owicki/Gries method, developed by Susan Owicki in her thesis [Owicki 1975], which was supervised by David

Gries. All these methods, though clothed in different notations, proved safety properties by the use of an invariant; they would be described in TLA as applications of Rule INV1.

Temporal logic was first used to reason about concurrency by Pnueli [1977]. It provided the first practical approach to proving more general liveness properties than simple termination. Pnueli introduced the simple temporal logic described in Section 3, with predicates as the only elementary formulas. Pnueli's logic was not expressive enough to describe all desired properties. It was followed by a plethora of proposals for more expressive logics, all obtained by introducing more powerful temporal operators. Pnueli [1979] was the first to describe a program by a temporal logic formula. He, and almost everyone else who followed him, represented programs by formulas that are not invariant under stuttering, so a finer-grained program could not implement a coarser-grained one. The observation that invariance under stuttering permits refinement first appeared in [Lamport 1983b].

The current use of primed and unprimed variables (or their equivalent) for describing "before" and "after" states of a program probably goes back to the early 1970s; we do not know where it first appeared. The idea of actually specifying a program operation by a relation between primed and unprimed variables appears to have been introduced independently by us [Lamport 1983a], Hehner [1984], and Shankar and Lam [1984]. These approaches all used the convention that variables not mentioned are not changed, so they had the inherent complexity epitomized by the observation that $y' = y'$ is not equivalent to true.

10.2 Comparison with Related Formalisms

The correctness of an algorithm does not depend on the formalism in which the algorithm and its properities are expressed. A proof of correctness should be essentially the same regardless of the formalism in which it is expressed. A sufficiently informal proof can usually be expressed quite easily in any formalism. Formalisms tend to differ in the ease with which the proof can be formalized. They also differ in how practical they are for writing large specifications. (Formal verification of large specifications is a difficult and rarely attempted task; it seems premature to draw any conclusions about its practicality.)

10.2.1 *Methods Based on a Program Text.* Extensions of the standard Floyd/Hoare method to concurrent programs prove only invariance properties and termination [Apt and Olderog 1990]. Methods for proving more general temporal properties have been developed for Unity [Chandy and Misra 1988] and other toy languages [Manna and Pnueli 1991]. It is straightforward to translate a proof in any of these methods into a TLA proof. However, none of the common approaches based on proving properties of a program text can express the concept of one program implementing another.

Toy programming languages are not very good for representing real algorithms, even if the toy language resembles the language in which the algorithm is implemented. The most convenient abstraction of a real program may require language features, such as atomic operations on complex data structures, that are seldom provided by toy languages. Moreover, one must reason about the complete system, including components such as file servers and communication lines that are not part of the program itself. A language designed for expressing programs may be ill suited for describing other components of the system. In particular, it can be awkward, or even impossible, to express the complex liveness properties needed to describe these components with the simple fairness assumptions built into a programming language.

Programming languages also lack the abstraction mechanisms needed to manage the complexity of large specifications. Consider how we encapsulated the lower-level details of Program 2 by defining Go, P, and V in Section 9.2. Such definitions are a common method of structuring TLA specifications. In a programming language, Go, P, and V would correspond to procedures with side effects that are executed as part of the current atomic operation. Such procedures are missing from all programming languages, both toy and real, that we know of.

10.2.2 *Automata-Based Approaches.* The shortcomings of conventional programming languages inspired the use of various forms of abstract automata [Lam and Shankar 1984; Lynch and Tuttle 1987]. Liveness properties of the automata are expressed either by fairness conditions on sets of actions or by formulas in a standard temporal logic.

A suitable language for describing automata should permit abstractions such as Go. However, automata-based methods lack the ability to manipulate specifications as mathematical formulas. For example, it seems impossible to write a real-time specification by expressing timing constraints as an automaton that is combined with the untimed specification.

Specifying liveness properties of automata is problematic. Although more expressive than the built-in fairness assumptions of a programming language, fairness conditions on sets of actions cannot conveniently specify all desired liveness properties. Temporal logic is expressive enough, but we know of no way to check for machine closure if the fairness requirement can be an arbitrary temporal logic formula. As explained in Section 5.3, the lack of machine closure often indicates an error.

Automata-based methods can prove that one specification implements another. The proofs should be essentially the same as the corresponding proof in TLA. We do not know of any completely formal proof system for an automata-based method, so we cannot say what problems may arise in formalizing the proofs. There is reason to believe that machine closure of specifications is necessary for completeness [Abadi and Lamport 1991]. The method of Lam and Shankar [1984] is known to be incomplete because it lacks rules for introducing certain kinds of auxiliary variables.

10.2.3 *Temporal Logics.* Many forms of temporal logic have been proposed for specifying and reasoning about concurrent algorithms. The most popular ones are probably Unity logic [Chandy and Misra 1988] and the logic of Manna and Pnueli [1991]. These logics share with TLA all the advantages that come from representing an algorithm as a formula.

TLA differs from other temporal logics because it is based on the principle that temporal logic is a necessary evil that should be avoided as much as possible. Temporal formulas tend to be harder to understand than formulas of ordinary first-order logic, and temporal logic reasoning is more complicated than ordinary mathematical (nonmodal) reasoning.

A typical TLA specification contains one $\Box$, one $\boldsymbol{\exists}$, and a few WF and/or SF formulas—even if the specification is hundreds of lines long. Most of the specification consists of definitions of predicates and actions. Temporal reasoning is reduced to a bare minimum; it is used almost exclusively for proving liveness properties. Formalizing liveness proofs is temporal logic's forte.

TLA avoids the extensive use of temporal operators that characterizes other temporal logic specification methods by relying instead on internal variables—variables that are hidden with the existential quantifier $\exists$. With any temporal logic, existential quantification over flexible variables is neeeded to express most nontrivial specifications. For example, it is necessary for specifying an n-element buffer, where n is a parameter (free rigid variable) of the specification (Note 19). Unity logic lacks a hiding operator, so it is inadequate as a formal specification language. Instead, Unity is used informally by adding the necessary internal variables and pretending that they are hidden [Misra 1990].

All logics that include existential quantification over flexible variables in principle have essentially the same expressive power. TLA can express all formulas invariant under stuttering that Manna and Pnueli's logic can. However, their logic can also express formulas that are not invariant under stuttering. Such formulas yield specifications that cannot be refined. Although all TLA formulas are expressible in Manna and Pnueli's logic, there is no simple translation from TLA to their logic because its quantification operator is not invariant under stuttering.

Notes

Note 1. We find set theory to be the most natural basis for a logic of actions, in which case Val is the collection of all sets. However, a "smaller" collection Val is adequate for many purposes.

Note 2. Formally, $+$ is an *operator*, and $x + y$ is an abbreviation for $+(x, y)$. An expression is either a variable, a constant symbol, or an expression of the form $o(e_1, \ldots, e_n)$ where o is an operator and the e_i are expressions. An operator o has a meaning $[\![o]\!]$, and the meaning of an expression is defined inductively—for example, $s[\![+(e_1, e_2)]\!]$ equals $[\![+]\!](s[\![e_1]\!], s[\![e_2]\!])$. If we base the logic on set theory, all the operators we need, such as $+$, can be defined in terms of the following four primitive ones: $\wedge$, $\neg$, $\in$, and ε (Hilbert's "choice" operator [Leisenring 1969]). In the discussion, we do not distinguish between $+$ and $[\![+]\!]$, so we write $s[\![x + y]\!] = s[\![x]\!] + s[\![y]\!]$.

Note 3. When proving the validity of an action by ordinary reasoning, x and x' must be considered distinct variables. For example, let $\mathcal{A}$ be the action $(x = y) \Rightarrow (x' = y')$. Naive substitution of equals for equals might lead one to think that $\mathcal{A}$ is valid. However, $s[\![\mathcal{A}]\!]t$ does not equal true if s is a state such that $s[\![x]\!] = s[\![y]\!]$ and t is a state such that $t[\![x]\!] \neq t[\![y]\!]$.

Note 4. Just as we do not bother distinguishing the constant symbol Nat from its meaning $[\![\mathsf{Nat}]\!]$, the set of naturals, we do not distinguish the rigid variables m and n from their meanings, which are first-order variables of the semantics.

Note 5. Formally, we should distinguish actions from temporal formulas and boolean operators on actions from boolean operators on temporal formulas. Letting

$\theta(\mathcal{A})$ denote the temporal formula that we now write $\mathcal{A}$, we should rewrite (9) as

$$\langle s_0, s_1, s_2, \ldots \rangle [\![\theta(\mathcal{A})]\!] \stackrel{\Delta}{=} s_0 [\![\mathcal{A}]\!] s_1$$

Letting $\vee$ denote disjunction of temporal formulas, we would then notice that the temporal formula we now write $\mathcal{A} \vee \mathcal{B}$ can denote either $\theta(\mathcal{A} \vee \mathcal{B})$ or $\theta(\mathcal{A}) \vee \theta(\mathcal{B})$. However, these two formulas are equivalent, which is why we can get away with writing $\mathcal{A}$ instead of $\theta(\mathcal{A})$ and using the same symbols for boolean operators on actions and temporal formulas.

Note 6. Formula Φ of Figure 2 asserts that every step of Program 1 increments either x or y, but not both. We could allow simultaneous incrementing of x and y by simply redefining $\mathcal{M}$ to equal $\mathcal{M}_1 \vee \mathcal{M}_2 \vee \mathcal{M}_{12}$, where

$$\mathcal{M}_{12} \stackrel{\Delta}{=} (x' = x + 1) \wedge (y' = y + 1)$$

However, there is no reason to complicate Φ in this way. In representing the execution of $x := x + 1$ by a single step, we are already modeling a complex operation as one event. Nothing would be gained by allowing the additional possibility of representing the executions of two separate statements as a single step.

Note 7. To write the state function $\langle x, y \rangle$, we must assume that any pair of values is a value. More generally, we assume that $\langle \mathsf{c}_1, \ldots, \mathsf{c}_n \rangle$ is a value, for any values $\mathsf{c}_1, \ldots, \mathsf{c}_n$.

Note 8. More precisely, we define $\Diamond \langle \mathcal{A} \rangle_f$ to be an abbreviation for $\neg \Box [\neg \mathcal{A}]_f$. The calculation shows that $\neg \Box [\neg \mathcal{A}]_f$ is equivalent to the RTLA formula $\Diamond \langle \mathcal{A} \rangle_f$ obtained from the definitions (7) of $\Diamond$ and (15) of $\langle \mathcal{A} \rangle_f$.

Note 9. We have told a white lie; $\mathcal{M}_1$ is not equivalent to $\langle \mathcal{M}_1 \rangle_{\langle x, y \rangle}$. For example, suppose there is a value ∞ such that $\infty + 1$ equals ∞, and let s be a state in which x has the value ∞. Then the pair s, s is an $\mathcal{M}_1$ step, but not an $\langle \mathcal{M}_1 \rangle_{\langle x, y \rangle}$ step. However, it is true that definitions (14) and (16) are equivalent, because $\mathit{Init}_\Phi \wedge \Box[\mathcal{M}]_{\langle x, y \rangle}$ implies that the values of x and y are always natural numbers, and $\mathsf{n} + 1 \neq \mathsf{n}$ is true for any natural number n.

Note 10. Observe that *Enabled* $\langle \mathcal{M}_1 \rangle_{\langle x, y \rangle}$ is not equivalent to true. For example, $\langle \mathcal{M}_1 \rangle_{\langle x, y \rangle}$ is not enabled in a state in which x equals ∞ (see Note 9). However, $\langle \mathcal{M}_1 \rangle_{\langle x, y \rangle}$ is enabled in any state in which x is a natural number, so $\mathit{Init}_\Phi \wedge \Box[\mathcal{M}]_{\langle x, y \rangle}$ implies $\Box \mathit{Enabled}\ \langle \mathcal{M}_1 \rangle_{\langle x, y \rangle}$. Hence, (16) and the definition of Φ in Figure 3 are equivalent.

Note 11. The hypothesis of STL1 means that F is a propositional tautology or is derivable by the laws of propositional logic from provable formulas.

Note 12. It is somewhat surprising that this completeness result holds even though rules STL1–STL6 are not enough to prove all tautologies of simple temporal logic.

Note 13. It is not necessary for "abc" $\in \mathsf{Nat}$ to be false. Formula (30) is true even if "abc" should happen to equal 135. By not assuming that strings and numbers are disjoint sets, we allow implementations in which strings and numbers share a common representation—for example, as strings of bits. We do, however, assume that "abc" does not equal "xyz".

Note 14. As an example of the problems introduced by types, consider a variable b of type "array of boolean". If e is an expression, then $b[e]$ is a state predicate. What is the meaning of this state predicate when the value of e is not in the index set of b? There are several possible answers, none of which we find pleasant. This problem is not addressed in the books by Manna and Pnueli [1991] and Chandy and Misra [1988], which allow boolean array variables. It is this very problem that led us to make booleans distinct from values, so state predicates are distinct from state functions.

Note 15. In the definition of $\mathcal{S}(m, v)$, the symbols m and v are parameters. The expression $\mathcal{S}(adr, val)$ denotes the formula obtained by substituting adr for m and val for v in this definition.

Note 16. Quantification over rigid variables can be defined in terms of quantification over flexible variables by

$$\exists\, \mathsf{c} : F \;\triangleq\; \boldsymbol{\exists}\, x : F(x/\mathsf{c}) \land \Box[\mathsf{false}]_x$$

where c is a rigid variable and x is any flexible variable that does not occur in the temporal formula F.

Note 17. The fact that TLA is a typeless logic and LP is typed caused no problem; just two basic LP types, corresponding to TLA's values and booleans, were used. The actual LP proofs are somewhat cumbersome because LP is designed to use types instead of sets. However, translating from a typed version of TLA into LP would be quite difficult—unless one used LP's type system for TLA. Adopting LP's type system, with its lack of subtyping and union types, would make it quite awkward to represent many simple algorithms in TLA.

Note 18. To be rigorous, we would have to write $pc[i]$ instead of pc_i, replacing the two variables pc_1 and pc_2 by a single "array" variable pc. (Formally, pc is a variable whose value is a function with domain $\{1, 2\}$.)

Note 19. A specification of a buffer should describe the interface operations of adding and removing elements from the buffer, but should not describe the internal state of the buffer. In all the temporal logics we know of, the only way to specify an n-element buffer is to describe how the interface operations interact with the internal state, and then hide the internal state by existential quantification.

ACKNOWLEDGMENTS

Martín Abadi helped to develop TLA and is the source of many of the ideas presented here, including the use of an untyped logic. Frits Vaandrager helped me clarify many parts of the exposition and made a valiant but unsuccessful attempt to persuade me that types are useful in logic. Fred Schneider, Reino Kurki-Suonio, Urban Engberg, and Peter Grønning caught a number of errors in earlier versions. Bryan Olivier prevented me from augmenting STL1–STL6 with a redundant rule. A study group at Aarhus University that included Douglas Gurr, Carolyn Brown, and Henrik Andersen provided helpful comments. Jeff Line, Peter Hancock, Dominique Mery, and Jack Wiledon discovered some typographical errors. Peter Ladkin requested Note 11. Luca Cardelli and Cynthia Hibbard suggested some improvements. I was helped in preparing the current version by two refer-

ees, whose carefully considered and detailed comments went well beyond the call of duty.

References

ABADI, M. AND LAMPORT, L. 1993. Conjoining specifications. Res. Rep. 118, Digital Equipment Corp., Systems Research Center, Palo Alto, Calif.

ABADI, M. AND LAMPORT, L. 1992. An old-fashioned recipe for real time. Research Report 91, Digital Equipment Corp., Systems Research Center, Palo Alto, Calif.[6]

ABADI, M. AND LAMPORT, L. 1991. The existence of refinement mappings. *Theor. Comp. Sci. 82,* 2 (May), 253–284.

ALPERN, B. AND SCHNEIDER, F. B. 1985. Defining liveness. *Inf. Process. Lett. 21,* 4 (Oct.), 181–185.

APT, K. R. 1981. Ten years of Hoare's logic: A survey—part one. *ACM Trans. Programm. Lang. Syst. 3,* 4 (Oct.), 431–483.

APT, K. R. AND OLDEROG, E.-R. 1990. *Verification of Sequential and Concurrent Programs*. Springer-Verlag, New York.

ASHCROFT, E. A. 1975. Proving assertions about parallel programs. *J. Comput. Syst. Sci. 10*, 110–135.

BURCH, J. R., CLARKE, E. M., MCMILLAN, K. L., DILL, D. L., AND HWANG, L. J. 1992. Symbolic model checking: 10^{20} states and beyond. *Inf. Comput. 98,* 2 (June), 142–170.

CHANDY, K. M. AND MISRA, J. 1988. *Parallel Program Design*. Addison-Wesley, Reading, Mass.

DE BAKKER, J. W., HUIZING, C., DE ROEVER, W. P., AND ROZENBERG, G. (Eds.) 1992. *Real-Time: Theory in Practice*, *Lecture Notes in Computer Science*, vol. 600. Proceedings of a REX Real-Time Workshop, Springer-Verlag, Berlin.

DIJKSTRA, E. W. 1968. The structure of the "THE"-multiprogramming system. *Comm. ACM 11,* 5 (May), 341–346.

DIJKSTRA, E. W. 1976. *A Discipline of Programming*. Prentice-Hall, Englewood Cliffs, N. J.

ENGBERG, U., GRØNNING, P., AND LAMPORT, L. 1992. Mechanical verification of concurrent systems with TLA. In *Computer-Aided Verification*, Lecture Notes in Computer Science, Proceedings of the 4th International Conference, CAV'92. Springer-Verlag, Berlin.

FLON, L. AND SUZUKI, N. 1978. Consistent and complete proof rules for the total correctness of parallel programs. In *Proceedings of 19th Annual Symposium on Foundations of Computer Science*. IEEE, New York, 184–192.

FLOYD, R. W. 1967. Assigning meanings to programs. In *Proceedings of the Symposium on Applied Math. Vol. 19*, American Mathematical Society, Providence, Rhode Island, 19–32.

[6] An earlier version, without proofs, appeared in [de Bakker et al. 1992, pages 1–27].

Garland, S. J. and Guttag, J. V. 1989. An overview of LP, the Larch Prover. In *Proceedings of the 3rd International Conference on Rewriting Techniques and Applications*, N. Dershowitz, Ed. *Lecture Notes on Computer Science*, vol. 355. Springer-Verlag, New York, 137–151.

Hehner, E. C. R. 1984. Predicative programming. *Comm. ACM 27,* 2 (Feb.), 134–151.

Hoare, C. 1969. An axiomatic basis for computer programming. *Comm. ACM 12,* 10 (Oct.), 576–583.

Keller, R. M. 1976. Formal verification of parallel programs. *Comm. ACM 19,* 7 (July), 371–384.

Lam, S. S. and Shankar, A. U. 1984. Protocol verification via projections. *IEEE Trans. Softw. Eng. SE-10,* 4 (July), 325–342.

Lamport, L. 1993. Hybrid systems in TLA$^+$. In *Hybrid Systems*, R. L. Grossman, A. Nerode, A. P. Ravn, and H. Rischel, Eds. *Lecture Notes in Computer Science*, vol. 736. Springer-Verlag, Berlin, 77–102.

Lamport, L. 1983a. Specifying concurrent program modules. *ACM Trans. Program. Lang. Syst 5,* 2 (April), 190–222.

Lamport, L. 1983b. What good is temporal logic. In *Information Processing 83: Proceedings of the IFIP 9th World Congress*, R. E. A. Mason, Ed. North-Holland, Amsterdam, 657–668.

Lamport, L. 1977. Proving the correctness of multiprocess programs. *IEEE Trans. Softw. Eng. SE-3,* 2 (Mar.), 125–143.

Leisenring, A. C. 1969. *Mathematical Logic and Hilbert's ε-Symbol.* Gordon and Breach, New York.

Lipton, R. J. 1975. Reduction: A method of proving properties of parallel programs. *Comm. ACM 18,* 12 (Dec.), 717–721.

Lynch, N. and Tuttle, M. 1987. Hierarchical correctness proofs for distributed algorithms. In *Proceedings of the 6th Symposium on the Principles of Distributed Computing*, ACM, New York, 137–151.

Manna, Z. and Pnueli, A. 1991. *The Temporal Logic of Reactive and Concurrent Systems.* Springer-Verlag, New York.

Misra, J. 1990. Specifying concurrent objects as communicating processes. *Sci. Comput. Program. 14,* 2–3, 159–184.

Owicki, S. 1975. *Axiomatic proof techniques for parallel programs.* Ph. D. thesis, Cornell University, Ithaca N. Y.

Pnueli, A. 1979. The temporal semantics of concurrent programs. In *Semantics of Concurrent Computation*, G. Kahn, Ed. *Lecture Notes in Computer Science*, vol. 70. Springer-Verlag, New York, 1–20.

Pnueli, A. 1977. The temporal logic of programs. In *Proceedings of the 18th Annual Symposium on the Foundations of Computer Science*, IEEE, New York, 46–57.

Schroeder, M. D., Birrell, A. D., Burrows, M., Murray, H., Needham, R. M., Rodeheffer, T. L., Satterthwaite, E. H., and Thacker, C. P. 1990. Autonet: A high-speed, self-configuring local area network using point-to-point links. Res. Rep. 59, Digital Equipment Corporation, Systems Research Center, Palo Alto, Calif.

Shankar, A. U. and Lam, S. S. 1984. Time-dependent communication protocols. In *Principles of Communication and Networking Protocols*, S. S. Lam, Ed. IEEE Computer Society Press, Los Alamitos, Calif., 504–519.

Received November 1992; revised November 1993; accepted December 1993.