Reporte de trabajo:

Nomenclatura y definiciones básicas de Redes de Petri

Ariel Sabiguero – asabigue@fing.edu.uy
Instituto de Computación
Facultad de Ingeniería

Abstract

Este reporte es un resumen de conceptos y notación empleada en el formalismo de Redes de Petri, realizado como subproducto del proyecto Clemente Estable Nº 4072 "Modelado y construcción de una máquina paralela virtual con componentes de bajo costo".
Es el resultado de la primer etapa del mismo en la que se realizó un estudio de herramientas matemáticas adecuadas para el modelado de performance. El presente documento introduce someramente notación y algunos aspectos de interés de las redes de Petri. El lector interesado en profundizar en el estudio de las mismas podrá continuar su estudio en cualquiera de las referencias bibliográficas que se presentan en este trabajo

Palabras clave

Redes de Petri, Redes de Petri temporizadas

Índice

Introducción

Las Redes de Petri (Petri Nets) fueron introducidas por C.A. Petri in 1962. Su éxito se debe básicamente a la simplicidad de su mecanismo básico, si bien, la representación de grandes sistemas es costosa. Se han hecho diversas extensiones al modelo básico de forma de simplificar su aplicación en diferentes campos.

Numerosos autores han extendido el modelo básico introduciendo el concepto de tiempo (Redes de Petri Temporizadas o Timed Petri Nets), con lo que se pueden lograr análisis cuantitativos de performance de los sistemas. Cuando se utilizan variables aleatorias exponenciales para especificar el comportamiento temporal del modelo, el modelo se denomina Red de Petri Estocástica (Stochastic Petri Net, SPN). Se puede probar que, bajo ciertas hipótesis, las Redes de Petri Estocásticas son isomórficas a Cadenas de Markov de tiempo contínuo homogéneas.

La capacidad de las Redes de Petri Estocásticas para describir simultáneamente el paralelismo y la sincronización en sistemas multiprocesadores se presenta como una alternativa viable cuando el nivel de detalle necesario hace inadecuados los modelos basados en teoría de colas.

Redes de Petri Standard

La estructura de una Red de Petri *standard* es un grafo bipartito que incluye un conjunto de *lugares P (places)*, un conjunto de *transiciones T (transitions)* y un conjunto de arcos dirigidos *A*. Un lugar *p* es una *entrada (input)* para una transición *t* si existe un arco incidente a la transición *(p,t)*. Un lugar *p* es una *salida (output)* de una transición *t* si existe un arco incidente de la transición al lugar *(t,p)*. El conjunto de arcos puede ser particionado en el conjunto de arcos de entrada a transiciones A_i y el conjunto de arcos de salida de transiciones A_o.

Formalmente:

$$
\begin{aligned}
PN &= (P, T, A) \\
P &= \{p_1, p_2, \ldots, p_n\} \\
T &= \{t_1, t_2, \ldots, t_m\} \\
A_i &\subset P \times T \\
A_o &\subset T \times P \\
A &= A_i \cup A_o
\end{aligned}
$$

i.1. *Marcado de Redes de Petri*

Las Redes de Petri pueden contener *fichas (tokens)*, gráficamente dibujados como puntos negros. Una Red de Petri con tokens se llama *Red de Petri con marcado*. El estado de una Red de Petri con marcas está definido por el número m_i de tokens contenidos en cada lugar l_i. El estado de la Red de Petri se llama usualmente el marcado y se denota por $M = \{m_1, m_2, ..., m_n\}$.

Una definición formal de una Red de Petri con marcado es:

$$PN = (P, T, A, M_0)$$
$$P = \{p_1, p_2, ..., p_n\}$$
$$T = \{t_1, t_2, ..., t_m\}$$
$$A_i \subset P \times T$$
$$A_o \subset T \times P$$
$$A = A_i \cup A_o$$
$$M_0 = \{m_{01}, m_{02}, ..., m_{0n}\}$$

dónde denota el número de fichas en el lugar p_i en el marcado inicial M_0.

Es posible definir la *ejecución* de una Red de Petri con marcas. La misma se define a través de las siguientes reglas:

1. Una transición está habilitada cuando todos sus lugares de entrada contienen al menos una ficha
2. Una transición habilitada se puede *disparar (fire)*, removiéndose un token de cada lugar de entrada y colocando uno en cada lugar de salida.
3. Cada disparo de una transición modifica la distribución de las fichas, y por ello produce un nuevo marcado en la red.

De aquí en más notaremos m_{pi} al número de tokens en el lugar p_i, $\#(p_a,t_b)$ vale 1 si existe un arco desde el lugar p_a a la transición t_b y 0 en cualquier otro caso; de forma similar, $\#(t_a,p_b)$ vale 1 si existe un arco desde la transición t_a al lugar p_b y 0 en cualquier otro caso. Una transición t_b está habilitada si

$$m_{pi} \geq \#(p_i, t_b) \forall p_i \in P,$$

Ecuación i.1-1

y el resultado del disparo de t_b es un nuevo marcado $M' = \{m'_1, m'_2, ..., m'_n\}$ definido como $m'_{pi} = m_{pi} - \#(p_i, t_b) + \#(t_b, p_i) \forall p_i \in P$

Ecuación i.1-2

i.2. Extensiones a las Redes de Petri standard

Las extensiones fueron introducidas para aumentar el poder de modelado de la herramienta. En esta sección consideraremos en esta sección son las definiciones de *arcos múltiples* y *arcos inhibidores.*

Decimos que tenemos *arcos múltiples* cuando se relaja la condición de unicidad de arcos en el grafo bipartito y pasamos a tener un multigrafo bipartito. Las ecuaciones que definen el estado habilitado de una transición (

Ecuación i.1-1,Ecuación i.1-2) mantienen su validez modificando la definición de $\#(p_a,t_b)$ y $\#(t_a,p_b)$ de la siguietnte forma:

$\#(p_a,t_b)$ denota el número de arcos desde el lugar p_a a la transición t_b.

$\#(t_a,p_b)$ denota el número de arcos desde la transición t_a al lugar p_b.

Los *arcos inhibidores* se representan gráficamente con un círculo en la punta (en lugar de la flecha), y su función es la de inhibir transiciones. La condición de disparo puede generalizarse diciendo que todos sus lugares de entrada normales contienen al menos una ficha y no hay ninguna ficha en sus lugares de entrada inhibidores. El disparo de una transición remueve una ficha de cada uno de sus lugares de entrada normales.

Estas dos extensiones tienen una naturaleza teórica muy diferente y por ello sus consecuencias. La introducción de arcos múltiples es esencialmente una convención que permite –en algunos casos- dibujar grafos más simples. La introducción de arcos inhibidores aumentan el poder expresivo del modelo.

Las Redes de Petri con arcos inhibidores y múltiples será considerada como la base de aquí en más.

i.3. Propiedades de las Redes de Petri.

Un marcado *M’* se dice *alcanzabe inmediatamente* desde *M* si *M’* puede ser obtenido disparando una transición habilitada en *M*. La ejecución de la Red de Petri permite definir una secuencia de marcados $\{M_0,M_1,M_2,\ldots\}$ y una secuencia de transiciones $\{t_1,t_2,\ldots\}$. Un marcado *M”* se dice *alcanzable* desde *M* si existe una secuencia de disparos de transiciones que mueve el estado de la Red de Petri desde *M* a *M”*.

Definimos el *conjunto alcanzable* $R(M_0)$ de una Red de Petri como el conjunto de todos los marcados que son alcanzables desde M_0. El conjunto alcanzable de una Red de Petri puede ser representado por un árbol, que se puede construir colocando el marcado inicial M_0 como raíz y como hijos todos los marcados inmediatamente alcanzables desde éste y repitiendo este procedimiento para cada marcado que se agrega. Existe la posibilidad de generar estructuras infinitas cuando a partir de un marcado *M’* se puede alcanzar un marcado *M”* y viceversa. En estos casos dejamos de expandir el árbol apenas llegamos a un marcado agregado anteriormente. Llamaremos a estos marcados *marcados duplicados*. Durante la construcción de árbol de alcance, podemos encontrar también *marcados muertos*, aquellos en los que ninguna transición está habilitada. Los marcados duplicados y muertos constituyen las hojas el árbol.

Una Red de Petri con marcas se dice *segura* si el número de fichas en cada lugar ≤ 1 para todos los marcados de $R(M_0)$. Una Red de Petri se dice *k-acotada* si el número de fichas en cada lugar es $\leq k$ para todos los marcados de $R(M_0)$.

Una Red de Petri con arcos múltiples y arcos inhibidores tienen un conjunto de estados alcanzables finito y por ello puede ser representado mediante una máquina de estados finitos. Agregando arcos inhibidores (capacidad de test cero) a Redes de Petri que no son *k-acotadas* se le da al modelo el poder computacional de una Máquina de Turing, permitiendo por ello modelar cualquier sistema computable.

Una Red de Petri se dice *estrictamente conservativa* si, para todos los marcados de $R(M_0)$, la suma de fichas en la Red es constante, es decir, las fichas se mueven por la red sin ser creadas o destruidas.

$$\forall M_k \in R(M_0), \sum_{i=1}^{n} m_{ki} = k$$

Ecuación i.3-1

Esto implica que para cada transición el número de arcos de entrada es igual al número de arcos de salida.

Una Red de Petri se dice *conservativa* con respecto a un vector de componentes no negativas $W=\{w_1,w_2,w_3,\ldots,w_n\}$ si,

$$\forall M_k \in R(M_0), \sum_{i=1}^{n} w_i m_{ki} = k$$

Ecuación i.3-2

Cuando una Red de Petri se utiliza para modelar sistemas reales, el concepto de conservación es de suma relevancia, pues generalmente las fichas son asociadas a recursos, los cuales, generalmente, no pueden variar (p.e. el número de procesadores), y como consecuencia, su número no puede variar en la Red.

Una transición $t_I \in T$ se dice *viva* si, para todo $M \in R(M_0)$, hay un marcado M', alcanzable desde M, que habilita la transición t_I. Una Red de Petri se dice *viva* si todas sus transiciones t_I son vivas. El concepto de vida es muy importante en sistemas de computadoras, y usualmente es asociado con la ausencia de estados de bloqueos mutuos o *deadlocks*.

El subconjunto de las Redes de Petri que se obtiene exigiendo que, para todo par de transiciones t_I y $t_2 \in T$, con $t_I \neq t_2$, y para cualquier marcado alcanzable, el disparo de t_I no puede deshabilitar el disparo de t_2. Estas redes son llamadas *persistentes*.

Redes de Petri con tiempos

Las Redes de Petri standard no incluyen concepto alguno de tiempo, por ello, solamente es posible describir solamente la estructura lógica de los sistemas y no su evolución temporal.

La introducción del tiempo en el modelo permite la descripción de comportamientos dinámicos de los sistemas, considerando la evolución de estados y la duración de cada acción tomada por el sistema. Hay múltiples formas diferentes de introducir el concepto de tiempo.

Una primer posibilidad consiste en asociar a cada transición un número que indica, en alguna unidad temporal adecuada, el *retardo (delay)* entre la habilitación y el disparo de la transición. Una Red de Petri Temporizada (*timed Petri Net TPN*) puede ser definida como:

$$TPN = \{P, T, A, M_0, \Theta\}$$

donde P, T, A y M_0 se definen como antes y $\Theta=(\theta_1, \theta_2,\ldots,\theta_m)$ es el conjunto de retardos asociados con las transiciones.

Una segunda posibilidad para la introducción del concepto de tiempo consiste en asignar un retardo θ al proceso de convertir una ficha en disponible luego que la misma llega a un nuevo lugar. Por ello, cada ficha puede estar en uno de dos estados: *disponible* y *no disponible*; solamente fichas disponibles habilitan transiciones. La falta de disponibilidad de una ficha modela el tiempo utilizado desarrollando una actividad. En esta abstracción, el tiempo es asociado a los lugares.

La adición de tiempos en las Redes de Petri es un proceso crítico y se deberá prestar atención especial a la comprensión total de la semántica del modelo y a los detalles de su comportamiento.

Existe una amplia variedad de extensiones adicionales, que básicamente consisten an adicionar tiempos a las diferentes componentes del grafo bipartito que constituye la red.

Redes de Petri Estocásticas

Los modelos de performance probabilísticos tratan de representar el comportamiento de sistemas determinísticos complejos por medio de procesos estocásticos. De esta forma es posible evitar una detallada descripción determinística de las operaciones del sistema, sustituyéndola por asunciones probabilísticas, que capturen la esencia del sistema.

Las Redes de Petri Estocásticas (SPN) se obtienen asociando con cada transición en una Red de Petri una variable aleatoria con distribución exponencial que exprese el retardo desde la habilitación hasta el disparo de la transición. Eliminando las variables aleatorias de una SPN se obtiene la Red de Petri asociada.

Consideremos una SPN y un marcado M en el cual múltiples transiciones están simultáneamente habilitadas. La transición que tiene asociado el retardo más breve disparará primero. La SPN alcanza un nuevo marcado M', en el cual algunas transiciones estuvieron habilitadas en el marcado M pero que no fueron disparadas y pueden aun estar habilitadas. Debido a la propiedad de *falta de memoria* de las variables aleatorias exponencialmente distribuidas, obtenemos una distribución de vida igual a la distribución del retardo de disparo en sí mismo. Se puede asumir que la actividad asociada con cada transición *recomienza* con cualquier nuevo marcado. Esto es válido inclusive cuando se está modelando actividades que se suceden en forma continua: el modelo no *"siente"* la repetición de actividades asociadas con una transición.

Una definición formal de una SPN es:

$SPN = \{P, T, A, M_0, L\}$

donde P, T, A y M_0 se definen como antes y $L=(l_1, l_2,\ldots,l_m)$ es el conjunto de tasas de retardos asociados con las transiciones, posiblemente dependientes del marcado, asociadas con las transiciones de la Red de Petri. Cuando sea necesario, la dependencia con un marcado dado M se representará como $l_j(M)$.

Se puede probar que, debido a la propiedad de falta de memoria de la distribución exponencial de los retardos en los disparos, las SPN son isomórficas a cadenas de Markov de tiempo continuo. En particular, una SPN *k-acotada* es isomórfica a una MC finita. La misma se puede obtener aplicando las siguientes reglas:

1. El espacio de estados S de la MC corresponde al conjunto de alcance $R(M_0)$ de la Red de Petri asociada con la SPN $(M_i \leftrightarrow i)$.
2. La tasa de transición del estado i (correspondiente a Mi) al estado j (Mj) es $q_{ij} = \sum_{k \in H_{ij}} l_k$ donde H_{ij} es el conjunto de transiciones habilitadas por el marcado Mi, cuyos disparos generan el marcado Mj.

Una SPN se dice *ergódica* si genera una CTMC. Es posible mostrar que una SPN es ergódica si M_0, el marcado inicial, es alcanzable desde cualquier $M_i \in R(M_0)$. Si la SPN es ergódica, es posible calcular la probabilidad de distribución de marcados en el estado estacionario resolviendo la ecuación matricial

$\pi Q=0$

con la restricción adicional

$\Sigma \pi_i=1$

donde Q es el generador infinitesimal cuyos elementos se obtienen por el método de construcción de la MC anterior y π es el vector de probabilidades del estado estacionario. A partir de la distribución de probabilidades del estado estacionario es posible obtener estimaciones cuantitativas del comportamiento de la SPN.

Redes de Petri Estocásticas Generalizadas

Frecuentemente no es deseable asociar un tiempo aleatorio a cada transición. Naturalmente se tiende a asociar tiempos con las transiciones más lentas o aquellas que se cree que tienen un mayor impacto en la performance global del sistema y a asumir instantáneas aquellas muy breves o cuya duración se estima que no tendrá impacto en la performance global.

Un ejemplo típico puede ser el caso en el cual las actividades que componen el sistema poseen diferencias en duración de órdenes de magnitud. La opción se convierte en particularmente conveniente si el número de estados de la Cadena de Markov asociada se reduce.

Las Redes de Petri Estocásticas Generalizadas (GSPN) se obtienen permitiendo que las transiciones pertenezcan a dos clases diferentes: *inmediatas* y *temporizadas*. Las transiciones inmediatas disparan en tiempo cero una vez que están habilitadas. Las transiciones temporizadas disparan luego de un tiempo de habilitación aleatorio, exponencialmente distribuido.

Una definición formal para las GSPN se obtiene haciendo que los elementos del arreglo *L* sean solamente elementos *m'*, aquellos asociados a transiciones temporizadas.

Si el conjunto de transiciones habilitadas *H* incluye solamente transiciones temporizadas, entonces la transición t_i ($i \in H$) dispara con probabilidad

$$\frac{l_i}{\sum_{k \in H} l_k}$$

Ecuación i.34-1

como en el caso de SPN. Si *H* involucra transiciones temporizadas e inmediatas simultáneamente, únicamente dispararán transiciones inmediatas. Si *H* involucra cero o más transiciones temporizadas y solamente una transición inmediata, solamente disparará la transición inmediata. Cuando *H* involucra varias transiciones inmediatas, es necesario especificar la función de densidad de probabilidad en el conjunto de transiciones inmediatas habilitadas de a cuerdo a la cual se obtiene la transición que dispara. El subconjunto *H* que involucra todas las transiciones inmediatas habilitadas, junto con las probabilidades de distribución es llamado *switch aleatorio (random switch)*.

Glosario

alcanzable, 5

arcos inhibidores, 4

arcos múltiples, 4

conjunto alcanzable, 5

disparo, 4

ejecución, 4

fichas, 3

inmediatamente alcanzable, 5

random switch, 9

Red de Petri con marcado, 3

Red de Petri Estocástica, 3

Redes de Petri, 3

Redes de Petri Temporizadas, 3

retardo, 6

SPN, 3

SPN ergódica, 8

switch aleatorio, 9

TPN, 6

Bibliografía

[ABC1] M. Ajmone Marsan, G. Balbo, G. Conte – "Performance Models of Multiprocessor Systems" – MIT Press – ISBN 0-262-01093-3 – 1986

[LIN1] Christoph Lindemann – "Performance Modelling with Deterministic and Stochastic Petri Nets" – 1998 – ISBN 0 471 97646 6