9

High-Level Petri
Nets—Extensions,
Analysis, and Applications

Xudong He* and Tadao
Murata

School of Computer Science*
Florida International University,
Florida, USA

Department of Computer Science*
University of Illinois at Chicago,
Chicago, Illinois, USA

9.1 Introduction ................................................................. 000
9.2 High-Level Petri Nets .................................................... 000
9.3 Temporal Predicate Transition Nets .................................... 000
9.4 PZ Nets ........................................................................ 000
9.5 Hierarchical Predicate Transition Nets ............................... 000
9.6 Fuzzy-Timing High-Level Petri Nets ................................ 000

9.1 Introduction

Petri nets are an excellent formal model for studying concurrent and distributed systems and have been widely applied in many different areas of computer science and other disciplines (Murata, 1989). There have been over 8000 publications on Petri nets (refer to Website http://www.daimi.au.dk/PetriNets/). Since Carl Adam Petri originally developed Petri nets in 1962, Petri nets have evolved through four generations: the first-generation low-level Petri nets primarily used for modeling system control (Reisig, 1985a), the second-generation high-level Petri nets for describing both system data and control (Jensen and Bozenburg, 1991), the third-generation hierarchical Petri nets for abstracting system structures (He and Lee, 1991. He, 1996; Jensen, 1992), and the fourth-generation object-oriented Petri nets for supporting modern system development approaches (Agha, 2001). Petri nets have also been extended in many different ways to study specific system properties, such as performance, reliability, and schedulability. Well-known examples of extended Petri nets include timed Petri nets (Wang, 1998) and stochastic Petri nets (Marsan et al., 1994; Haas, 2002). In this article, we present several extensions to Petri nets based on our own research work and provide analysis techniques for these extended Petri net models. We also discuss the intended applications of these extended Petri nets and their potential benefits.

9.2 High-Level Petri Nets

In the past few years, a concerted effort by the worldwide Petri net community has resulted in an international standard on defining high-level Petri nets (HLPN) (ISOLTEC, 2002) that will profoundly help to facilitate and promote the research and applications of these nets. In the following sections, we follow as closely as possible the notations, concepts, and definitions given in the standard documentation to introduce high-level Petri nets, which are used later to define our extensions.
9.2.1 The Syntax and Static Semantics of High-Level Petri Nets

An HLPN is a structure:

\[ N = (NG, Sig, V, H, \text{Type}, AN, M_0). \]

Here:

- \( NG = (P, T; F) \) is a net graph, with:
  - \( P \) a finite set of nodes, called places;
  - \( T \) a finite set of nodes, called transitions, disjoint from \( P(P \cap T = \emptyset) \);
  - \( F \subseteq (P \times T) \cup (T \times P) \) a set of directed edges called arcs, known as the flow relation.
- \( Sig = (S, O) \) is a Boolean signature, where \( S \) is a set of sorts and \( O \) is a set of operators defined in the Annex A of ISO/IEC (2002).
- \( V \) is an \( S \)-indexed set of variables, disjoint from \( O \).
- \( H = (S_I, O_H) \) is a many-sorted algebra for the signature \( Sig \), defined in this list:
  - \( \text{Type} : P \to S_H \) is a function that assigns types to places.
  - \( \text{AN} = (A, TC) \) is a pair of net annotations.
- \( A : F \to \text{TERM}(O \cup V) \) such that for all \( (p, t), (t, p) \in F \) and all bindings \( \alpha, \text{Val}_\alpha(A(p, t)), \text{Val}_\alpha(A(t, p)) \in \mu \text{Type}(p) \).
  - \( \text{TERM}(O \cup V), \alpha, \text{Val}_\alpha \) and \( \mu \text{Type}(p) \) are defined in Annex A of ISO/IEC (2002).
  - \( A \) is a function that annotates each arc with a term that when evaluated (for some binding) results in a multiset over the associated place's type.
- \( TC : T \to \text{TERM}(O \cup V)_{\text{bool}} \) is a function that annotates transitions with Boolean expressions.
- \( M_0 : P \to \bigcup_{p \in \mu \text{Type}(p)} \) such that for all \( p \in P \), \( M_0(p) \in \mu \text{Type}(p) \) is the initial marking function that associates a multiset of tokens (of the correct type) with each place.

The above definitions are directly from ISO/IEC 2002. Basically, HLPN has three essential parts: (1) a net graph \( NG \) defining the syntax, (2) an underlying algebraic specification \( (Sig, V, H) \) defining the semantic domain, and (3) a net inscription \( (Type, AN, M_0) \) mapping syntactic entities to their semantic denotations. By restricting the underlying algebraic specification and/or the net inscription, different variations of high-level Petri nets can be obtained.

9.2.2 Dynamic Semantics

**Marking**

The marking \( M \) of the HLPN is defined in the same way as the initial marking:

\[ M : P \to \bigcup_{p \in \mu \text{Type}(p)} \] such that \( \forall p \in P \), \( M(p) \in \mu \text{Type}(p) \).

**Enabling**

- **Enabling of a Single Transition Mode**
  A transition \( t \in T \) is enabled in a marking \( M \) for a particular assignment of \( x_t \) to its variables and satisfies the transition condition, \( \text{Val}_\alpha(\text{TC}(t)) = \text{true} \) known as a mode of \( t \), if:

\[ \forall p \in P \text{Val}_\alpha(p_T) \leq M(p), \]

where \( \forall (u, v) \in (P \times T) \cup (T \times P), u \not= v \) for \( (u, v) \not\in F \).

- **Concurrent Enabling of Transition Modes**
  Let \( x_t \) be an assignment for the variables of transition \( t \in T \) that satisfies its transition condition, and then denote the set of all assignments for transition \( t \) by \( \text{Assign}_t \). Define the set of all transition modes to be \( TM = \{ (t, x_t) | t \in T, x_t \in \text{Assign}_t \} \) and a step to be a finite nonempty multiset over \( TM \).

A step \( X \) of transition modes is enabled in a marking, \( M \), if:

\[ \forall p \in P \sum_{(t, x_t) \in X} \text{Val}_\alpha(p_T) \leq M(p). \]

Thus, all of the transition modes in \( X \) are concurrently enabled if \( X \) is enabled. Enabling of a single transition mode is a special case of concurrent enabling of transition modes.

The enabling condition of a transition specifies that enough right tokens are available to make the transition happen.

**Transition Rule**

If \( t \in T \) is enabled in mode \( x_t \) for marking \( M \), \( t \) may occur in mode \( x_t \). When \( t \) occurs in mode \( x_t \), the marking of the net is transformed to a new marking \( M' \), denoted in \( M[t, x_t > M'] \), according to the following rule:

\[ \forall p \in P(M'(p) = M(p) - \text{Val}_\alpha(p_T) + \text{Val}_\alpha(T, p)). \]

If a step \( X \) is enabled in marking \( M \), then the step can occur, resulting in a new marking of \( M' \) denoted by \( M[X > M'] \), where \( M' \) is given by:

\[ \forall p \in P(M'(p) = M(p) - \sum_{(t, x_t) \in X} \text{Val}_\alpha(p_T) + \sum_{(t, x_t) \in X} \text{Val}_\alpha(T, p)). \]

The firing rule defines the effect of firing a transition, which consumes specific tokens according to the mode in the preset (input places) of the transition and generates new tokens in the postset (output places) of the transition.

**Behavior of A HLPN \( N \)**

An execution sequence \( M_0 | X_0 > M_1 | X_1 > \ldots \) of \( N \) is either finite when the last marking is terminal (no more enabled transition in the last marking) or infinite, in which each \( X_i \) is
9.3 Temporal Predicate Transition Nets

Petri nets are a model-oriented formal method and are well suited for modeling the dynamic behaviors of concurrent and distributed systems. Petri net specifications reveal system design structures and thus provide guidelines for system implementation. Furthermore, Petri net specifications are executable and support system simulation and testing in addition to formal analysis. It is not easy, however, to directly and explicitly express behavioral properties using Petri nets. On the other hand, property-oriented formal methods, such as temporal logic, are ideal for specifying and analyzing behavioral properties. It would be great to have a hybrid formal method by integrating the strengths of model-oriented and property-oriented formal methods. In recent years, integrating formal methods has become a major research area (Clark and Wing, 1996). There have been several published results on integrating Petri nets and temporal logic (Anttila et al., 1983; Diaz et al., 1983; He and Lee, 1990; He, 1992; Mandrioli et al., 1996; Suzuki and Lu, 1989). Most of the earlier work was either lacking a systematic approach or using a low-level Petri net model. In the following subsections, we describe our results on integrating predicate transition nets [GL81] and first-order linear time temporal logic [MP92].

9.3.1 Definition of Temporal Logic

In this section, we define a linear time first-order temporal logic (LITFOTL) in the style of Lamport (1994a) based on a given PrTN. Let \( N = (NG, Sig, V, H, Type, AN, M_0) \) be a PrTN.

Values, States, Variables, and Predicates

The set of values is the multiset of tokens defined by the ground terms \( TERM(O) \) of \( N \). Multisets can be viewed as partial functions. For example, multiset \( \{3a, 2b\} \) can be represented as \( \{a=3, b=2\} \).

The set of state variables is the set \( P \) of places of \( NG \), which change their meanings during the executions of \( N \). The arity of a place \( p \) is determined by its type \( Type(p) \).

The set of states \( St \) is the set of all reachable markings \( M_0 \geq M \) of \( N \). A marking is a mapping from the set of state variables into the set of values. We use \( M[p] \) to denote the value of \( p \) under state (marking) \( M \).

Since state variables take partial functions as values, they are flexible function symbols. We can access a particular component value of a state variable. There is a problem, however, associated with partial functions (i.e., many values are undefined). The above problem can easily be solved by extending state variables into total functions in the following way: for any \( n \)-ary state variable \( p \), any tuple \( c \in TERM(O) \), and any state \( M \), if \( p(c) \) is undefined under \( M \), then let \( M[p(c)] = 0 \). The above extension is consistent with the semantics of PrTN. Furthermore we can consider the meaning \( [p(c)] \) of the function application \( p(c) \) as a mapping from states to \( \text{Nat} \) using a postfix notation for function application \( M[p(c)] \).

Rigid Variables, Rigid Function, and Predicate Symbols

Rigid variables are individual variables that do not change their meanings during the executions of \( N \). All rigid variables occurring in our temporal logic formulas are bound (quantified), and they are the only variables that can be quantified. Rigid variables are variables appearing in the label expressions and constraints of \( N \).

Rigid function and predicate symbols do not change their meanings during the executions of \( N \). The set of rigid function and predicate symbols is defined in \( V \) of \( N \).

State Functions, Predicates, and Transitions

A state function is an expression built from values, state variables, rigid function, and predicate symbols. For example, \( [p(c) + 1] \) is a state function, where \( c \) and \( 1 \) are values, \( p \) is a state variable, and \( + \) is a rigid function symbol. Since the meanings of rigid symbols are not affected by any state, for any given state \( M \), \( M[p(c) + 1] = M[p(c)] + 1 \).

A predicate is a Boolean-valued state function. A predicate \( p \) is satisfied by a state \( M \) if and only if \( M[p] \) is true.

A transition is a particular kind of predicate that contains primed state variables (e.g., \( [p'(c) = p(c) + 1] \)). A transition relates two states (an old state and a new state), where the unprimed state variables refer to the old state and where the primed state variables refer to the new state. Therefore, the meaning of a transition is a relation between states. The term transition used here is a temporal logic entity. Although it reflects the nature of a transition in a PrTN net \( N \), it is not a transition in \( N \). For example, given a pair of states \( M \) and \( M' \), \( M[p'(c) = p(c) + 1]M' \) is defined by \( M'[p(c)] = M[p(c)] + 1 \). Given a transition \( t \), a pair of states \( M \) and \( M' \) is called a “transition step” if \( M[p]M' \) equals true.

We can easily generalize any predicate \( p \) without primed state variables into a relation between states by replacing all unprimed state variables with their primed versions such that \( M'[p]M' \) equals \( M'[p] \) for any states \( M \) and \( M' \).

Temporal Formulas

Temporal formulas are built from elementary formulas (predicates and transitions) using logical connectives \( \neg \) and \( \land \) (and
The semantics of temporal logic is defined on infinite sequences of states that are extracted from the execution sequences of PrTN, where the last marking of a finite execution sequence is repeated infinitely many times at the end of the execution sequence. For example, for an execution sequence \(M_0 \mid X_0 > M_1 \mid X_1 \ldots > M_n\), the following infinite sequence behavior \(\sigma = << M_0 \ldots M_n, M_n, \ldots >>\) is obtained. We denote the set of all possible behaviors obtained from a given PrTN as \(St^\infty\). Let \(u\) and \(v\) be two arbitrary temporal formulas; \(p\) be an m-ary place; \(t\) be a transition; \(x_1, x_2, \ldots, x_n\) be rigid variables; \(\sigma = << M_0, M_1, \ldots >>\) be a behavior; and \(\sigma^k = << M_k, M_{k+1}, \ldots >>\) be a \(k\)-step-shifted behavior sequence. We define the semantics of temporal formulas recursively as follows:

\[
\begin{align*}
(1) \; & \sigma[p(x_1, \ldots, x_n)] \equiv M_0[p(x_1, \ldots, x_n)] \\
(2) \; & \sigma[t] \equiv M_0[t] \circ M_1 \\
(3) \; & \neg \sigma[u] \equiv \neg \sigma[u] \\
(4) \; & \sigma[u \land v] \equiv \sigma[u] \land \sigma[v] \\
(5) \; & \forall x \sigma[u] \equiv \forall x \sigma[u] \\
(6) \; & \sigma[u] \equiv \forall u \in \text{Nat} \; \sigma[u]
\end{align*}
\]

Intuitively temporal operator always \(\square\) means every state in a state sequence, and its dual operator sometimes \(\lozenge\) means some future state in a state sequence. The relationship between these two temporal operators is:

\[-\square \equiv \lozenge.\]

A temporal formula \(u\) is said to be satisfiable denoted as \(\sigma[u] = u\), if there is an execution \(\sigma\) such that \(\sigma[u]\) is true (i.e., \(\sigma = u \Leftrightarrow \exists \sigma \in St^\infty. \sigma[u]\)). A temporal formula \(u\) is valid with regard to \(N\), denoted as \(N = u\), if it is satisfied by all possible behaviors \(St^\infty\) from \(N\): \(N = u \Leftrightarrow \forall \sigma \in St^\infty. \sigma[u]\).

9.3.2 Temporal Predicate Transition Net

A temporal predicate transition net (TPrTN) is a tuple \(TN = (NG, \text{Sig}, V, H, \text{Type}, AN, M_0, f)\), where \(N = (NG, \text{Sig}, V, H, \text{Type}, AN, M_0)\) is a PrTN, and \(f\) is a LTFTOTL formula that constrains the execution of \(TN\). The semantics of \(TN\) are defined to be the set of execution sequences \(\Sigma = \{\sigma|\sigma \in St^\infty \land \sigma[f]\}\).

It is easy to see that a TPrTN \(TN\) is a PrTN \(N\) with a temporal logic formula \(f\). The temporal logic formula is defined according to the net graph \(NG\) using logical connectives and operators, which can be viewed as a part of the underlying algebraic specification (Sig, V, H). The temporal logic formula \(f\) is evaluated using the dynamic behaviors of \(N\). By incorporating a temporal formula into the definition of a PrTN, we are able to explicitly specify and verify many system properties such as fairness.

9.3.3 An Example of TPrTN

The Five Dining Philosophers problem is a classical example for studying behavioral properties of concurrent processes. There are five philosophers sitting around a round table, and there is one chopstick between two adjacent philosophers. Each philosopher is either thinking or eating. In order for a philosopher to eat, he needs to pick up two chopsticks next to him. There are several interesting issues with regard to a system model of this simple problem:

- Is the system deadlock free? A deadlock occurs when the system cannot progress anymore due to a formation of a waiting cycle.
- Has the system enforced mutual exclusion? Mutual exclusion ensures that no two neighboring philosophers can eat at the same time.
- Is the system live-lock free? A live lock occurs when some philosopher has no chance to eat anymore from a certain point and thus starves to death.

The following TPrTN models the Five Dining Philosophers problem, shown in Figure 9.1, in which five philosophers are denoted by five integer tokens 0 to 4 and five chopsticks are also denoted by five integer tokens 0 to 4. Places \(p_1\) and \(p_2\) stand for the Thinking and Eating states of philosophers, respectively. Place \(p_2\) defines the availability of chopsticks. Transitions \(t_1\) and \(t_2\) stand for the actions of Pick up and Put down chopsticks, respectively.

The algebraic definition of the net \(TN = (NG, \text{Sig}, V, H, \text{Type}, AN, M_0, f)\) is as follows:

\[
P = \{p_1, p_2, p_3\}, \; T = \{t_1, t_2\},
\]

\[
F = \{(p_1, t_1), (p_2, t_1), (t_1, p_3), (p_3, t_1), (t_2, p_1), (t_2, p_1), (t_2, p_2)\},
\]

\[
\text{Type}(p_1) = \text{Type}(p_2) = \text{Type}(p_3) = \text{INT},
\]

\[
A(p_1, t_1) = x, \; A(p_2, t_1) = [x, y], \; A(t_1, p_3) = x, \; A(p_3, t_2) = x, \; A(t_2, p_1) = x, \; A(t_2, p_2) = [x, y],
\]

\[
TC(t_1) = TC(t_2) = y = x \oplus 1,
\]

\[
M_0(p_1) = M_0(p_2) = \{0, 1, 2, 3, 4\}, \; \text{and} \; M_0(p_3) = \{\}.
\]
where INT is a sort defined in Sig, x and y are variables in \( V \), + is a modulo 5 addition operation given in Sig and defined in the algebra \( H \), and \( y = x + 1 \) is a term in algebra \( H \).

Temporal formula \( f \) is defined as follows:

\[
f = \forall x (\square (p_1(x) = 1 \land p_2(x) = 1 \land p_3(x+1) = 1) \Rightarrow \square (p'_1(x) = 1 \land p'_2(x) = 0 \land p'_3(x+1) = 0)).
\]

Temporal formula \( f \) defines the strong fairness (Manna and Pnueli, 1992) that captures the enabling condition and the firing result of transition \( t \) with every mode (defined by \( \forall x \)). Intuitively, it states that any philosopher who wants to eat infinitely many times (defined by the temporal operator sequence \( \exists \omega \)) will eat infinitely many times.

### 9.3.4 Analysis of TPrTNs

System behavioral properties can be divided into two major categories: safety properties and liveness properties (Manna and Pnueli, 1992). Widely accepted formal definitions of safety and liveness properties were given in [AS85 (Alpern and Schneider, 1985)]. A safety property, which is different from the concept sainess in Petri net literature (Murata, 1999) needs to hold in every state and in every state sequence. The canonical form of a safety property is characterized by a temporal formula \( \square \omega \). A liveness property, which is different from the concept transition liveness in Petri net literature (Murata, 1989), needs to hold in some future state in every state sequence. The canonical form of a liveness property is characterized by a temporal formula \( \exists \omega \).

In He and Lee (1990) and He and Ding (1992), we developed an axiomatic approach of using temporal logic to analyze PrTNs. The essential ideas are to derive system-dependent temporal inference rules from PrTN transitions. These inference rules capture the causal relationships of transition firings. Thus, executions of a PrTN become temporal logic deductions in the derived axiomatic system. We have proved that safety properties (Manna and Pnueli, 1992) can be effectively analyzed using this axiomatic approach. Furthermore, we have shown how to analyze liveness properties (Manna and Pnueli, 1992) using the above temporal formula by the (1991).

Given \( TN = (NG, \mathit{Sig}, V, H, \mathit{Type}, AN, M_0, f) \), we can obtain the following system-dependent temporal axiom system \( L \):

1. A system-dependent axiom captures the initial marking \( M_0 \): \( \land_{p \in P} (IC_p) \), \( (IC) \)
2. or each place \( p \) of arity \( k \), let \( y_1, \ldots, y_k \) be new variables not used in any label and \( t \) be any transition in the postset \( \bullet p \) with label \( L(p, t) = \{ m_1(x_{11}, \ldots, x_{1k}), \ldots, m_k(x_{n1}, \ldots, x_{nk}) \} \). We construct an inference rule as follows: \( p(y_1, \ldots, y_k) > p'(y_1, \ldots, y_k) \)

where \( ET_t \) represents the causal relationship of firing transition \( t \). Intuitively, each inference rule states if a place loses tokens from one marking to another, then some relevant transition must have fired (He and Ding, 1992). The variables are universally quantified and are omitted for simplicity. The above inference rules only reflect the local state changes. Together with system independent axioms and inference rules (Manna and Pnueli, 1992, 1995), we can use IC and PRp to prove a variety of safety properties.

The following is the derived system-dependent temporal axiom system of the Five Dining Philosophers problem shown in Figure 9.1:

\[
p_0(0) = 1 \land p_1(1) = 1 \land p_2(2) = 1 \land p_3(3) = 1 \land p_4(4) = 1 \land p_5(5) = 1 \land p_0(1) = 1 \land p_1(2) = 1 \land p_2(3) = 1 \land p_3(4) = 1 \land p_4(5) = 1 \land p_5(6) = 1. \quad (IP)
\]

\[
p_0(x) > p'_1(x) - p_1(x) \geq 1 \land p_2(x) \geq 1 \land p_3(y) \geq 1 \land y = x + 1 \land p_1(x) = p_0(x) - 1 \land p_2(x) = p_1(x) + 1. \quad (PR_{p_1})
\]

\[
p_0(x) > p'_2(x) - p_2(x) \geq 1 \land p_3(x) \geq 1 \land p_4(y) \geq 1 \land y = x + 1 \land p_2(x) = p_0(x) - 1 \land p_3(x) = p_2(x) + 1 \land p_4(y) = p_3(y) + 1. \quad (PR_{p_2})
\]

A TPrTN is deadlock free if and only if there is at least one transition being enabled in any reachable state or if a normal terminal state has been reached. Deadlock freedom is a safety property. The deadlock freedom property of the Five Dining Philosophers problem is formulated as follows:

\[
\square \exists x, y(p_1(x) \geq 1 \land y = x + 1). \\
(\forall)
\]

A TPrTN net is livelock free if and only if for any transition enabled infinitely often with a mode \( \alpha \), the same transition with mode \( \alpha \) will fire infinitely often. Livelock freedom is a liveness property. One of the liveliness properties of the Five Dining Philosophers problem is that every individual philosopher will eventually get a piece of meal (starvation free), which can be expressed by the following temporal formula:

\[
\forall x \square \omega (p_2(x) = 1). \\
(**)
\]

The proofs of the above formulas involve logical deductions using inference rules of ordinary first-order logic and temporal logic as well as system-dependent inference rules listed previously. Because logical deductions are machine-oriented and not suitable for human comprehension, we omit the proofs of the
above properties (*) and (**). Interested readers can find the proofs of (*) and (***) in He (1991) and He and Ding (1992).

9.4 PZ Nets

A widely accepted formal notation for specifying the functionality of sequential systems is Z (Spivey, 1992). The Z is based on typed set theory and first-order logic and, thus, offers rich type definition facility and supports formal reasoning. However, Z does not support an effective definition of concurrent and distributed systems, and Z specifications do not have explicit operational semantics. Many researchers have attempted to combine Z with other concurrent models, including CSP (Benjamin, 1990), CCS (Taguchi and Araki, 1997), and temporal logic (Duke and Smith, 1989; Clarke and Wing, 1996; Evans, 1997; Lamport, 1994b) in recent years. Several works attempted to integrate Petri nets and Z (Van Hee et al., 1991; He, 2001). In van Hee et al. (1991); Z was used to specify (1) the metamodel of restricted hierarchical colored Petri nets (that allows super transitions but not super places) and (2) the transitions of specific colored Petri nets. A complete specification consists of a hierarchical net structure, one global state schema for defining all places, and one operation schema for each transition. Schemas of transitions are piped to obtain an operational semantics through the use of input and output variables. In He (2001), a formal method of integrating Petri nets with Z was presented with the objectives to extend Petri nets with data and function definition capabilities through an underlying Z specification and to extend Z with an explicit operational semantics and concurrency mechanisms through Petri nets. In the following subsections, we briefly describe the basics of Z notation and introduce the results given in He (2001).

9.4.1 A Brief Overview of Z

Besides elementary data types, Z allows the user to introduce other primitive types, called given types. Given types are in capital letters and enclosed by a pair of brackets that do not require further definitions. For example, [IDEN] introduces a given type IDEN.

Part of Z is an essential notation called schema from which new types and their properties can be defined. A Z schema has a boxed structure often consisting of two major parts: the declaration part and the predicate part (optional) as follows:

Name
Declaration part
Predicate part

The name of a schema is global and can be used in the declaration part of other schemas to reuse the variable definitions.

The declaration part defines local variables of the schema in the form: var: Type (the collection of these variable definitions forms the signature of the schema), and thus a reference of this variable in another schema needs to be prefixed by its schema name followed by a period (i.e., Name. var: A variable name can be a simple name (defining the old value), a name with a dash (defining the new value of the same name without the dash), a name with a question mark? (an input from the external environment), or a name with an exclamation point! (an output to the external environment). If a schema includes both the name of a schema S and its dashed version S', the following notation is used: ΔS when the value of some variable defined in S has been changed by the required processing or ΕS when nothing in S is changed by the required processing.

The predicate part specifies the constraint (invariant) of a definition or the precondition (subformulas without dashed variables) and postcondition (subformulas containing dashed variables) of some processing in terms of first-order logic formulas. Subformulas on separate lines in the predicate part are conjoined by default. The predicate thus defines the property of the schema.

State Schemas and Operation Schemas

A schema can be used to define the abstract state of a system, called a state schema in the sequel, when the state machine model is used. The predicate part in this case defines the data invariant among involved variables.

Each state schema S needs an initial value that is defined by an initialization schema that enumerates the initial values of state variables in the predicate part.

A schema SC can be used to define an operation, called an operation: schema in the sequel, which includes both a state schema name S and its version S' in its declaration part. However, ΔS or ΕS is often used instead of S and S'. The predicate part can be divided into the pre-condition part (denoted by pre-SC in the sequel) and the post-condition part (denoted by post-SC in the sequel).

Operations on Schemas

New schemas can be built using existing schemas besides the inclusion mentioned in the previous paragraphs.

1. **Schema disjunction:** Schema disjunction has the form New = Old1 ∨ Old2, where New, Old1, and Old2 are schema names. The declaration part of new schema New is obtained by merging the declaration parts of Old1 and Old2, and the predicate part of New is obtained by making a disjunction of the predicate part of Old1 with that of Old2.

2. **Schema conjunction:** Schema conjunction has the form New = Old1 ∧ Old2, where New, Old1, and Old2 are schema names. The declaration part of new schema New is obtained by merging the declaration parts of Old1 and Old2, and the predicate part of New is obtained by making a conjunction of the predicate part of Old1 with that of Old2.
(3) **Schema composition:** Let Old1 and Old2 be two schemas; a new schema of the form New = Old1; Old2 can be defined and has the following meaning:
- The signature of New is the inputs and outputs of both Old1 and Old2, together with the nondashed variables in Old1 and dashed variables of Old2;
- The property of Old1 is included in New, but all the dashed names are redecorated with a decorator not used in either Old1 and Old2;
- The property of Old2 is included in New, but all the nondashed names are decorated with the same decorator as was used in Old1; and
- The newly decorated names are hidden with an existential quantifier.

Other operations related to schemas can be found in work by Spivey (1992), including using a schema as a type in the declaration part of other schemas, hiding declarations in a schema, and reusing the operations defined in one schema in other related schemas through promotion.

### 9.4.2 Definition of PZ Nets

Let $Z = (Z_P, Z_I, Z_T)$ be a collection of Z schemas. If we let $z$ be a Z schema, we use `name(z)`, `sig(z)`, and `prop(z)` to denote the name, the signature part (as a typed set of mappings), and property part of $z$, respectively in the sequel. The $Z_I$ schemas in $Z$ satisfy the following conditions:

- $\forall z_1, z_2 \in Z_P \cdot (z_1 \neq z_2 \Rightarrow \text{sig}(z_1) \cap \text{sig}(z_2) = \emptyset)$;
- $\forall z_1, z_2 \in Z_I \cdot (z_1 \neq z_2 \Rightarrow \text{sig}(z_1) \cap \text{sig}(z_2) = \emptyset)$; and
- $|Z_P| = |Z_I|$.

The first two conditions state that the signatures of $z$ schemas are pair-wise disjoint, and the last condition states that the number of $z$ schemas in $Z_P$ is the same as that in $Z$ (i.e., one-to-one correspondence).

A **PZ net** is a tuple $ZN = (NG, Sig, V, H, Type, AN, M_0, Z)$.

Here, $Z$ is a collection of Z schemas satisfying the above conditions, $AN = (Type, A, TC)$, and:

- **Type:** $P \to Z_P$ is one-to-one mapping, giving the data definition of $ZN$. For each place $p \in P$, Type maps $p$ to a unique Z schema $z \in Z_P$ such that $p = \text{name}(z)$. The type of $p$ is defined by the signature of $z$.
- **TC:** $T \to Z_T$ is one-to-one mapping, providing the functionality definition of $ZN$. For each transition $t \in T$, $C$ maps $t$ to a Z schema $z \in Z_T$ such that $t = \text{name}(z)$. The functional processing of $t$ is defined by the property of $z$. Furthermore, the following constraint is satisfied. For any $p \in P, t \in T$:
  1. If $(p, t) \in F$, then $\text{sig}(\text{Type}(p)) \cap \text{sig}(\text{TC}(t)) \neq \emptyset$.
  2. If $(t, p) \in F$, then $\text{sig}(\text{Type}(p)) \cap \text{sig}(\text{TC}(t)) \neq \emptyset$.
- **A:** $F \to pVar$ is the control flow definition of $ZN$, where $pVar$ is the set of all hidden variables (through quantification) in $Z_T$. Let $Var(t)$ denote the set of hidden variables in $TC(t)$ for any transition $t$, let $S(v)$ denote the sort (or type) of a variable $v$, and let $S(V)$ denote the set of the sorts of the variables in the sorted set $V$. Then, the following constraints are satisfied for any $p \in P$:
  1. $\tilde{A}(p, t) \subseteq Var(t)$ and $\tilde{A}(t, p) \subseteq Var(t)$, where
    $$\tilde{A}(x, y) = \begin{cases} A(x, y) & \text{if } (x, y) \in F \\ \emptyset & \text{otherwise} \end{cases}$$
  2. For any $x \in \tilde{A}(p, t) (\text{or } \tilde{A}(t, p)) \cdot S(x) \subseteq S(\text{sig}(\text{Type}(p)))$.
  3. $\text{prop}(\text{TC}(t)) \Rightarrow \text{sig}(\text{Type}(p)) = \text{sig}(\text{Type}(p)) - \tilde{A}(p, t) \cup \tilde{A}(t, p)$.
- $M_0: P \rightarrow Z_0$ is a Type-respecting (i.e., $\text{sig}(\text{Type}(p)) = \text{sig}(M_0(p))$) initial marking of $ZN$, where $Z_0$ is a set of Z schemas defining the initial state of the system. $M_0(p)$ is the Z schema in $Z_I$ defining the initial marking of place $p$.

The following is a simple example of a PZ net specification of the familiar Five Dining Philosophers problem. The overall system structure is shown in Figure 9.2, which is the same as Figure 9.1 except for the renaming.

The Z schemas ($Z_P, Z_I, Z_T$) are the following:

1. $Z_P = \{\text{thinking}, \text{chopstick}, \text{eating}\}$:
   - **Thinking**
     - `tphil`: `PHIL`
   - **Chopstick**
     - `chop`: `CHOP`
   - **Eating**
     - `ephil`: `(PHIL × CHOP × CHOP)`
     - left: `PHIL` → `CHOP`
     - right: `PHIL` → `CHOP`

2. $Z_T = \{\text{pickup}, \text{putdown}\}$:
   - **Pick up**
     - $\Delta$ Thinking
   - **Chopstick**

![FIGURE 9.2](image-url) A PZ Net Specification of the Five Dining Philosophers Problem.
$\Delta$ Eating
\[ \exists x: \text{PHIL} \bullet \exists y_1, y_2: \text{CHOP} \bullet 3z: \text{PHIL} \times \text{CHOP} \times \text{CHOP} \bullet (x \in \text{tphl} \land y_1 \in \text{chop} \land y_2 \in \text{chop} \land y_1 = \text{left}(x) \land y_2 = \text{right}(x) \land (z' = \langle x, y_1, y_2 \rangle \land \text{tphl}' = \text{tphl} \setminus \{x\} \land \text{chop}' = \text{chop} \setminus \{y_1, y_2\} \land \text{ephil}' = \text{ephil} \cup \{z'\} \land \text{left}' = \text{left} \land \text{right} = \text{right}) \]

Put down

$\Delta$ Thinking
$\Delta$ Chopstick
$\Delta$ Eating
\[ \exists x': \text{PHIL} \bullet \exists y_1', y_2': \text{CHOP} \bullet 3z: \text{PHIL} \times \text{CHOP} \times \text{CHOP} \bullet (z \in \text{ephil} \land z = \langle x, y_1', y_2' \rangle \land \text{tphl}' = \text{tphl} \cup \{x'\} \land \text{chop}' = \text{chop} \cup \{y_1', y_2'\} \land \text{ephil}' = \text{ephil} \setminus \{z\} \land \text{left}' = \text{left} \land \text{right} = \text{right}) \]

(3) $Z_i = \text{Init\_Thinking}, \text{Init\_Chopstick}, \text{Init\_Eating}$:

\begin{itemize}
  \item Init\_Thinking
    \[ \text{tphl} = \{0, 1, 2, 3, 4\} \]
  \item Init\_Chopstick
    \[ \text{chop} = \{0, 1, 2, 3, 4\} \]
  \item Init\_Eating
    \[ \text{Eating} \]
    \[ \text{ephil} = \emptyset \]
    \[ \text{left} = \{0, 0, 2, 2, 3, 3\} \]
    \[ \text{right} = \{0, 0, 1, 1, 2, 3, 3\} \]
  \end{itemize}

Type = \{p_1 \rightarrow \text{Thinking}, p_2 \rightarrow \text{Chopstick}, p_3 \rightarrow \text{Eating}\}

\[ \text{TC} = \{t_1 \rightarrow \text{Pickup}, t_2 \rightarrow \text{Putdown}\} \]

$A \equiv \{(p_1, t_1) \rightarrow \{x\}, (p_2, t_1) \rightarrow \{y_1, y_2\}, (t_1, p_3) \rightarrow \{z\}, (p_3, t_2) \rightarrow \{z\}, (t_2, p_1) \rightarrow \{x'\}, (t_2, p_2) \rightarrow \{y_1', y_2'\}\}$

M_0 = \{p_1 \rightarrow \text{Init\_Thinking}, p_2 \rightarrow \text{Init\_Chopstick}, p_3 \rightarrow \text{Init\_Eating}\}.

\section{9.4.3 PZ Net Analysis}

In He (1995, 2001), a structural induction technique was proposed for analyzing safety properties of PZ nets, which is based on an invariance inference rule from work by Manna and Pruei (1995). This technique is briefly discussed in the following subsections.

**Formalizing Invariant Properties**

Let $|M_0|^\omega$ denote the set of all valid marking sequences extracted from all execution sequences of a given PZ net, and let $\sigma$ be a valid marking sequence with $|\sigma|$ as the length of the sequence and $\sigma(i)$ as the $i$th state (marking) in $\sigma$. W (a first order logic formula) is an invariant property if and only if the following holds:

\[ \forall \sigma: \sigma(i) = |W| \land \sigma(i) = |\sigma| \land W(\sigma(i)) \Rightarrow W(\sigma(i)) \]

Thus, a safety property holds in every state (marking) of every valid marking sequence. The above formulation can be simplified to the following equivalent version in terms of the set of all reachable markings only:

\[ \forall M: M \in |M_0| \Rightarrow (M \models W) \]

**Proving Invariant Properties**

In Manna and Pruei (1995), several temporal logic-based inference rules for invariance (or safety properties) were given. Among the rules, the following basic invariance rule (in its state validity form) is reformulated in terms of PZ nets and is essential.

**The Basic Invariance Rule**

B1. $M_0 \Rightarrow W$

B2. $C(t): a \land W \Rightarrow W'$ for every $t \in T$ and occurrence $x$

In the basic invariance rule, premise B1 requires that the initial marking $M_0$ imply property $W$, and premise B2 requires that all transitions preserve $W$. $C(t)$ is the Z schema associated with $t$ that defines the enabling condition (precondition) and the firing result (postcondition) of $t$. $W'$ is obtained from $W$ by changing the names of variables to their dashed version. Based on premises B1 and B2, we conclude that $W$ is valid or is satisfied under any reachable marking from $M_0$.

We provide the following procedure to use the above inference rule:

**Step 1.** Prove the initial marking as satisfying a system property formula.

**Step 2.** Assume the system property formula holds after $k$ events in a state $M$.

**Step 3.** Prove the system property formula holds after $k + 1$ events in any directly reachable state $M'$ from $M$.

It is easy to see that steps 2 and 3 in the temporal induction proof technique fulfill premise B2 of the invariance rule. Furthermore, we only need to consider the firing of a relevant transition and the associated $Z$ schema with regard to the given property under the guidance of the net structure during step 3.
For example, a system deadlock may occur when a particular place has a special marking. To show the system does not have the deadlock, we only need to show that transitions connected to this place cannot result in this special marking. Therefore, the proof is in general local in the sense that only a subset of transitions needs to be considered. Similar ideas have been also explored in other temporal analysis techniques (He and Lee, 1990; He and Ding, 1988). In general, logical, net structural, and net behavioral reasonings are needed to prove an invariant property. The above temporal induction technique is demonstrated in the following example.

The deadlock freedom property of the Five Dining Philosophers problem is formulated as follows:

$$\Box (\exists x \in \text{phil}(p_3(x) \geq 1 \lor \exists x \in \text{tphil}.(\text{right}(x) \in p_2 \land \text{left}(x) \in p_2)).$$

This problem explains that at any state, a philosopher is eating, which ensures the enabledness of transition Put down, or a philosopher is thinking, which ensures the enabledness of transition Pick up.

The above formula can be rewritten as follows without using the temporal operator:

$$\forall M \in [M_0, \ldots, \exists x \in \text{phil}(p_3(x) \geq 1 \lor \exists x \in \text{tphil}.(\text{right}(x) \in p_2 \land \text{left}(x) \in p_2)).$$

Thus, we need to prove the following formula using the structural induction technique:

$$\exists x \in \text{tphil}(p_3(x) \geq 1 \lor \exists x \in \text{phil} \text{ right}(x) \in p_2 \land \text{left}(x) \in p_2) \quad (*)$$

Here is the proof outline of the formula (*):

**Step 1** Under the initial marking $M_0$, Init_Thinking, Init_Chopstick, and Init_Eating endures:

$$\exists x \in \text{phil}.(\text{right}(x) \in p_2 \land \text{left}(x) \in p_2)$$

**Step 2** Assume (*) holds after $k$ transitions in a state $M$.

**Step 3** Prove (*) holds after $k + 1$ transitions in a state $M'$ such that $M[t/\alpha > M']$.

**Case 1** Firing transition Pick up with $\alpha = [x/p_{h1}, y_1/c_{h1}, y_2/c_{b2}]$ and from the postcondition of schema Pick up with $c_{h1} \in \text{phil}' \text{ in } M'$. Thus, $\exists x \in \text{phil}(p_3(x) \geq 1)$ is true in $M'$, and hence (*) holds in $M'$.

**Case 2** Firing transition putdown $\alpha = [z/\in p_{h1}, c_{h1}, c_{b2}]$:

From the precondition of put down with left ($p_{h1} = c_{h1}$) and right ($p_{h1} = c_{b2}$) from the postcondition of schema put down with $p_{h1} \in tphil$, $c_{h1} \in \text{chop}'$, and $c_{b2} \in \text{chop}'$. Thus, $\exists x \in \text{tphil}.\text{right}(x) \in p_2 \land \text{left}(x) \in p_2$ is true in $M'$, and hence (*) holds in $M'$.

Therefore, (*) has been proven.

### 9.5 Hierarchical Predicate Transition Nets

The development of hierarchical predicate transition nets (HPrTNs) was motivated by the need to construct specifications for large systems using Petri nets (Reisig, 1987) and inspired by the development of modern high-level programming languages and other hierarchical and graphical notations, such as data flow diagrams (Yourdon, 1989) and statecharts (Harel, 1988). Similar work on introducing hierarchies into colored Petri nets was given in Jensen (1992, 1995). With the introduction of hierarchical structures into predicate transition nets, the resulting net specifications are more understandable, and the specification construction process becomes more manageable. HPrTNs were used in specifying several systems, including an elevator system (He and Lee, 1991), a library system (He and Yana, 1992), and a hurried dining philosophers system (He and Ding, 2001). HPrTNs can be analyzed directly by using a structural induction technique combining structural, behavioral, and logical reasoning (He, 2001) and can be translated into program skeletons in a concurrent and parallel object-oriented programming language C++ (He 2000c) and Java (Lewandowski and He, 1998, 2000). A complete formal definition of HPrTNs was given in He (1996). In the following subsections, basic concepts and notation of HPrTNs are briefly introduced.

#### 9.5.1 Definition of HPrTNs

An HPrTN is a structure:

$$HN = (NG, \text{Sig}, \text{V}, H, \text{Type}, \text{AN}, M_0, \rho).$$

Here:

- $NG = (P, T, F)$ is a net graph. $P$ and $T$ are finite sets of places and transitions such that $P \cap T = \emptyset$. Elements in $P$ are represented by solid and dotted circles. Similarly, elements in $T$ are represented by solid and dotted boxes. Solid circles or boxes are elementary nodes, and dotted circles and boxes are super nodes. In particular, we identify two subsets $IN \subseteq P \cup T$ and $OUT \subseteq P \cup T$ such that $IN$ contains the heads of all incoming nonterminating arcs (an arc inside a super node is a nonterminating arc if one of its ends is connected to the boundary of the super node) and $OUT$ contains the tails of all outgoing nonterminating arcs. Nodes in $IN \cup OUT$ are called interface nodes. We use $\bullet$ to denote the set of the presets of all elements in $IN$ (i.e., $\bullet IN = \{ n | n \in IN \}$), and $OUT \bullet$ to denote the set of the postsets of all elements in $OUT$. $F$ is the set of arcs and is called the flow relation, satisfying the conditions: $P \cap F = \emptyset$, $F \cap T = \emptyset$, and $F \subseteq (\bullet IN \times \bullet IN \cup P \times T \cup T \times P \cup \text{OUT} \times \text{OUT} \bullet)$. An arc $f$ can be uniquely identified by a pair of nodes $(n_1, n_2)$ denoting its source and sink, in which $n_1(n_2)$ may denote the preset (postset) of $n_2(n_1)$ when $f$ is a nonterminating arc.
• \( \text{Sig}, V, H \), Type are defined as in HLPN in Section 9.2.
• \( \mathcal{A}N = (A, TC) \) is a pair of annotation sets.

A is a function that annotates each arc with a term that when evaluated (for some binding) results in a multiset over the associated place's type. Furthermore, all simple labels of a compound label must have distinct identifiers, and all simple labels of arcs connected to the same node must have distinct identifiers. Because compound labels define data flows as well as control flows, the following basic control flow patterns (He and Lee, 1991) must be correctly labeled: (1) data flowing into and out of an elementary transition must take place concurrently, and (2) data flowing into and out of an elementary predicate can occur at different times. Furthermore, data flows between different levels of hierarchies must be balanced (i.e., a simple label occurs in a nonterminating arc if and only if it also appears in an arc with the same direction connected to the enclosing super node).

• \( TC \) is a function that annotates transitions with Boolean expressions. A super transition is an abstraction of low-level actions, and its meaning is thus completely defined by the low-level refinement. Therefore, the constraint of a super transition is true by default (it is conceivable that a nontrivial constraint for a super transition might be useful; however, in general it is very difficult to define such a constraint and also very difficult to interpret the constraint with regard to the operational (dynamic) semantics of the super transition).

• \( M_0 \) is the initial marking function.

• \( \rho P \cup T \rightarrow \rho (P \cup T) \) is a hierarchical mapping that defines the hierarchical relationships among the nodes in \( P \) and \( T \); this mapping also satisfies the constraint that the interface nodes \( \in \text{IN} \cup \text{OUT} \) be all predicates if their parent node is a predicate or all transitions if their parent node is a transition. For any node \( n, \rho(n) \) defines the immediate descendant nodes of \( n \). The ancestor and descendants of any node can be easily expressed by using well-known relations, such as transitive closure on a state node in an HPrTN is local to its parent and can be uniquely identified by prefixing its ancestors’ names separated with periods to its own name; however, often its own name is referred whenever a no name clash occurs.

The enabling condition of an elementary transition is defined exactly as that of an HLPN in Section 9.2. A super transition is enabled if at least one of its interface child transitions in \( \text{IN} \) is enabled and its firing is defined by an execution sequence of its child transitions; thus, its behavior is fully defined by its child transitions. The firing rule of a transition is formally defined in He (1996). Two transitions (including the same transition with two different occurrence modes) can fire concurrently if they are not in conflict (the firing of one of them disables the other). Conflicts are resolved nondeterministically. The firing of an elementary transition is atomic, and the firing of a super transition implies the firing of some elementary transition and may not be atomic. We define the behavior of an HPrTN to be the set of all possible maximal execution sequences containing only elementary transitions. Each execution sequence represents consecutively reachable markings from the initial marking in which a successor marking is obtained through a step (firing of some enabled transitions) from the predecessor marking.

Figure 9.3 shows an HPrTN specification of the Five Dining Philosophers problem.

Type(Thinking) = \( \varphi \) Eating = \( p \) (PHIL), Type(Used) = \( p \) (CHOP),
Type(Chop)) = \( \varphi \) (Avail) \cup \varphi \) (Used), Type(Chasing) = \( p \) (PHIL \times \text{CHOP} \times \text{CHOP})

\[ A(f_1) = < 1, \text{re} >, A(f_2) = < 2, \text{re} >, A(f_3) = < 3, \text{ph} >, \]
\[ A(f_4) = < 4, \text{ph} >, A(f_5) = < 5, \text{ph} >, A(f_6) = < 6, \text{ph} >, \]
\[ A(f_7) = < 7, \{A, B\}, A >, A(f_8) = < 8, \{A, B\}, A >, \]
\[ A(f_9) = A(f_5) \times A(f_3), A(f_{10}) = A(f_4) \times A(f_8), \]
\[ A(f_11) = A(f_2) \times A(f_5), A(f_{12}) = A(f_3) \times A(f_4), \]
\[ A(f_{13}) = A(f_1) \times A(f_2), A(f_{14}) = A(f_2) \times A(f_3), \]
\[ TC(\text{Pickup}) = \{\text{ph} = \text{re}[1]\} \land \{\text{ch}2 = \text{re}[2]\} \land \{\text{ch}2 = \text{re}[3]\}, \]
\[ TC(\text{Putdown}) = \{\text{ph} = \text{re}[1]\} \land \{\text{ch}1 = \text{re}[2]\} \land \{\text{ch}1 = \text{re}[3]\}, \]

Figure 9.3 Figure 9.3 Figure? If not, please provide an introductory sentence.
9.5.2 System Modeling Using HPrTNs

HPrTNs can be used to model systems using the traditional structured approach (Yourdon, 1989) or modern object-oriented approach (Booch, 1994). He and Ding (2001) proposed an approach to realize various object-oriented (OO) concepts in HPrTNs. Furthermore, He (2000a, 2000b) and Dong and He (2001) applied HPrTNs to formalize several diagrams in the Unified Modeling Language (UML)—a software industry standard second-generation object-oriented modeling language. In the following sections, we briefly describe how to model several key OO concepts in HPrTNs.

Classes

One of the central ideas of OO paradigm is data encapsulation captured by the class concept. A class is essentially an abstract data type with a name, a set of related data fields (or attributes), and a set of operations on the data fields. It is straightforward to use a predicate to denote a data field (structure) and a transition to represent an operation in Petri nets. The current value of a data field is determined by the tokens of the denoting predicate under the current marking. The meaning or definition of an operation is specified by the constraint associated with the denoting transition. HPrTNs were originally developed for structured analysis that provides separate mechanisms for data abstraction and processing abstraction through super predicates and super transitions, respectively. Therefore, we can use a super predicate and super transition pair in an HPrTN to capture the notion of a class although it is adequate to define a class by using a super predicate when there is no externally visible operation or using a super transition when there is no externally visible attribute. This view of class is a major improvement over the view in (He and Ding, 1996), where a class was represented by a super predicate only. In this view, the interface of the class is defined by the super predicate and the super transition. The super predicate defines data and internal operations of the class, while the super transition mainly defines the externally visible operations of the class. The corresponding subnets further define the internal structures of the data and the operations. The net inscription defines the meanings of net components through predicate types, token values, and transition constraints. When the resulting HPrTN is simple enough, there is no need to separate the super nodes from their subnets (i.e., the subnets are directly embedded inside the super nodes). An attribute or operation is externally visible if the corresponding denoting predicate or transition is an interface node (i.e., connected with a nonterminating arc). It should be noted, however, that not every super predicate or transition needs to be considered as a class. A super predicate or transition may simply denote a data abstraction or operation abstraction as originally intended; for example, a super predicate can be used to hide the internal states of an attribute that is defined by several related predicates, and a super transition can be used to define alternative implementations of an operation to realize operation overloading or overriding. Thus, our approach supports the coexistence of various modeling paradigms.

Based on the above analysis, we use the following C + + -like class schema to document a class defined by the super node(s) in an HPrTN (it is worth noting that the class schema is only used for understanding purpose and does not add functionality to the given HPrTN:

\[
\text{class Name [superclass(es)]} \\
\text{public:} \\
\text{predicates and transitions} \\
\text{[private:} \\
\text{predicates and transitions],}
\]

where brackets [...] denote optional items. Predicates and transitions listed in both public and private sections are those contained in the super node(s). The name(s) of the super node(s) are used to form the class name.

In the HPrTN shown in Figure 9.3, both super transition Phil and super predicate Chop can be viewed as classes. Thus, the following class definitions can be obtained:

\[
\text{class Chop} \\
\text{[public:} \\
\text{Avail, Used} \\
\text{],} \\
\text{class Phil} \\
\text{[public:} \\
\text{Pick up, Put down} \\
\text{private:} \\
\text{Thinking, Eating} \\
\text{].}
\]

Objects

An instance or object of a class has its own copy of data while sharing operations with other objects of the same class. To distinguish an object from other objects, a unique identifier is needed for each object.

In an HPrTN, an object is essentially defined by a set of tokens related through the same identifier, thus, the sort of any predicate \( p \) needs to contain a component sort of relevant identifiers (i.e., Type \( (p) = p(ID \times \ldots) \)). Different objects of a class share the same class data structures (i.e., tokens with different identifiers can reside in a predicate at the same time in an HPrTN). In general, however, objects of the same class cannot interact with each other directly. The above problem can be easily solved by defining a subexpression comparing token identifiers in the constraint of each transition. Movements of tokens and/or changes of token values while maintaining the object identifier indicate state changes of the object.

In the HPrTN shown in Figure 9.3, there are \( k \) philosopher objects with identifiers of sort PHIL, and there are \( k \) chopstick objects with identifiers of sort CHOP.
Class Reference Relation

Classes work together to fulfill the functionality of the underlying system. A class can use the operations and/or data provided by other classes.

In HPrTNs, the interface of a transition includes a box with a name and the labels of relevant arcs (the label identifiers determining the calling context and the flow expressions specifying parameters). The meaning of an elementary transition is defined by its constraint, and the meaning of a super transition is defined through its corresponding subnet.

It is easy to model a class reference by adding some arc when a class needs to access some public attribute of another class. For example, Figure 9.4 illustrates simple class reference relationships where some operation in class $C_1$ ($t_1$ and $p_1$) uses some public attributes defined in class $C_2$ ($p_2$ and $t_2$), and some operation in class $C_2$ uses some public attributes defined in class $C_1$. Figure 9.3 contains simple reference relationships between classes Phil and Chop.

To define an operation in one class using another operation in a different class, we cannot simply add an arc since Petri nets do not allow direct connections between transitions. As discussed earlier, there are two main ways to handle class reference relationships in the existing research works: (1) to fuse the two operations in two classes into one such that only synchronized communication is allowed (Biberstein and Buchs, 1995; Battison et al., 1995; Lakos, 1995b) and (2) to create some places inside one class to hold parameters to simulate message passing and function calls (Bastida, 1995; Lakos, 1995b) which supports asynchronous communication.

HPrTNs support both synchronous and asynchronous communications through reference places to model different communication protocols. These reference predicates do not belong to any class and can be viewed as connectors in software architecture languages (Shaw and Garlan, 1996). It is quite easy to model a function call through passing two messages by using one reference predicate to hold the input values and another to hold output results; another easy way to model a function call is by defining the calling operation (function) as a super transition whose subnet has at least two transitions to handle sending and receiving values.

Figure 9.5 shows the general pattern of a function call from class $C_1$ containing $t_1$ to class $C_2$ containing $t_2$, in which $p_1$ and $p_2$ are reference places. The above pattern defines a one-way synchronized communication (i.e., the caller must wait for the callee to continue its execution), whereas a simple message passing from an operation in one class to an operation in another class in general defines an asynchronous communication.

Figure 9.6 defines a general synchronization pattern such that two operations in class $C_1$ containing $t_1$ and $C_2$ containing $t_2$ must execute mutually exclusively, where $p$ is a reference predicate with an initial dummy token.

It is quite natural and easy to define class reference relationships by using the decomposition and synthesis techniques of HPrTNs discussed in He and Lee (1991).

Class Inheritance Relation

Another major feature of the OO paradigm is class inheritance relation that captures the generalization–specialization relationships in the real world (Coad and Yourdon, 1991). A class inheritance relationship exists between a superclass and a subclass such that the subclass inherits data structures as well as operation definitions from the superclass without the need to define them again. Thus, class inheritance relation supports a flexible and managed way to reuse existing data structures and operations.

A class inheritance relation is realized in HPrTNs through the reuse of the net structures of inherited super nodes and the net inscription of inherited elementary nodes (the sorts of predicates, the label expressions of relevant arcs, and the constraints of transitions) defined in an existing HPrTN denoting a class. The inherited predicates and transitions, however, are explicitly represented or embedded in the subclass to clearly define its role, the same convention was used by Lakos (1995b).

An inherited element in a subclass has a name of the form...
super_node.element_name, where super_node is the partial name of the superclass and element_name is the internal name of the element in the superclass. Renaming of relevant arcs are also necessary to reflect the current context and to ensure flow balance. It is clear that inheritance does not reduce the size of an HPrTN specification since inherited elements are embedded; an alternative way to embedding is through delegation, as explained by Abade and Cardelle (1996). However, the advantages are obvious since the meaning or structure (the most difficult part in writing an HPrTN specification) of an inherited element is already available and is obtained without any additional effort, furthermore, many known properties of the inherited element might be maintained through inheritance (structural properties are surely kept, but behavioral properties may need additional validation). It is worth noting that (1) only public components of a superclass can be inherited; (2) inheritances from multiple superclasses are supported, and an element can be inherited by multiple subclasses because no ambiguity will occur due to the naming convention; and (3) a redefined (overriding) operation is considered as a new operation in a subclass and is distinguished from an inherited operation such that an overriding operation in a subclass has the same name as the overridden operation in the superclass, this distinction between inheritance and overriding was also made by Abadi and Cardelli (1996).

Figure 9.7 shows a class inheritance relationship defined as follows:

```cpp
class p1 and t1
{ public:
    p2, t2, t3
};
class p3 and t3:
{ public:
    t3, p1 \cdot p2, t1 \cdot t2
    private:
    p4, p5, t6
};
```

**Figure 9.7** An Example of Class Inheritance Relation.

**Polymorphism**

OO paradigm also supports **polymorphism** such that an operation's name (with possibly different signatures) may have different meanings or behaviors (implementations) through inheritance or overriding.

Polymorphism can be achieved in HPrTNs in two different yet related ways. First, polymorphism is a major feature of the underlying many-sorted algebra \( H \) of an HPrTN; detailed discussions of algebraic specifications and polymorphism can be found in work by Ehrig and makr, (1985). The same operation symbol in \( H \) is used for many derived sorts. A simple example is the overloaded equality (\( = \)) operator when an algebraic specification \( H \) contains two elementary data types (or classes) INT and CHAR with a single parameterized definition of the equality (\( = \)). Second, polymorphism can be accomplished through net structure and inscriptions. An operation provides overriding capability if its constraint distinguishes a superclass object and a subclass object (or two objects from two different subclasses with the same superclass) and processes them differently. To realize polymorphism in an HPrTN, a shared predicate can be used to hold tokens of a superclass as well as tokens of subclasses, and the shared predicate is connected to the transition defined in the subclass and its inherited (or overriding) versions in the subclasses. The constraint of the original transition is only satisfied by the tokens of the superclass, and the constraint of each inherited (or overriding) transition is only satisfied with tokens of the subclass containing the transition.

Figure 9.8 shows the general pattern of realizing polymorphism through net structure in HPrTNs, in which \( p \) is a shared place, \( t_1 \) is a part of the superclass, and \( t_2 \) is a part of the subclass, and \( op_1^* \) is either an inherited or an overriding version of \( op_1 \), as shown in Figure 9.8. Furthermore, operation overriding can be achieved through a super transition in an HPrTN such that the firing of a particular component transition is determined partly by the (dynamic) instantiation of an object identifier (and thus its sort). The use of the above case-like net structure in realizing polymorphism can be avoided in the implementation of an HPrTN.

**Figure 9.8** Polymorphism Through Choice Structure.
9.6 Fuzzy-Timing High-Level Petri Nets

9.6.1 Definition of FTHN

A Fuzzy-timing high-level petri net (FTHN) is a structure:

\[ FN = (N, D, FT) \]

where \( N \) is an HLPN defined in Section 9.2; \( D \) is the set of all fuzzy delays, \( d_{tp}(\tau) \), associated with arcs \((t, p)\) from each transition \( t \in T \) to its output place \( p \); and \( FT \) is the set of all fuzzy time stamps. A fuzzy time stamp \( \pi(\tau) \in FT \) is associated with each token and each place. A fuzzy time stamp \( \pi(\tau) \) is a fuzzy time function or possibility distribution giving the numerical estimate of the possibility that a particular token arrives at time \( \tau \) in a particular place. Any type of possibility distribution can be represented or approximated by using a number of trapezoidal distributions. Thus, we use the trapezoidal possibility distribution specified by the five parameters, \( h(a, b, c, d) \) as shown in Figure 9.9, where \( h \) is the height having the following properties: \( 0 \leq h \leq 1 \), \( h = 1 \) for an event (arrival of a token) that has occurred or will occur and \( h < 1 \) for an event that will not necessarily occur. A so-called fuzzy number is represented by the triangular distribution \( h(a, b, d) \), which is a special case of the trapezoidal form with \( b = c \). A deterministic interval between \( a \) and \( d \) denoted \([a, d]\) can be represented by \((a, a, d, d)\), a special case of trapezoidal form with \( a = b = c = d \), and \( h = 1 \). In addition, given an arbitrary-shaped possibility distribution, we can approximate it with the union of a number of trapezoidal distributions (Zhou and Murata, 1999).

In Zhou et al. (2000), FTHN is extended by adding a time interval with a possibility value \( p \) in the form of \( p[\alpha, \beta] \) to each transition. That is, each transition is associated with a firing interval denoted \( p[\alpha, \beta] \), where the default interval is \([0, 0]\) (a transition definitely fires as soon as it is enabled). If a transition \( t \) is enabled at time instant \( \tau \), it may not fire before time instant \( \tau + \alpha \) and must fire before or at time instant \( \tau + \beta \). Possibility \( p \) is a value in the interval \([0, 1]\), where \( p = 1 \) if transition \( t \) is not in structural conflict with any other transition, and \( p \) can be less than 1 when we want to assign different chances to transitions in conflict. Here, 'structural conflict' means that a transition \( t \) shares some input place with another transition that can be enabled simultaneously with transition \( t \); firing one transition will disable the other transition.

9.6.2 Computation for Updating Fuzzy Time Stamps

Suppose that a transition \( t \) is enabled by \( n \) tokens and the fuzzy enabling time \( e_i(\tau) \) of transition \( t \) is computed by \( e_i(\tau) = \text{latest} \{\pi_i(\tau), i = 1, 2, \ldots, n\} \), where latest is the operator that constructs the "latest-arrival-lowest-possibility distribution" from \( n \) distributions (Murata, 1996; Murata et al., 1999). The \( \pi_i(\tau) \) is the fuzzy time stamp to which the enabling token arrives at the \( i \)th input place of transition \( t \). When there are \( m \) transitions in structural conflict that are enabled with their fuzzy enabling times, \( e_i(\tau), i = 1, 2, \ldots, t, \ldots, m \), and with their possibility intervals, \( p[\alpha_i, \beta_i] \), we compute the fuzzy occurrence time \( o_i(\tau) \) of transition \( t \) whose fuzzy enabling time is \( e_i(\tau) \) as follows:

\[ o_i(\tau) = \text{min} \{e_i(\tau) \oplus p[\alpha_i, \beta_i]\}, \]

where \( \oplus \) is the operator that constructs the "earliest-arrival-highest-possibility distribution" from \( m \) distributions, \( (\text{min}, \text{Murata, 1996; Murata et al., 1999}), \) \( \oplus \) is the extended addition (Dubois and Prade, 1989). We compute the fuzzy time stamp \( \pi_{\tau}(\tau) \), which is the fuzzy time distribution at which a token arrives at the transition \( t \) output place \( p \), as follows:

\[ \pi_{\tau}(\tau) = o_i(\tau) \oplus d_{\tau}(\tau) = h_1(o_1, o_2, o_3, o_4) \oplus h_2(d_1, d_2, d_3, d_4) \]

where \( h_{\tau}(\tau) \) is the fuzzy delay associated with the arc \((t, p)\) (Murded, 1996). When there are no transitions in conflict with transition \( t \) with its possibility interval \( p[\alpha_t, \beta_t] \), the fuzzy occurrence time is given by \( o_i(\tau) = e_i(\tau) \oplus \alpha_t, \beta_t \), where we set \( p = 1 \) since no conflict exists. We use the following formulas as approximate computations of the earliest and latest operations:

\[ \text{earliest} \{e_i(\tau), i = 1, 2, \ldots, n\} = \text{earliest}\{h_1(o_1, o_2, o_3, o_4), i = 1, 2, \ldots, n\} = \max \{h_1\} \left[ \min \{o_i\}, \min \{e_2\}, \min \{e_3\}, \min \{e_4\}\right], i = 1, 2, \ldots, n \]

\[ \text{latest}\{\pi_i(\tau), i = 1, 2, \ldots, n\} = \text{latest}\{h_1(\pi_1, \pi_2, \pi_3, \pi_4), i = 1, 2, \ldots, n\} = \min \{h_1\} \left[ \max \{\pi_i\}, \max \{\pi_2\}, \max \{\pi_3\}, \max \{\pi_4\}\right], i = 1, 2, \ldots, n \]

Using the above procedure, we compute and update fuzzy time stamps \( \pi(\tau) \), fuzzy enabling times \( e(\tau) \), and fuzzy occurrence times \( o(\tau) \) each time a transition firing (atomic action) occurs, starting from the initial (given) fuzzy timestamps of tokens in the initial marking \( M_0 \) and initially specified fuzzy delays, \( d_{\tau}(\tau) \). (See Zhou et al. [2000].)
9.6.3 Intended Application Areas and Application Examples of FTHNs

As seen in Subsections 9.6.1 and 9.6.2, the essence of FTHNs is the computation of updating fuzzy time stamps, and this computation involves only additions and comparisons of real numbers. Thus, computation can be done very fast, and FTHNs are suitable for performance analysis in real-time applications. The notion of fuzzy timing is highly flexible in that it can capture imprecise or incomplete knowledge regarding time behavior with specified or given possibility distributions, as well as the more conventional deterministic and probabilistic knowledge. However, it is expected that the traditional probabilistic approach and the FTHN method using possibility theory are complementary, rather than competitive, as possibility theory is considered to be complementary to probability theory (Zadeh, 1995). We expect that the FTHN method is more scalable than the traditional stochastic approaches because the FTHN computations are done using real arithmetic operations.

Some examples of FTHN applications include the following. A real-time network protocol used in local area networks (LANs) is modeled using FTHNs in (Murata, et al., 1999). Here we are interested in evaluating the worst-case performance in a given network, where propagation delays, times for processing message frames, and other such processes are specified as trapezoidal fuzzy-time functions. The fuzziness is due to the uncertain length (thus uncertain delay) of each message. In addition, FTHN models are used for performance evaluation of manufacturing systems, where the abilities of machines and/or workers involved in a manufacturing process are fuzzily known (Watanuki, 1999). Another area in which FTHN models are applied is the synchronization of multimedia systems (Zhou and Murata, 1998, 2001). Multimedia systems integrate a variety of media with different temporal characteristics (e.g., time-dependent media, such as video, audio, or animation) and time-independent media (e.g., text, graphics, and images). Thus, synchronization is a critical issue in these systems. The temporal specification has to be properly represented for presentation reviewing and planning by the user as well as for storing purposes. Multimedia synchronization has a time-critical problem because it must guarantee the temporal constraints for presenting the media items. Thus, there is a benefit in using FTHN methods to specify and analyze the temporal relations and specifications. The FTHN method has been used to present a new fine-grained temporal FTHN model for distributed multimedia synchronization (Zhou and Murata, 2001). The FTHN method has also been applied to a video-on-demand (VOD) system, which is a continuous playback multimedia application in which constant real-time media data are required for smooth real-time presentation. Thus, timing is a critical issue, and the response time is the only timing that has direct interaction with the subscribers and also impacts the quality-of-service (QoS), as discussed by Murata and Chen (2000). A nontrivial example of FTHN application is found in Zhou et al. (2000), where the FTHN method has been used to model networked virtual-reality systems, such as Cave Automatic Virtual Environment (CAVE) and Narrative Immersive Constructionist/Collaborative Environments (NICE) projects at the University of Illinois at Chicago. Using the FTHN models, various simulations have been conducted to study real-time behaviors, network effects, and performance (latencies and jitters) of the NICE. The simulation results are consistent with data and measurements obtained experimentally. This study shows how powerful FTHN models are in specifying and analyzing performance and how helpful the performance analysis is in improving a real-time networked-virtual-environment system design process.

References

Booch, G. (1994). Object-oriented analysis and design with applica-
tions. (2d. ed.), Benjamin/Cummings.


