

REDES DE PETRI CONTINUAS: EXPRESIVIDAD, ANÁLISIS Y CONTROL DE UNA CLASE DE SISTEMAS LINEALES CONMUTADOS

M. Silva ¹ L. Recalde ¹

*Instituto de Investigación en Ingeniería de Aragón, I3A
Dep. de Informática e Ingeniería de Sistemas,
Universidad de Zaragoza, España,
{silva, lrecalde}@unizar.es.*

Resumen: Gracias a la existencia de potentes teorías de análisis y síntesis, así como a su directa representabilidad gráfica, las redes de Petri constituyen uno de los formalismos más aceptados en las aplicaciones de ingeniería en las que convienen modelos (de eventos) discretos. No obstante, a veces surgen problemas de decidibilidad o, en sistemas con grandes poblaciones, de enorme complejidad computacional. En este trabajo se presenta la urdimbre y algunos flecos de la teoría desarrollada para una relajación fluida o continua de los modelos discretos y dinámicos. Con ella se puede abordar eficientemente el estudio de sistemas de otra forma no abordables, al tiempo que se establecen algunos puentes a conceptos y resultados del análisis y síntesis de prestaciones, o de la observación y control de sistemas denominados continuos, aunque las clases de modelos que se obtienen son técnicamente, a veces también conceptualmente, híbridos. Obviamente, el precio que se paga por toda relajación es una pérdida de fidelidad en los modelos. En este sentido se ha de tener en cuenta que, aunque insólito en aplicaciones de ingeniería, los modelos discretos pueden ser no continuizables, del mismo modo que no todo sistema continuo no-lineal se puede aproximar satisfactoriamente por uno lineal (piénsese, por ejemplo, en sistemas caóticos). *Copyright ©2007 CEA-IFAC*

Palabras Clave: Red de Petri, ciclo de vida, relajación matemática, teoría estructural, análisis cualitativo, análisis cuantitativo, optimización paramétrica, observación y control dinámico.

1. INTRODUCCIÓN

Las redes de Petri (RdP) constituyen una familia de formalismos bien conocidos para el modelado, análisis y síntesis de sistemas concurrentes y dis-

tribuidos formalizados como discretos (Peterson 1981, Brams 1983, Silva 2002, Silva 1993, Jensen 1997, David y Alla 2004, Valk y Girault 2003). Su centralidad conceptual se constata al considerar que son definibles axiomáticamente (por el propio Petri), como Sistema de Adición de Vectores (Vector Addition System, VAS), a través de la Teoría de Regiones de un grafo etiquetado (por codificación del conjunto de nodos-estados globales), o a

¹ Los desarrollos presentados se han beneficiado en gran parte de los proyectos CICYT y FEDER DPI2003-06376 y DPI2006-15390, así como de la subvención del Gobierno de Aragón al GISED como Grupo de Excelencia en Investigación (2003-2007).

partir de la lógica lineal (lógica no monótona de Girard) (Silva y Teruel 1996).

Partiendo de la base de que es impropio hablar de sistemas continuos o discretos, ya que éstas son características de los modelos, cabe apuntar que por abuso del lenguaje se han lexicalizado las dos expresiones, así como la de sistema híbrido. La matematización como sistemas continuos dinámicos se remonta al menos a finales del siglo XVII, principios del XVIII, época en que se configura el *cálculo sublime* (cálculo diferencial e integral) merced al impulso, entre otros, de Newton o Leibnitz, así como el *cálculo de variaciones*, donde de mencionar un único apellido habría que optar por el de Bernoulli, sabiendo, que en realidad nos referimos a toda una saga (Sussmann y Willems 1997). En suma, más de trescientos años de perspectiva avalan estos desarrollos.

Sin embargo, la visión de sistemas dinámicos como (*de eventos*) *discretos* es mucho más reciente. Disciplina vertebrada en la intersección de la Automática, la Investigación de Operaciones y la Informática, se puede considerar su nacimiento dentro de la segunda mitad del pasado siglo XX. Por ello, entre otras cosas, pero muy principalmente por la necesaria multifuncionalidad a lo largo del ciclo de vida, no existe para ellos un marco formal único, que sea aceptado con idéntico consenso al que existe para las ecuaciones diferenciales en el ámbito de los sistemas modelados como continuos.

En los sistemas discretos se asume que las variables de estado, memorización del pasado dinámico, se pueden modelar como discretas, codificándose, según los formalismos, con alfabetos varios, o numéricamente, por ejemplo. En las redes de Petri el estado, denominado *marcado*, se codifica con variables: *booleanas*, en las redes Elementales; *numéricas* (con naturales), en las redes Lugar/Transición; o con *polinomios* en ciertas redes de Alto Nivel (por ejemplo, las redes Coloreadas). Dicho esto, los desarrollos que se presentan en este trabajo parten del nivel de abstracción intermedio, de entre los tres mencionados. En cualquier caso, aceptando que puede convenir que un sistema sea modelado en un momento determinado como discreto, cabe observar que los cambios de valor de su estado serán bruscos, puesto que se “salta” de un valor discreto a otro. Esos cambios se producen merced a *eventos* (sobre apuntes lingüísticos acerca de cuestiones relativas a “sistemas de eventos discretos”, véase (Silva 2006b)). Eso hace que los modelos basados en redes de Petri se construyan empleando dos clases de objetos: los *lugares* (que designan a las variables de estado locales) y las *transiciones* (cuya actividad-eventos- puede cambiar el valor de algunas de las anteriores). Se denomina *marcado del sistema*

discreto dinámico al conjunto de los valores de las variables denominadas lugares.

Las *redes de Petri* se definen como modelos en los que estados y eventos están en alternancia. Se representan gráficamente mediante grafos bipartidos, como ocurre también con las *redes de Colas* (RdC), o incluso con formalismos que manejan visiones continuas como son los *diagramas de Forrester* (DdF). Los almacenes (de órdenes de ejecución, de productos parcialmente elaborados, de recursos...) se denominan *colas* (RdC), *depósitos* (DdF), o *lugares* (RdP); la evolución en los anteriores se realiza merced a *estaciones* donde residen *servidores* (RdC), *válvulas* (DdF) o *transiciones* (RdP). En todos ellos subyace una lógica de evolución del tipo producción/consumo: para producir algo, se necesita consumir otros elementos, bien definitivamente, o bien temporalmente (por ejemplo, recursos reutilizables, que son retornados tras la operación). La *localidad* de los estados parciales (de donde el marcado global es un estado distribuido), y de los eventos hacen que en la construcción de modelos se puedan emplear estrategias *descendentes* y *modulares* (o *ascendentes*).

Las redes de Petri no son un simple formalismo, del mismo modo que el diseño y realización de un sistema no es algo que se produzca monolíticamente en una simple fase o etapa. Dicho de otro modo, el *ciclo de vida* de los sistemas (en particular de producción) comprende muy diversas fases, y su creciente complejidad hace que el empleo de métodos formales sea cada vez más importante. De forma tradicional se procede empleando para cada fase un formalismo diferente (grafos o diagramas de estado, máquinas de estado algorítmicas, redes de colas, PERTs o grafos conjuntivos/ disyuntivos...), entre los considerados específicamente más “adecuados”. No obstante, es bien sabido que la “optimización” de las partes de un sistema no lleva forzosamente a un óptimo global, algo que aquí se puede aplicar a la secuencia de visiones o fases de diseño y operación del mismo. Por ello, se aboga por la idea de emplear un *paradigma* formal de modelado que permita acometer económicamente el trabajo en su conjunto, facilitándose la coherencia entre fases, así como el establecimiento de sinergias entre las mismas. Es decir, se trata de optimizar globalmente el ciclo de vida.

El paradigma de modelado basado en redes de Petri se puede ver construido sobre la base de los diferentes niveles de abstracción de los modelos (con variables booleanas, numéricas, polinomios...) (Silva y Teruel 1996) y de las interpretaciones (asociación de significados a los lugares y transiciones), así como la conexión del modelo con el mundo exterior (guardas o condiciones lógicas externas, eventos, temporizaciones, procedimien-

tos de ordenación a niveles superiores...) (Silva y Teruel 1998, Silva 2006a). A la *coherencia* y *economía* que supone el empleo de diferentes modelos de un mismo paradigma, hay que añadir beneficiosos efectos de *sinergia* en los análisis de los diferentes modelos por fases (la comprensión de una determinada propiedad en uno de ellos permite mejorar la de otras en otros modelos). Por ello, en nuestra estrategia de investigación siempre procuramos aplicar esta visión de conjunto, desplazándonos en los difusos límites conceptuales y metodológicos de las teorías de Control, Informática, e Investigación de Operaciones, realizando pequeñas transgresiones que dotan a la investigación de alicientes adicionales.

Las visiones plasmadas por los modelos (de eventos) discretos son particularmente interesantes en sistemas realizados artificialmente (de manufactura, de transporte, logísticos, de coordinación de trabajos...). Ahora bien, si el sistema está muy poblado (i.e., existen muchos elementos, sean activos o pasivos: piezas, vehículos, clientes, posiciones en almacenes...) el análisis del mismo puede ser computacionalmente muy dificultoso, incluso imposible en la práctica, bien por incapacidad de decidir, bien por problemas de *explosión de estados* (globales). En suma, a pesar de las potencias de cálculo disponibles, los problemas pueden ser difícilmente tratables en la práctica. Ello lleva a la idea de *relajar* el modelo. Es decir, a asumir una cierta pérdida de fidelidad, con la esperanza de ganar en capacidad de análisis, en particular reducir drásticamente la complejidad computacional asociada a los procedimientos de análisis o síntesis.

En el caso presente, se relajan las redes Lugar/Transición por *fluidificación* o *continuización*. La idea no es nueva en el ámbito de los modelos formales. Por ejemplo, se utiliza cuando se pasa de un *problema lineal de programación en los enteros* a otro *problema de programación lineal en los reales* (eventualmente no negativos), lo que finalmente supone algoritmos polinómicos de cálculo, o al relajar una red de colas en una fluida (Mandelbau y Chen 1991, Chen y Yao 2001). En el caso de redes de Petri se relaja el disparo de las transiciones, lo que lleva a que las variables naturales que representan el marcado de los lugares tomen valores en los reales no negativos. La presentación de las primeras ideas de fluidificación en redes de Petri se remontan a dos décadas atrás, con motivo del *8th European Workshop on Applications and Theory of Petri Nets* (Zaragoza, junio, 1987), ahora *International Conference on Applications and Theory of Petri Nets*. René David (INP Grenoble) las planteaba inspirado en la fluidificación de redes de colas (para una revisión en marco amplio, (David y Alla 1992)), mientras que con José Manuel Colom, proponíamos el uso

sistemático de la programación lineal y geometría convexa para el estudio de sistemas discretos, a partir de su ecuación de estado (una visión de conjunto de lo perseguido, y lo obtenido en esta segunda aproximación se puede ver en (Silva *et al.* 1998)). Más recientemente, se reflejan resultados de ambas aproximaciones en (David y Alla 2004).

En esencia, partiendo de un modelo de red discreto, la fluidificación se aplica al disparo de una transición (ahora se podrá disparar media vez, o una y tres cuartos, por ejemplo), con lo que su contador de disparos y el marcado de los lugares en su entorno devienen variables continuas. Si se fluidifican todas las transiciones, el modelo de red se dice *continuo* (aunque técnicamente sea híbrido, como se verá posteriormente). Si no, coexistirán marcados enteros y continuos, y la red será denominada *híbrida*. En este trabajo se resumen algunas de las líneas desarrolladas en el ámbito de las redes continuas, siempre buscando la coherencia con lo obtenible del modelo discreto del que se asume se puede derivar, y empleando al máximo los resultados de la teoría *estructural* de análisis de las redes, es decir, resultados en los que el estado (marcado) inicial es, en esencia, un parámetro; dicho de otro modo, considerar simplemente como un último recurso el empleo de técnicas de simulación o de enumeración de estados, aproximaciones a las que, sin embargo, en el estado actual de conocimientos, hay que acudir con cierta frecuencia en el desarrollo de la ingeniería correspondiente.

Finalmente, cabe apuntar que existen modelos alternativos híbridos basados en redes de Petri. Por un lado, fluidificando sólo unos pocos lugares, sobre la base de una red estocástica, línea propuesta desde la Universidad de Duke (Horton *et al.* 1998); en el LAAS (Champagnat *et al.* 1998, Champagnat *et al.* 2001) se consideran modelos híbridos basados en una red discreta, asociándole a los marcados de lugares sistemas de ecuaciones diferenciales (próximo a ésta: (Demonogdin y Koussoulas 1998)). Muy centrado en las líneas desarrolladas por nuestro grupo en Zaragoza desde 1999, con la colaboración de colegas de diversas universidades, el presente trabajo parte recordando qué es una red de Petri autónoma (no interpretada) discreta y cómo se fluidifica (o continuiza). Posteriormente se hace énfasis en técnicas de análisis, así como en las diferencias que pueden exhibir los modelos discreto y continuo asociados a una misma red. Se destacan aspectos como la decidibilidad y la relativamente baja complejidad algorítmica para problemas de análisis y síntesis frecuentes. La introducción de una base temporal asociada al disparo de transiciones plantea problemas porque la elección de la semántica más apropiada (problema de modelado al fin y al cabo) puede llevar a conclusiones relativamente diferen-

tes. En particular se exploran la semántica de *finitos servidores* (“pocos” servidores, pero “muchos” clientes), y la de *infinitos servidores* (“muchos” servidores y “muchos” clientes); también se menciona una semántica basada en el *producto de marcados*, derivable de la de infinitos servidores, de especial interés para modelos poblacionales. La potencia expresiva con infinitos servidores es tal que se pueden simular *máquinas de Turing*, con lo que ello arrastra de indecidibilidades; por otro lado, el empleo de productos permite la simulación de modelos *caóticos*. Por último, la búsqueda de interpretaciones a nivel de red de conceptos como *observabilidad*, *control óptimo en régimen permanente* (problema de complejidad polinómica si todas las transiciones son controlables y la función objetivo es lineal), o *control dinámico* (esencialmente, pero no únicamente basado en aproximación MPC, Model Predictive Control), cierran la presentación.

2. REDES DE PETRI AUTÓNOMAS DISCRETAS Y CONTINUAS. MARCO CONCEPTUAL.

2.1 Red de Petri discreta.

Definición 1. (Red de Petri discreta). Una red de Petri es una estructura definida como $\mathcal{N} = \langle P, T, \mathbf{Pre}, \mathbf{Post} \rangle$ donde P y T son conjuntos disjuntos de lugares y transiciones respectivamente, y $\mathbf{Pre}, \mathbf{Post}$ son las matrices de pre- y post-incidencia en $\mathbb{N}^{|P| \times |T|}$.

Una RdP se puede representar gráficamente como un grafo bipartido dirigido: los lugares se representan como círculos, y las transiciones como barras o rectángulos. $\mathbf{Pre}[p, t] = w > 0$ significa que hay un *arco* de p a t con peso o multiplicidad w , y $\mathbf{Post}[p, t] = w > 0$ representa un arco de t a p con peso w . Dado un nodo $v \in P \cup T$, se define $\bullet v$ ($v \bullet$), como el conjunto de sus nodos de entrada (salida). Para que la RdP represente un sistema dinámico hay que definir un estado inicial.

Definición 2. Un sistema dinámico $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ está compuesto por una red de Petri \mathcal{N} y un *marcado inicial* $\mathbf{m}_0 \in \mathbb{N}^{|P|}$, que no es más que el estado distribuido inicial.

La evolución del marcado (estado) se basa en una regla de disparo que responde a una lógica de consumo/producción de recursos; se puede enunciar como: si existen suficientes recursos puede (no tiene por qué) hacerse la evolución.

Una transición t está *sensibilizaba* en un marcado \mathbf{m} si $\mathbf{m} \geq \mathbf{Pre}[P, t]$; y su disparo lleva a un nuevo marcado $\mathbf{m}' = \mathbf{m} + \mathbf{C}[P, t]$. Esto se denota como

$\mathbf{m} \xrightarrow{t} \mathbf{m}'$, y se dice que \mathbf{m}' es un marcado alcanzable (desde \mathbf{m}). El espacio de alcanzabilidad es el conjunto de marcados alcanzables desde \mathbf{m}_0 , y se denota como $CA(\mathcal{N}, \mathbf{m}_0)$.

Dada una RdP $\langle P, T, \mathbf{Pre}, \mathbf{Post}, \mathbf{m}_0 \rangle$, con matriz de incidencia $\mathbf{C} = \mathbf{Post} - \mathbf{Pre}$, si \mathbf{m} es alcanzable por el disparo de una secuencia $\sigma \in T^*$, es decir, si $\mathbf{m}_0 \xrightarrow{\sigma} \mathbf{m}$, entonces: $\mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \sigma$ donde $\sigma[t]$ es el número de disparos de t en σ (el contador de disparos). A esta ecuación, con variables en los *naturales*, se le conoce como *ecuación fundamental* o *ecuación de estado*.

Desgraciadamente, aun interpretándola en los naturales, satisfacer la ecuación de estado es condición necesaria pero no suficiente, en general, para la alcanzabilidad; los marcados que son solución de la ecuación pero no son alcanzables se denominan *espurios*. Por ejemplo, en el sistema dinámico de la Figura 1, el marcado $[0, 1, 0, 0]$ verifica la ecuación de estado con vector de disparos $[0, 1, 1]$, pero no es alcanzable ya que no existe una secuencia disparable formada por t_2 y t_3 . Debido a la existencia de marcados espurios, resulta que a través de la ecuación de estados sólo se podrán semi-decidir las propiedades analizadas. No obstante, existen subclases de redes en las que no hay marcados espurios (o estos no son singularmente problemáticos).

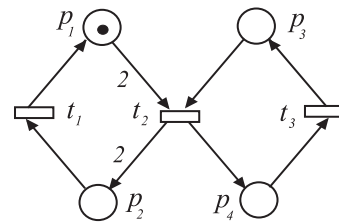


Figura 1. El marcado $[0, 1, 0, 0]$ es una solución espuria de la ecuación de estado.

Los anuladores derecho e izquierdo de la matriz de incidencia \mathbf{C} son dos objetos estructurales que reciben el nombre de T- y P-flujos, respectivamente. Si son no negativos se denominan semiflujos. Si existe un P-semiflujo estrictamente positivo, es decir $\mathbf{y} > \mathbf{0}$ tal que $\mathbf{y} \cdot \mathbf{C} = \mathbf{0}$, se dice que la red es conservativa, y si existe un T-semiflujo positivo, es decir $\mathbf{x} > \mathbf{0}$ tal que $\mathbf{C} \cdot \mathbf{x} = \mathbf{0}$, se dice que la red es consistente.

Los flujos son importantes porque inducen relaciones de invariancia enraizadas en la estructura que son útiles para razonar acerca del comportamiento del sistema. Por ejemplo, si $\mathbf{y} \geq \mathbf{0}$ verifica $\mathbf{y} \cdot \mathbf{C} = \mathbf{0}$, entonces *para todo marcado inicial* \mathbf{m}_0 , todo marcado alcanzable \mathbf{m} verifica: $\mathbf{y} \cdot \mathbf{m} = \mathbf{y} \cdot \mathbf{m}_0 + \mathbf{y} \cdot \mathbf{C} \cdot \sigma = \mathbf{y} \cdot \mathbf{m}_0$. Es decir, los P-semiflujos corresponden a leyes de conservación de marcas. Análogamente, si $\mathbf{x} \geq \mathbf{0}$ verifica $\mathbf{C} \cdot \mathbf{x} = \mathbf{0}$, se tiene que $\mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \mathbf{x} = \mathbf{m}_0$. Es

decir, los T-semiflujos corresponden a *potenciales* secuencias repetitivas. Los semiflujos cuyo soporte no contiene otros semiflujos, y tal que el mínimo común denominador de sus componentes es 1 se llaman semiflujos mínimos. Por ejemplo, el sistema de la Figura 2 tiene 2 T-semiflujos mínimos $[1\ 1\ 0\ 0\ 1\ 1]$, y $[0\ 0\ 1\ 1\ 0\ 0]$, pero no se pueden disparar aislados.

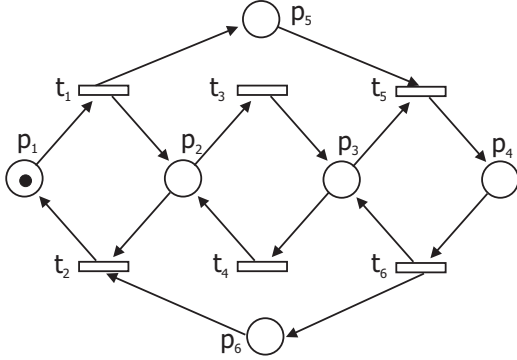


Figura 2. El T-semiflujo mínimo $[1\ 1\ 0\ 0\ 1\ 1]$ no se puede disparar aisladamente.

Una transición t es *viva* si desde todo marcado alcanzable puede llegar a dispararse, es decir, para todo $\mathbf{m} \in \text{CA}(\mathcal{N}, \mathbf{m}_0)$ existe un marcado \mathbf{m}' alcanzable desde \mathbf{m} en el que t está sensibilizada. Un sistema es vivo si todas sus transiciones lo son. La vivacidad asegura que todas las acciones se pueden volver a realizar. Una red es *estructuralmente viva* si existe al menos un marcado inicial (finito) que hace el sistema vivo. Dicho de otro modo, una red estructuralmente no viva no admite ningún marcado que la haga viva, de donde se infiere una patología estructural.

Un sistema es *libre de bloqueo* si en todo marcado alcanzable hay alguna transición sensibilizada (algo análogo a que el motor “no se gripa”). Claramente la ausencia de bloqueo es una condición necesaria para vivacidad.

Un sistema es *reversible* si desde todo marcado alcanzable es posible volver al marcado inicial (y por tanto a cualquier otro marcado alcanzable).

Un sistema es *limitado* si todos sus lugares son limitados, es decir, para cualquier marcado alcanzable el número de marcas en cada lugar está acotado. Una red es *estructuralmente limitada* si para todo marcado inicial el sistema es limitado. Utilizando el Lemma de Farkas (Murty 1983) se puede demostrar que es equivalente que una red sea estructuralmente limitada o que exista $\mathbf{y} > \mathbf{0}$ tal que $\mathbf{y} \cdot \mathbf{C} \leq \mathbf{0}$.

Dos transiciones están en conflicto si el disparo de una limita o impide el disparo de la otra. Más formalmente, $t, t' \in T$ están en conflicto si existen $k, k' \in \mathbb{N}$ tal que $\mathbf{m} \geq k \cdot \mathbf{Pre}[P, t]$ y $\mathbf{m} \geq k' \cdot \mathbf{Pre}[P, t']$, pero $\mathbf{m} \not\geq k \cdot \mathbf{Pre}[P, t] +$

$k' \cdot \mathbf{Pre}[P, t']$. Para que esto ocurra, es necesario que los predecesores de t y t' no sean disjuntos ($\bullet t \cap \bullet t' \neq \emptyset$), y en ese caso decimos que t y t' están en *conflicto estructural*.

La relación de conflicto estructural no es transitiva. Por ejemplo, en la Figura 3 las transiciones t_1 y t_2 , y las transiciones t_2 y t_3 están en relación de conflicto estructural, pero t_1 y t_3 no están en relación de conflicto estructural, dado que no comparten ningún lugar de entrada. Se define la relación de *conflictos acoplados* como su clausura transitiva, y es una relación de equivalencia. El conjunto de todas estas clases de equivalencia se denota \mathcal{A} (*Conjunto de Conjuntos de Conflictos Acoplados*). Por ejemplo, en la Figura 3 t_1 , t_2 y t_3 están en relación de conflicto acoplado, $\mathcal{A} = \{\{t_1, t_2, t_3\}\}$. Dada una red, si $\mathbf{Pre}[P, t] = \mathbf{Pre}[P, t'] \neq \mathbf{0}$, se dice que t y t' están en relación de *conflicto igualado*. Esta es una relación de equivalencia, y establece una partición en el conjunto de las transiciones, \mathcal{I} (*Conjunto de Conjuntos de Conflictos Igualados*). En la Figura 3 ninguna transición está en relación de conflicto igualado con otra, $\mathcal{I} = \{\{t_1\}, \{t_2\}, \{t_3\}\}$. En cambio, en la Figura 4, t_1 y t_2 están en relación de conflicto igualado, $\mathcal{I} = \{\{t_1, t_2\}\}$.

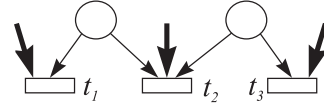


Figura 3. Relación de conflicto acoplado.

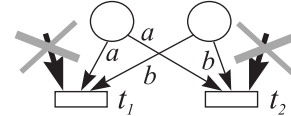


Figura 4. Relación de conflicto igualado.

Muchas subclases de redes se definen restringiendo o eliminando la mezcla de elecciones y sincronizaciones. Por ejemplo, en las redes *Sin Elección* cada lugar tiene como mucho una transición de salida ($\forall p \ |p^\bullet| \leq 1$), y en las redes *Sin Sincronización* cada transición tiene como mucho una entrada ($\forall t \ |\bullet t| \leq 1$). Las redes de *Conflictos Igualados* son redes en las que si dos transiciones tienen un lugar de entrada común, todas sus entradas son exactamente iguales ($\bullet t \cap \bullet t' \neq \emptyset \Rightarrow \mathbf{Pre}[P, t] = \mathbf{Pre}[P, t']$); si además todos los arcos tienen peso 1, reciben el nombre de redes de *Elección Libre* (porque no existen condicionantes externos a la resolución *local* y *libre* del conflicto). Los *grafos marcados* son redes en las que cada lugar tiene una única transición de entrada y una de salida ($\forall p \ |p^\bullet| = |p^\circ| = 1$), y los arcos tienen peso 1. Otra clase importante son las redes *mono-T-semiflujo*, que son redes conservativas, consistentes con un único T-semiflujo mínimo \mathbf{x} (es decir,

para cada transición t_i hay un único T-semiflujo tal que $\mathbf{x}[t_i] = 1$). En este tipo de redes la relación entre el número de disparos de las transiciones, cuando el número de disparos tiende a infinito, queda completamente definida por la estructura de la red. La Figura 5 representa la relación entre algunas de estas clases.

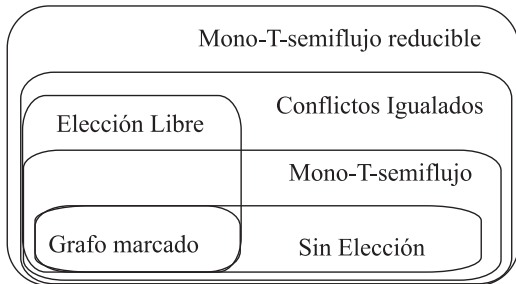


Figura 5. Algunas relaciones entre subclases de redes.

Ligado a las sincronizaciones, aparece el concepto de *configuración*: una configuración asigna a cada transición uno de sus lugares de entrada. Por lo tanto, el número de configuraciones posibles está acotado por $\prod_{t \in T} |\bullet t|$. El espacio de alcanzabilidad puede dividirse en *regiones* en función de las configuraciones “activas”.

Definición 3. Sea \mathcal{N} una red de Petri. Una *configuración* \mathcal{C}_k de \mathcal{N} es una aplicación de T en P tal que $\mathcal{C}_k(t) \in \bullet t$, es decir, elige para cada transición uno de sus lugares de entrada. Si \mathcal{C}_k es una configuración, \mathcal{R}_k denota la región correspondiente:

$$\mathcal{R}_k = \left\{ \mathbf{m} \in (\mathbb{R}_{\geq 0})^{|P|} \mid \forall t \in T, \frac{\mathbf{m}[\mathcal{C}_k(t)]}{\mathbf{Pre}[\mathcal{C}_k(t), t]} = \min_{p \in \bullet t} \left\{ \frac{\mathbf{m}[p]}{\mathbf{Pre}[p, t]} \right\} \right\}$$

Estas regiones son poliedros y, exceptuando la frontera, son disjuntos.

Por ejemplo, la red de la Figura 2 tiene 4 configuraciones, según si se asocia a p_2 ó p_6 a t_2 y p_3 ó p_5 a t_5 . Utilizaremos (p_i, t_j) para señalar que el lugar $p_i \in \bullet t_j$ está asociado a t_j en la configuración.

Puede ocurrir que en una sincronización uno de los lugares de entrada en realidad nunca limite el disparo. Es decir, que existan regiones vacías para el sistema con ese marcado inicial. Dado un sistema, si un lugar nunca es el único que limita el disparo de sus transiciones de salida, se dice que es un lugar implícito, y puede eliminarse sin modificar el comportamiento del sistema.

Definición 4. Sea una red $\langle P \cup \{p\}, T, \mathbf{Pre}, \mathbf{Post} \rangle$ y \mathbf{m}_0 un marcado inicial. El lugar p es *implícito* para \mathbf{m}_0 si el conjunto de secuencias disparables en $\langle P, T, \mathbf{Pre}[P, T], \mathbf{Post}[P, T], \mathbf{m}_0[P] \rangle$ coincide con las disparables en el sistema original.

Ejemplo 1. Sea el modelo de un sistema flexible de manufactura de la Figura 6. Las piezas de tipo A se procesan primero en la máquina M1 y a continuación en la máquina M2, mientras que las piezas de tipo B se procesan primero en M2 y luego en M1. Los productos intermedios se almacenan en B_1A y B_1B, y los productos finales en B_2A y B_2B, respectivamente. La máquina M3 toma una parte A y una parte B y ensambla el producto final, que se almacena temporalmente en B_3. En B_3 hay espacio para almacenar hasta 10 productos. Puede haber hasta 10 piezas de tipo A (10 piezas de tipo B) que se encuentren en B_1A (B_1B), o siendo procesadas por M1 o M2. Las piezas se mueven en pallets, y se dispone de 20 pallets de tipo A y 15 pallets de tipo B. En suma, en este sencillo caso, se tienen tres máquinas secuenciales (autómatas finitos), cuatro almacenes intermedios y uno terminal. Las relaciones existentes son de exclusión mutua (actividades de M1 y M2) o de producción-consumo.

Con este marcado inicial todas las regiones contienen marcados. Sin embargo, si $\mathbf{m}_0[\text{Max_B}] > 15$, Max_B nunca restringe el disparo de S.M2.B, y por tanto existen regiones vacías. Como esta es la única transición de salida de Max_B, está claro que éste es un lugar implícito para este marcado inicial.

2.2 Red de Petri continua.

La estructura de una red de Petri continua es la misma que la de una red de Petri discreta. Lo único que cambia es la regla de disparo (y por lo tanto el marcado), que se relaja permitiendo disparar las transiciones en cantidades no enteras, pero positivas.

Definición 5. Un sistema continuo $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ está compuesto por una red de Petri “clásica” $\mathcal{N} = \langle P, T, \mathbf{Pre}, \mathbf{Post} \rangle$, y un *marcado* $\mathbf{m}_0 \in (\mathbb{R}_{\geq 0})^{|P|}$, es decir definido en los reales no negativos.

El concepto de *sensibilización* de una transición se extiende al caso continuo de forma inmediata: puesto que las transiciones *se pueden disparar en cualquier cantidad, no necesariamente entera*, para que se pueda disparar basta con que todos los lugares de entrada de la transición estén marcados, i.e., t está *sensibilizada* en \mathbf{m} si $\forall p \in \bullet t, \mathbf{m}[p] > 0$. Al igual que en los sistemas discretos, el *grado de sensibilización* mide la máxima cantidad en la que una transición se puede disparar en un paso, y viene dado por $\min_{p \in \bullet t} \{ \mathbf{m}[p] / \mathbf{Pre}[p, t] \}$. El disparo de t en una cantidad $\alpha \leq \min_{p \in \bullet t} \{ \mathbf{m}[p] / \mathbf{Pre}[p, t] \}$ lleva a un nuevo marcado $\mathbf{m}' = \mathbf{m} + \alpha \cdot \mathbf{C}[P, t]$.

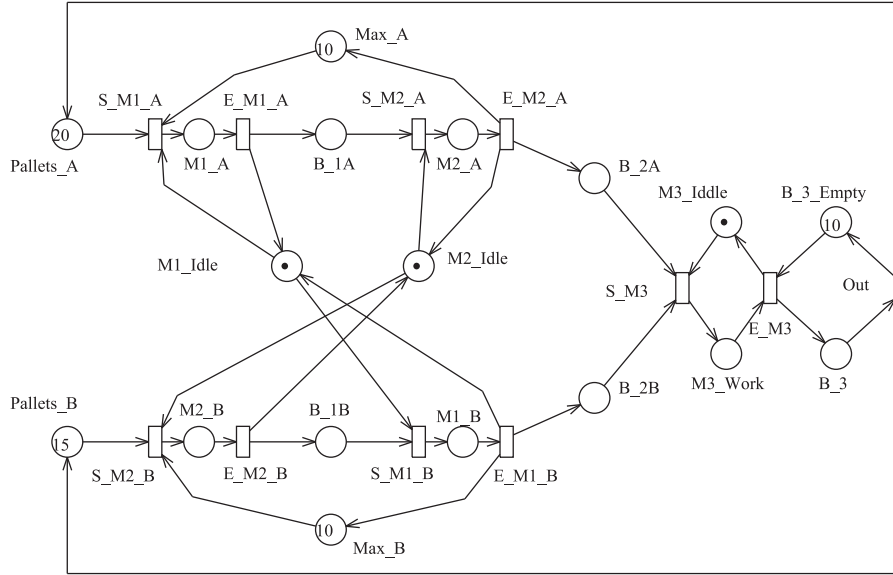


Figura 6. Modelo red de Petri de un sistema de manufactura.

Las propiedades meramente estructurales de las RdP discretas se transponen de forma inmediata a las continuas, por ejemplo, P- y T-(semi)flujos, consistencia, conservatividad, conexión del grafo, trampas, conjuntos de conflictos igualados \mathcal{I} y acoplados \mathcal{A} ...

En las propiedades comportamentales la extensión no es tan simple como cabría esperar en un primer momento. Por ejemplo, el disparo de transiciones en redes continuas se puede hacer en cantidades muy pequeñas, de forma que una red puede estar “casi” bloqueada sin llegar nunca a pararse completamente. Por ejemplo, en el sistema de la Figura 7, cada vez que se disparan t_1 y t_2 en la máxima cantidad posible, el marcado de p_1 se reduce a la mitad. Por lo tanto, se pueden encontrar un marcado tan pequeño como se desee, pero nunca se llega a cero (como en la clásica paradoja de Zenón).

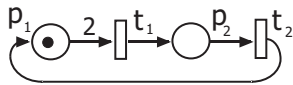


Figura 7. El sistema es estructuralmente no vivo como discreto, pero como continuo nunca se bloquea completamente con secuencias finitas de disparos.

Para mantener la coherencia con el caso discreto, parece razonable el considerar los marcados límite, i.e. a través de secuencias infinitas, como alcanzables. En (Recalde *et al.* 1999) se introdujo el concepto de alcanzabilidad en el límite, estudiado más en profundidad en (Júlvez *et al.* 2003).

Definición 6. Sea $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ un sistema RdP continuo. Se define $CA(\mathcal{N}, \mathbf{m}_0)$ como el Conjunto Alcanzable de marcados que se obtiene dispa-

rando una secuencia finita, y $\lim\text{-CA}(\mathcal{N}, \mathbf{m}_0)$ como su extensión a secuencias infinitas. Es decir, $\lim\text{-CA}(\mathcal{N}, \mathbf{m}_0)$ se forma añadiendo a $CA(\mathcal{N}, \mathbf{m}_0)$ los marcados que son límite de secuencias de marcados pertenecientes a $CA(\mathcal{N}, \mathbf{m}_0)$.

A partir del concepto de alcanzabilidad se extienden de forma inmediata las propiedades de vivacidad y ausencia de bloqueo definidas en el caso discreto.

Definición 7. Un sistema $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ es lim-vivo si para todo $\mathbf{m} \in \lim\text{-CA}(\mathcal{N}, \mathbf{m}_0)$ y para toda transición $t \in T$ existe un marcado sucesor $\mathbf{m}' \in \lim\text{-CA}(\mathcal{N}, \mathbf{m})$ tal que $\mathbf{m}'[\bullet t] > 0$.

Un sistema $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ es lim-libre-de-bloqueo si para todo marcado $\mathbf{m} \in \lim\text{-CA}(\mathcal{N}, \mathbf{m}_0)$ existe una transición $t \in T$ tal que $\mathbf{m}[\bullet t] > 0$.

Un sistema $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ es lim-limitado si existe $\mathbf{k} \in \mathbb{R}^{|P|}$ tal que todo marcado $\mathbf{m} \in \lim\text{-CA}(\mathcal{N}, \mathbf{m}_0)$ verifica $\mathbf{m} \leq \mathbf{k}$.

Obsérvese que aunque en la definición de limitación se ha empleado $\lim\text{-CA}$, se podía haber definido igual restringiéndolo a secuencias finitas, ya que cualquier cota válida para los marcados alcanzables con secuencias finitas será válida para los marcados límite.

El sistema de la Figura 7 permite también observar que en redes de Petri continuas es posible tener secuencias infinitas donde todos los marcados sean diferentes y estén limitados. Esto es algo que no ocurre con redes de Petri discretas donde no existen secuencias infinitas en las que los marcados estén acotados y no se repitan.

En redes de Petri discretas es posible que un sistema sea limitado, pero no estructuralmente limitado (ver Figura 4.3 de (Silva 2002)). Esto se debe a que pueden existir vectores de disparo que hagan crecer el marcado, pero que no sean nunca disparables con ese marcado inicial concreto. En el caso continuo, si todas las transiciones son disparables desde algún marcado alcanzable desde \mathbf{m}_0 , las dos propiedades son equivalentes, ya que se puede garantizar que hay un marcado alcanzable desde el que se puede disparar ese vector que hace crecer el marcado. Dicho de otro modo, el sistema continuo $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ es limitado si y sólo si la red \mathcal{N} es estructuralmente limitada.

Teorema 1. (Recalde *et al.* 1999) Sea $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ un sistema continuo en el que toda transición es disparable al menos una vez. Es equivalente:

- $\langle \mathcal{N}, \mathbf{m}_0 \rangle$ es limitado.
- \mathcal{N} es estructuralmente limitada.
- existe $\mathbf{y} > \mathbf{0}$ tal que $\mathbf{y} \cdot \mathbf{C} \leq \mathbf{0}$.
- Para todo $\mathbf{x} > \mathbf{0}$ tal que $\mathbf{C} \cdot \mathbf{x} \geq \mathbf{0}$, $\mathbf{C} \cdot \mathbf{x} = \mathbf{0}$.

Por lo tanto, si un sistema continuo es limitado, también lo es como discreto, aunque el contrario no siempre es cierto. Este teorema estructural, en sus dos últimas afirmaciones da visiones alternativas (basadas en lugares y transiciones, respectivamente) cuya equivalencia reposa en los conocidos teoremas de alternativas (Murty 1983).

Igual que ocurre en el caso discreto, en redes de Petri continuas no se verifican propiedades generales de monotonía. Por ejemplo, aumentar el marcado de un lugar puede hacer que una red viva deje de serlo. El sistema de la Figura 8 es vivo, pero si $\mathbf{m}_0[p_2] = 2$, disparando t_1 la red queda bloqueada. Sin embargo, sí que se cumplen propiedades de homotecia o de escalabilidad del marcado. Es decir, las propiedades del sistema continuo no varían si su marcado se multiplica por una constante, algo que no se cumple en general en el caso discreto. Por ejemplo, la red de la Figura 9(b) es viva como discreta con este marcado, pero no lo es si el marcado inicial se multiplica por 2. Esta propiedad se puede interpretar también como una condición que es lógico exigir al sistema para que su aproximación continua tenga sentido.

Otra propiedad de las redes de Petri continuas que no se verifica en el caso discreto es la *convexidad* de los espacios de alcanzabilidad (tanto usando secuencias finitas como infinitas).

Propiedad 1. Sea el sistema continuo $\langle \mathcal{N}, \mathbf{m}_0 \rangle$. Para todo \mathbf{m}_1 y $\mathbf{m}_2 \in \text{CA}(\mathcal{N}, \mathbf{m}_0)$ ($\mathbf{m}_1, \mathbf{m}_2 \in \text{lim-CA}(\mathcal{N}, \mathbf{m}_0)$), cualquier marcado $\mathbf{m} = \alpha \mathbf{m}_1 + (1 - \alpha) \mathbf{m}_2$, con $\alpha \in [0, 1]$ pertenece a $\text{CA}(\mathcal{N}, \mathbf{m}_0)$ (respectivamente, pertenece a $\text{lim-CA}(\mathcal{N}, \mathbf{m}_0)$).

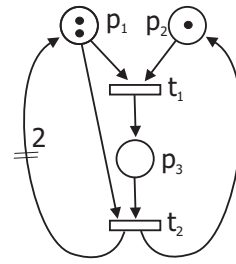


Figura 8. El sistema es vivo con este marcado, pero deja de serlo si se aumenta el marcado de p_2 a 2 marcas.

Sin embargo, hay que tener presente que la continuización es una relajación y hay propiedades que no tiene sentido estudiar en el modelo continuo. Por ejemplo, la exclusión mutua en el acceso a un recurso está basada en el carácter discreto. Además, en el caso continuo, algunas propiedades pueden confundirse y perder la distinción entre ellas. Por ejemplo, en condiciones bastante generales lim-vivacidad y lim-reversibilidad (es decir, para todo marcado $\mathbf{m} \in \text{lim-CA}(\mathcal{N}, \mathbf{m}_0)$ se verifica que $\mathbf{m}_0 \in \text{lim-CA}(\mathcal{N}, \mathbf{m})$) son equivalentes (Júlvez *et al.* 2003).

En (Júlvez *et al.* 2003) se estudia también la decidibilidad del problema de alcanzabilidad, demostrándose que, como ocurre en el caso discreto, es decidible (tanto para el caso de secuencias de disparo finitas como infinitas). En (Recalde *et al.* 2007) se demuestra que propiedades como vivacidad y ausencia de bloqueos son también decidibles. En general, el estudio de estas propiedades resulta computacionalmente mucho más sencillo que en el caso de redes discretas, ya que problemas de programación entera se transforman en problemas de programación lineal.

3. RED DE PETRI CONTINUA AUTÓNOMA: ANÁLISIS.

Como en el caso discreto (Silva *et al.* 1998), la ecuación de estado también puede usarse en el análisis de redes de Petri continuas, simplemente eliminando las restricciones de integralidad. Idénticamente, la ecuación de estado puede tener soluciones espurias. Sin embargo, bajo condiciones muy generales y frecuentes, en la práctica si la red es consistente ($\exists \mathbf{x} > \mathbf{0}$ tal que $\mathbf{C} \cdot \mathbf{x} = \mathbf{0}$) y todas las transiciones se pueden disparar, la ecuación de estado ¡no tiene soluciones espurias!, es decir, el conjunto $\text{lim-CA}(\mathcal{N}, \mathbf{m}_0)$ coincide con las soluciones de la ecuación de estado. Los detalles sobre la relación entre ambos conjuntos y otros con diferentes matices se pueden ver en (Recalde *et al.* 1999, Júlvez *et al.* 2003).

Teorema 2. (Recalde *et al.* 1999) Si $(\mathcal{N}, \mathbf{m}_0)$ es consistente, y toda transición se puede llegar a disparar,

$$\lim\text{-CA}(\mathcal{N}, \mathbf{m}_0) = \{\mathbf{m} \geq \mathbf{0} \mid \mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \boldsymbol{\sigma}, \\ \boldsymbol{\sigma} \geq \mathbf{0}\}$$

El espacio de alcanzabilidad es por tanto un politopo, y para decidir sobre la alcanzabilidad basta con resolver un problema de programación lineal, es decir, se decide con complejidad polinomial en tiempo.

Lamentablemente, la vivacidad de la red continua no es, en general, ni condición necesaria ni suficiente de vivacidad de la red discreta. Por ejemplo, la red de la izquierda en la Figura 9 es viva como continua, pero no como discreta, y la red de la derecha es viva como discreta, pero no es viva como continua. Sin embargo, en el primer caso el problema es que en la red discreta el marcado no es adecuado, y si se cambia a [20] la red es viva como discreta. Parece lógico por tanto que la propiedad que hay que buscar es vivacidad estructural, es decir, que la estructura de la red sea “correcta” (Recalde *et al.* 1999). Sin embargo, en el ejemplo de la derecha la red continua no es viva con ningún marcado inicial (tampoco si sólo se consideran secuencias de disparo finitas).

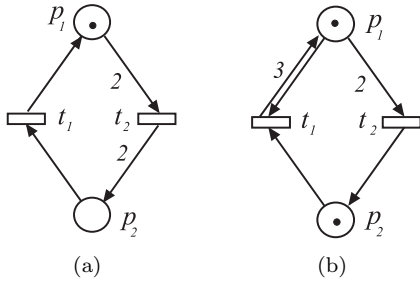


Figura 9. El sistema de la izquierda es vivo como continuo, pero no como discreto. El de la derecha es vivo como discreto, pero no como continuo con cualquier marcado inicial.

Teorema 3. (Recalde *et al.* 1999) Sea $(\mathcal{N}, \mathbf{m}_0)$ un sistema lim-vivo y limitado. Entonces, \mathcal{N} es estructuralmente viva y estructuralmente limitada como red discreta

Los teoremas del rango demostrados para el caso discreto enuncian condiciones necesarias o/y suficientes para que una red sea estructuralmente viva y limitada basadas en la estructura de la red (Silva *et al.* 1998). Utilizando el resultado del Teorema 3, las condiciones necesarias se extienden al caso continuo.

Teorema 4. (Recalde *et al.* 1999) Sea $(\mathcal{N}, \mathbf{m}_0)$ un sistema lim-vivo y limitado (vivo y limitado en el

caso discreto). Entonces, \mathcal{N} es consistente ($\exists \mathbf{x} > \mathbf{0}$ tal que $\mathbf{C} \cdot \mathbf{x} = \mathbf{0}$), conservativa ($\exists \mathbf{y} > \mathbf{0}$ tal que $\mathbf{y} \cdot \mathbf{C} = \mathbf{0}$) y $\text{rango}(\mathbf{C}) \leq |\mathcal{I}| - 1$.

La condición del rango en cierta forma acota la relación admisible entre las elecciones y las sincronizaciones. Si $\text{rango}(\mathbf{C}) > |\mathcal{I}| - 1$, existen decisiones libres que sin embargo no pueden ser tomadas sin tener en cuenta el resto, en el sentido de que elecciones incorrectas pueden llegar a bloquear el sistema o parte de él. Por ejemplo, en la red de la Figura 10(a), hay un conflicto libre entre t_1 y t_2 , pero la sincronización posterior hace que si se elige la misma transición demasiadas veces se llegue a bloquear el sistema. El hecho de que en secuencias infinitas t_1 y t_2 se deban disparar en una cierta proporción se refleja en que hay un único anulador derecho de la matriz de incidencia, es decir $\text{rango}(C) = 2$. Como $\mathcal{I} = \{t_1, t_2\}, \{t_3\}$, aplicando el teorema la red no puede ser lim-viva con ningún marcado inicial. En cambio, la red de la Figura 10(b) cumple las condiciones necesarias de vivacidad (estructural) ya que $\text{rango}(\mathbf{C}) = 2$ y $|\mathcal{I}| = 3$.

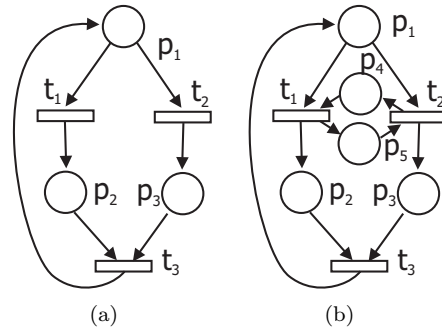


Figura 10. Redes consistentes y conservativas con (a) $|\mathcal{I}| = 2$, y $\text{rango}(C) = 2$, (b) $|\mathcal{I}| = 3$, y $\text{rango}(C) = 2$.

Para las redes mono-T-semiflujo, la condición del Teorema 4 puede mejorarse.

Teorema 5. (Júlvez *et al.* 2006) Sea \mathcal{N} una red mono-T-semiflujo. Si \mathcal{N} es estructuralmente lim-viva, entonces toda transición tiene al menos un lugar de entrada que no tiene otra transición de salida.

La red de la Figura 11 es mono-T-semiflujo, consistente y conservativa y cumple que $\text{rango}(\mathbf{C}) = 5 = |\mathcal{I}| - 1$, pero las transiciones t_1, t_2 y t_3 no tienen un lugar de entrada que no sea entrada de otra transición, luego no es estructuralmente lim-viva.

Las condiciones suficientes también pueden extenderse, y para el caso continuo incluso puede garantizarse lim-vivacidad para un conjunto de

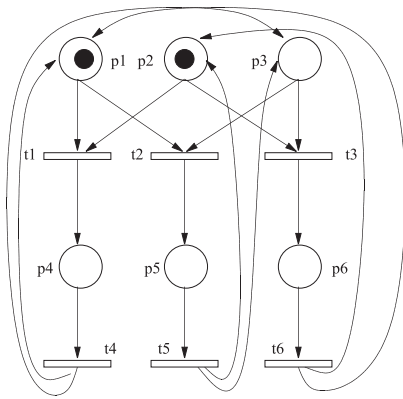


Figura 11. Red mono-T-semiflujo, consistente, conservativa y $\text{rango}(\mathbf{C}) = 5 = |\mathcal{I}| - 1$, pero no estructuralmente lim-viva.

marcados, algo que no se puede hacer en el caso discreto.

Teorema 6. (Recalde *et al.* 1999) Sea \mathcal{N} consistente, conservativa y $\text{rango}(\mathbf{C}) = |\mathcal{A}| - 1$. Entonces \mathcal{N} es estructuralmente viva como discreta y estructuralmente lim-viva como continua. Además, cualquier marcado que no deje vacío ningún P-semiflujo hace el sistema lim-vivo.

Esta condición suficiente se obtiene imponiendo condiciones más duras en la resolución de conflictos, transformando cada conflicto acoplado en un conflicto igualado (manteniendo la matriz de incidencia). Así, si la red que se obtiene es viva, también lo será la red original. La red que se obtiene es de Conflictos Igualados, para la cual el teorema del rango es condición necesaria y suficiente (Recalde *et al.* 1999).

Bajo las condiciones del Teorema 2, una condición necesaria y suficiente para que un sistema sea lim-libre-de-bloqueo es que no exista solución de la ecuación de estado sin ninguna transición sensibilizada. Por ejemplo, la red de la Figura 10(b) con marcado inicial $\mathbf{m}_0 = [1, 0, 0, 1, 0]$ es lim-libre-de-bloqueo si y sólo si el siguiente sistema no tiene solución:

$$\begin{aligned} \mathbf{m} &= \mathbf{m}_0 + \mathbf{C} \cdot \boldsymbol{\sigma} \geq \mathbf{0} \\ \min(\mathbf{m}[p_1], \mathbf{m}[p_4]) &= 0 \quad (t_1 \text{ no sensibilizada}) \\ \min(\mathbf{m}[p_1], \mathbf{m}[p_5]) &= 0 \quad (t_2 \text{ no sensibilizada}) \\ \min(\mathbf{m}[p_2], \mathbf{m}[p_3]) &= 0 \quad (t_3 \text{ no sensibilizada}) \end{aligned}$$

Para resolver este problema se pueden definir variables booleanas $\mu(i) = 0 \Leftrightarrow \mathbf{m}[p_i] = 0$. El sistema se bloquea si y sólo si $(\mu(1) + \mu(4)) \cdot (\mu(1) + \mu(5)) \cdot (\mu(2) + \mu(3)) = 1$. Como $\mathbf{m}[p_4] + \mathbf{m}[p_5] = 1$ (P-semiflujo), se tiene que $\mu(4) + \mu(5) = 1$, y se puede simplificar la expresión booleana que caracteriza el bloqueo: $\mu(1) \cdot (\mu(2) + \mu(3)) = 1$. Sean

$$\Sigma_{12} = \begin{cases} \mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \boldsymbol{\sigma} \geq \mathbf{0}, \boldsymbol{\sigma} \geq \mathbf{0} \\ \mathbf{m}[p_1] = \mathbf{m}[p_2] = 0 \end{cases}$$

$$\Sigma_{13} = \begin{cases} \mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \boldsymbol{\sigma} \geq \mathbf{0}, \boldsymbol{\sigma} \geq \mathbf{0} \\ \mathbf{m}[p_1] = \mathbf{m}[p_3] = 0 \end{cases}$$

El sistema es lim-libre-de-bloqueo si y sólo si Σ_{12} y Σ_{13} no tienen solución. La ausencia de solución es un problema lineal de complejidad polinómica, pero teóricamente el número de problemas es exponencial.

La condición de ausencia de bloqueo se puede interpretar en términos de regiones: se trata de verificar que en ninguna región exista un marcado de bloqueo.

4. TEMPORIZACIÓN DE MODELOS

4.1 Marco conceptual

Al igual que ocurre con las redes discretas, es posible añadir distintas interpretaciones temporales al disparo de las transiciones. Una forma habitual de introducir el tiempo en las redes discretas (modelo *markoviano* o *amnésico*) consiste en suponer que todas las transiciones tienen asociado un tiempo con función de distribución de probabilidad exponencial (Molloy 1982, Ajmone Marsan *et al.* 1995).

Una primera aproximación continua a la interpretación estocástica discreta consiste en asociar a cada transición una velocidad media. Esta aproximación de primer orden transforma un *modelo estocástico discreto* en un *modelo determinista continuo* (Silva y Recalde 2002, Recalde y Silva 2001). En particular es un sistema de ecuaciones diferenciales no lineales.

Como el marcado es el estado global distribuido del sistema continuo, si en la ecuación de estado hacemos explícita la dependencia del tiempo, se obtiene la ecuación $\mathbf{m}(\tau) = \mathbf{m}(0) + \mathbf{C} \cdot \boldsymbol{\sigma}(\tau)$, $\boldsymbol{\sigma}(\tau) \geq \mathbf{0}$, que representa la evolución temporal. Derivando se llega a

$$\dot{\mathbf{m}}(\tau) = \mathbf{C} \cdot \dot{\boldsymbol{\sigma}}(\tau) \stackrel{\text{def}}{=} \mathbf{C} \cdot \mathbf{f}(\tau)$$

donde $\mathbf{f}(\tau)$ es el flujo instantáneo de disparo de las transiciones (i.e., disparos por unidad de tiempo). Si la red tiene P-flujos (y en particular, si es conservativa), las filas de \mathbf{C} no son linealmente independientes, es decir, el orden del sistema es menor que el número de lugares.

Si la red autónoma es lím-viva (respectivamente lím-limitada), la red temporizada también lo será. Sin embargo, es posible que la red autónoma no sea lím-viva (o lím-limitada), y la red temporizada sí lo sea, ya que los marcados que se alcanzan en la red temporizada son sólo una parte de los que se alcanzan en la red autónoma.

Las dos semánticas habitualmente utilizadas en sistemas discretos, finitos o infinitos servidores, pueden aplicarse también en el caso continuo.

Bajo la semántica de *finitos servidores* (o *velocidad constante*) el flujo de cada transición tiene una cota superior (el número de servidores por la velocidad de cada uno). Si todos los lugares de entrada de una transición están marcados, la transición se dispara con esta velocidad. Si no es así, la velocidad está además limitada por el flujo que llega a los lugares de entrada no marcados. Bajo la semántica de *infinitos servidores*, o *velocidad variable*, el flujo de una transición es el producto de una constante (la velocidad de un servidor) por el grado de sensibilización de la transición (número de servidores activos). Es decir, $\mathbf{f}[t] = \lambda[t] \cdot \min_{p \in \bullet t} \{ \mathbf{m}[p] / \mathbf{Pre}[p, t] \}$.

Interpretando las transiciones como puntos de encuentro de clientes y servidores, parece que la semántica más adecuada depende del número relativo de clientes y servidores que haya en el modelo discreto. Suponiendo que cualitativamente hablando puede haber “muchos” o “pocos” de cada uno, la fluidificación se puede aplicar a los clientes, a los servidores, o a ambos. El Cuadro 1 representa los cuatro casos. Si hay pocos clientes (Pocos-Pocos y Pocos-Muchos en el Cuadro 1), el sistema no está realmente cargado, la fluidificación puede no ser adecuada y las transiciones “deberían” mantenerse como discretas. Si hay muchos clientes y muchos servidores (Muchos-Muchos), el modelo continuo con semántica de “infinitos” servidores (como metáfora de “muchos”) parece lo más razonable. Por otra parte, en el caso de muchos clientes y pocos servidores (Muchos-Pocos) la relajación es adecuada a nivel de los clientes y la semántica de “finitos” servidores (como metáfora de “pocos”) puede proporcionar una buena aproximación. Es importante resaltar que la semántica de finitos servidores corresponde a nivel conceptual a un esquema *híbrido*: la fluidificación se aplica solamente a los clientes, mientras que los servidores se mantienen discretos, contándose con números naturales (y se usan para definir la velocidad máxima de la transición). Sin embargo, la semántica de infinitos servidores realmente relaja tanto a los clientes como a los servidores. En cualquier caso, ambas clases de modelos de sistemas continuos obtenidos son técnicamente híbridos.

En el caso de redes discretas, es clásico el que la semántica de finitos servidores se puede simular utilizando semántica de infinitos servidores (Ajmone Marsan *et al.* 1995). La idea es que los servidores pueden hacerse explícitos en el modelo utilizando lugares conectados con arcos de ida y vuelta con las transiciones. Sin embargo, en la red continua el significado de las “marcas de clientes” y de las “marcas de servidores” es muy diferente. Las dos semánticas corresponden en la práctica dos formas diferentes de aproximar la red discreta.

El modelo de la Figura 12 representa un sistema de producción con dos líneas de trabajo (la de arriba con transiciones impares y la de abajo con transiciones pares). Se ha analizado como red de Petri discreta, utilizando para las transiciones distribuciones de tipo exponencial (modelo markoviano) y semánticas tanto de infinitos servidores como de un solo servidor. En ambos casos, al multiplicar el marcado de la red por k , la productividad/rendimiento (normalizado en el caso de infinitos servidores) se aproxima al que se obtiene cuando la red se considera continua (Cuadro 2). En este caso, con la red continua con semántica de infinitos servidores se obtiene una mejor aproximación del sistema original discreto. Puede observarse que el error se reduce al aumentar el marcado, por ejemplo es el 14% para $k = 1$, el 6% para $k = 3$ y el 2% para $k = 10$. La semántica (básica) de finitos servidores no tiene en cuenta las restricciones que aparecen asociadas al uso de recursos compartidos, que en este caso resultan bastante importantes.

En ambos casos, la evolución del sistema se puede describir como un conjunto de sistemas de ecuaciones diferenciales lineales en conmutación. Dicha conmutación no está asociada a un evento externo, sino que es la evolución del marcado el que la provoca. Para la semántica de finitos servidores, la conmutación ocurre cuando un lugar queda vacío. Para la semántica de infinitos servidores, la conmutación ocurre cuando en una sincronización el lugar de entrada asociado al mínimo cambia. En el caso de redes sin sincronizaciones (es decir, $\forall t \in T | \bullet t | = 1$) el comportamiento de la red bajo semántica de infinitos servidores se corresponde con un sistema lineal. Si el sistema tiene sincronizaciones, en cada región \mathcal{R}_k la evolución del sistema está definida por una ecuación lineal, es decir, $\dot{\mathbf{m}} = \mathbf{C} \cdot \mathbf{\Lambda} \cdot \mathbf{\Pi}_k \cdot \mathbf{m}$, donde $\mathbf{\Lambda}$ es una matriz diagonal $[\lambda[t_1] \cdots \lambda[t_{|T|}]]$ y

$$\mathbf{\Pi}_k[t_i, p_j] = \begin{cases} 1 & \text{si } (p_j, t_i) \text{ pertenece} \\ \mathbf{Pre}[p_j, t_i] & \text{a la configuración } \mathcal{C}_k \\ 0 & \text{en otro caso} \end{cases}$$

Ejemplo 2. La Figura 13 modela un recurso compartido (representado por el marcado de p_6) entre dos procesos secuenciales. La velocidad de las transiciones es $\lambda = [1 \ 2 \ 1 \ 1 \ 0,5]$. El comportamiento de la red discreta es el mismo si se utiliza semántica de finitos o infinitos servidores, ya que los servidores aparecen de forma explícita en el modelo. El comportamiento como continua sin embargo varía sustancialmente dependiendo de la semántica que se utilice.

Semántica de finitos servidores (e “infinitos” clientes). En \mathbf{m}_0 , los lugares de entrada de las transiciones t_1 y t_4 están marcados, y por lo tanto, $\mathbf{f}[t_1] = \mathbf{f}[t_4] = 1$. Las otras transiciones

Tabla 1. Fluidificación de las transiciones (Silva y Recalde 2004)

Clientes	Servidores	Semántica de la transición
Muchos(C)	Muchos(C)	semántica de infinitos servidores
Muchos(C)	Pocos(D)	semántica de finitos servidores
Pocos(D)	Pocos(D)	transiciones discretas
Pocos(D)	Muchos(C)	transiciones discretas

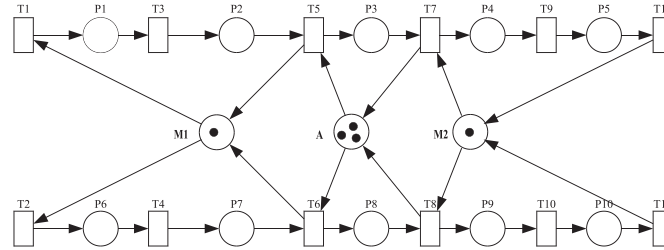


Figura 12. Esquema sencillo de un sistema de producción con dos máquinas (M1 y M2) y un almacén intermedio (A).

Tabla 2. Rendimiento en estado estacionario del sistema de la Figura 12, suponiendo que cada operación necesita 1 u.t. (Para $k=10$ y $k=50$, se han utilizado simulaciones.)

k	Marcados alcanzables	Infinitos servidores		Un servidor
		Rendimiento de cada t_i	Rendimiento/ k	Rendimiento de cada t_i
1	250	0,172	0,172	0,172
2	6.300	0,366	0,183	0,303
3	67.375	0,564	0,188	0,399
5	2.159.136	0,966	0,193	0,528
10	?	1,96	0,196	0,693
50	?	9,97	0,199	0,91
...
Continuo			0,2	1

tienen lugares de entrada desmarcados, y por tanto su flujo depende del flujo de entrada a estos lugares vacíos. Para t_2 , el flujo de entrada a p_3 (el único que está vacío) es 1, por lo tanto $f[t_2] = \min\{\lambda[t_2], 1\} = 1$. Por la misma razón, el flujo de la transición t_3 será 1. Para t_5 , el flujo de entrada a p_5 es 1, así que su flujo está limitado por su velocidad máxima de disparo, que es 0.5. Por lo tanto, las ecuaciones iniciales del sistema serán:

$$\begin{aligned} \dot{\mathbf{m}}[p_1](\tau) &= \mathbf{f}[t_3](\tau) - \mathbf{f}[t_1](\tau) = 0 \\ \dot{\mathbf{m}}[p_2](\tau) &= \mathbf{f}[t_5](\tau) - \mathbf{f}[t_4](\tau) = -0,5 \\ \dot{\mathbf{m}}[p_3](\tau) &= \mathbf{f}[t_1](\tau) - \mathbf{f}[t_2](\tau) = 0 \\ \dot{\mathbf{m}}[p_4](\tau) &= \mathbf{f}[t_2](\tau) - \mathbf{f}[t_3](\tau) = 0 \\ \dot{\mathbf{m}}[p_5](\tau) &= \mathbf{f}[t_4](\tau) - \mathbf{f}[t_5](\tau) = 0,5 \\ \dot{\mathbf{m}}[p_6](\tau) &= \mathbf{f}[t_3] + \mathbf{f}[t_5] - \mathbf{f}[t_2] - \mathbf{f}[t_4] = -0,5 \end{aligned}$$

Estas ecuaciones se verifican hasta $\tau = 2$, cuando $\mathbf{m}[p_6]$ y $\mathbf{m}[p_2]$ se quedan vacíos. El marcado en este momento es $[1 \ 0 \ 0 \ 0 \ 1 \ 0]$. Ahora t_1 y t_5 tienen todos sus lugares de entrada marcados, así que $\mathbf{f}[t_1] = 1$ y $\mathbf{f}[t_5] = 0,5$. Las transiciones t_2 y t_4 tienen un lugar de entrada vacío, pero además este lugar es el mismo, y por lo tanto tiene que especificarse cómo va a repartirse el flujo que llegue a p_6 entre las dos transiciones. Por ejemplo, supongamos que se decide que el

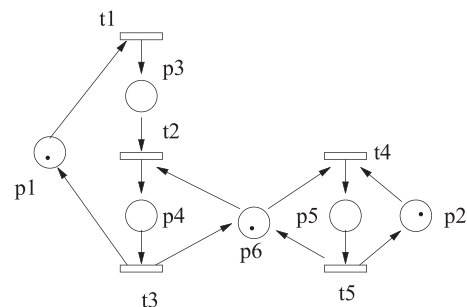


Figura 13. Red de Petri utilizada en el Ejemplo 2.

flujo de t_2 sea igual al de t_4 . Además, el flujo de salida de los lugares vacíos está limitado por los flujos de entrada. Esto lleva al siguiente sistema de ecuaciones:

$$\begin{aligned} \max \quad & \mathbf{f}[t_2] \\ \text{s.t.} \quad & \mathbf{f}[t_1] = 1 \quad (\text{entradas no vacías}) \\ & \mathbf{f}[t_5] = 0,5 \quad (\text{entradas no vacías}) \\ & \mathbf{f}[t_2] = \mathbf{f}[t_4] \quad (\text{resolución del conflicto}) \\ & \mathbf{f}[t_4] \leq \mathbf{f}[t_5] \quad (p_2 \text{ está vacío}) \\ & \mathbf{f}[t_2] \leq \mathbf{f}[t_1] \quad (p_3 \text{ está vacío}) \\ & \mathbf{f}[t_3] \leq \mathbf{f}[t_2] \quad (p_4 \text{ está vacío}) \\ & \mathbf{f}[t_2] + \mathbf{f}[t_4] \leq \mathbf{f}[t_3] + \mathbf{f}[t_5] \quad (p_6 \text{ está vacío}) \\ & \mathbf{f}[t_3] \leq 1 \quad (\text{velocidad máxima}) \\ & \mathbf{f}[t_2] \leq 2 \quad (\text{velocidad máxima}) \\ & \mathbf{f}[t_4] \leq 1 \quad (\text{velocidad máxima}) \end{aligned}$$

(1)

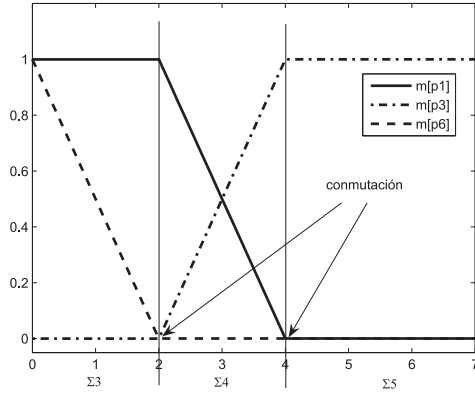


Figura 14. Evolución del sistema representado en la Figura 13 con semántica de finitos servidores

La ecuación asociada a p_2 vacío puede eliminarse, ya que se deduce de p_4 y p_6 vacíos. En términos de redes de Petri, se puede observar que p_2 es *implícito* (es decir, nunca es el único lugar que limita el disparo de t_4 (Dicesare *et al.* 1993), pág. 33), y por lo tanto no es necesario que aparezca. La solución de (1) es $\mathbf{f}[t_2] = \mathbf{f}[t_3] = \mathbf{f}[t_4] = 0,5$. Por lo tanto las ecuaciones que definen la evolución del sistema a partir de $\tau = 2$ son:

$$\begin{aligned}\dot{\mathbf{m}}[p_1](\tau) &= -0,5 \\ \dot{\mathbf{m}}[p_2](\tau) &= \dot{\mathbf{m}}[p_4](\tau) = \dot{\mathbf{m}}[p_5](\tau) = \dot{\mathbf{m}}[p_6](\tau) = 0 \\ \dot{\mathbf{m}}[p_3](\tau) &= 0,5\end{aligned}$$

En $\tau = 4$, p_4 se queda vacío. El marcado en este momento es $[0, 0, 1, 0, 1, 0]$. La transición t_5 tiene todos los lugares de entrada marcados, y por lo tanto $\mathbf{f}[t_5] = 1$. Utilizando un problema de programación lineal como antes se llega a $\mathbf{f}[t_1] = \mathbf{f}[t_2] = \mathbf{f}[t_3] = \mathbf{f}[t_4] = 0,5$, y que corresponde a un estado estacionario.

Semántica de infinitos servidores (e “infinitos” clientes). Bajo esta semántica, los flujos de las transiciones están definidos por:

$$\mathbf{f}[t_1](\tau) = \mathbf{m}[p_1](\tau) \quad (2a)$$

$$\mathbf{f}[t_2](\tau) = 2 \cdot \min \{ \mathbf{m}[p_3](\tau), \mathbf{m}[p_6](\tau) \} \quad (2b)$$

$$\mathbf{f}[t_3](\tau) = \mathbf{m}[p_4](\tau) \quad (2c)$$

$$\mathbf{f}[t_4](\tau) = \min \{ \mathbf{m}[p_2](\tau), \mathbf{m}[p_6](\tau) \} \quad (2d)$$

$$\mathbf{f}[t_5](\tau) = 0,5 \cdot \mathbf{m}[p_5](\tau) \quad (2e)$$

Habida cuenta de que p_2 es implícito $\mathbf{f}[t_4] = \min \{ \mathbf{m}[p_2], \mathbf{m}[p_6] \} = \mathbf{m}[p_6]$. Para $\tau = 0$ se cumple que $\mathbf{m}[p_3] < \mathbf{m}[p_6]$, y por lo tanto la configuración que está activa inicialmente es $\{(p_1, t_1), (p_3, t_2), (p_4, t_3), (p_6, t_4), (p_5, t_5)\}$, que tiene asociado el siguiente sistema lineal:

$$\begin{aligned}\dot{\mathbf{m}}[p_1](\tau) &= \mathbf{f}[t_3] - \mathbf{f}[t_1] = \mathbf{m}[p_4](\tau) - \mathbf{m}[p_1](\tau) \\ \dot{\mathbf{m}}[p_2](\tau) &= \mathbf{f}[t_5] - \mathbf{f}[t_4] = 0,5\mathbf{m}[p_5](\tau) - \mathbf{m}[p_6](\tau) \\ \dot{\mathbf{m}}[p_3](\tau) &= \mathbf{f}[t_1] - \mathbf{f}[t_2] = \mathbf{m}[p_1](\tau) - 2\mathbf{m}[p_3](\tau) \\ \dot{\mathbf{m}}[p_4](\tau) &= \mathbf{f}[t_2] - \mathbf{f}[t_3] = 2\mathbf{m}[p_3](\tau) - \mathbf{m}[p_4](\tau) \\ \dot{\mathbf{m}}[p_5](\tau) &= \mathbf{f}[t_4] - \mathbf{f}[t_5] = \mathbf{m}[p_6](\tau) - 0,5\mathbf{m}[p_5](\tau) \\ \dot{\mathbf{m}}[p_6](\tau) &= \mathbf{f}[t_3] + \mathbf{f}[t_5] - \mathbf{f}[t_2] - \mathbf{f}[t_4] = \\ &= \mathbf{m}[p_4](\tau) + 0,5\mathbf{m}[p_5](\tau) \\ &\quad - 2\mathbf{m}[p_3](\tau) - \mathbf{m}[p_6](\tau)\end{aligned}$$

En la Figura 15 se puede ver la evolución del marcado de este sistema. En $\tau \simeq 1,14$ $\mathbf{m}[p_3](\tau) = \mathbf{m}[p_6](\tau)$ y se cambia de configuración a $\{(p_1, t_1), (p_6, t_2), (p_4, t_3), (p_6, t_4), (p_5, t_5)\}$ que tiene asociado el siguiente sistema lineal:

$$\begin{aligned}\dot{\mathbf{m}}[p_1](\tau) &= \mathbf{m}[p_4](\tau) - \mathbf{m}[p_1](\tau) \\ \dot{\mathbf{m}}[p_2](\tau) &= 0,5\mathbf{m}[p_5](\tau) - \mathbf{m}[p_6](\tau) \\ \dot{\mathbf{m}}[p_3](\tau) &= \mathbf{m}[p_1](\tau) - 2\mathbf{m}[p_6](\tau) \\ \dot{\mathbf{m}}[p_4](\tau) &= 2\mathbf{m}[p_6](\tau) - \mathbf{m}[p_4](\tau) \\ \dot{\mathbf{m}}[p_5](\tau) &= \mathbf{m}[p_6](\tau) - 0,5\mathbf{m}[p_5](\tau) \\ \dot{\mathbf{m}}[p_6](\tau) &= \mathbf{m}[p_4](\tau) + 0,5\mathbf{m}[p_5](\tau) \\ &\quad - 3\mathbf{m}[p_6](\tau)\end{aligned}$$

Finalmente el sistema alcanza un estado estacionario, con $\mathbf{m} = [0,4, 0,6, 0,2, 0,4, 0,4, 0,2]$.

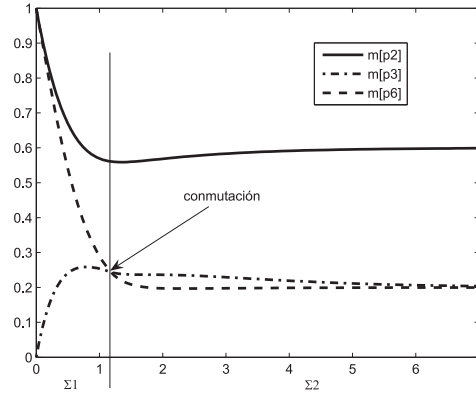


Figura 15. Evolución del sistema representado en la Figura 13 con semántica de infinitos servidores

Los sistemas (no forzados) que se obtienen con ambas semánticas son positivos. Es decir, no es necesario imponer la restricción de que el flujo y el marcado sean no negativos, sino que viene garantizado de forma natural por la propia semántica (Silva y Recalde 2003).

Otras semánticas pueden aparecer de forma natural en dominios de aplicación concretos. Por ejemplo, en *dinámica de poblaciones* el producto de los marcados de los lugares de entrada de una transición puede ser útil para expresar un flujo basado en sumas de tuplas de elementos que se encuentran en una sincronización (Silva y Recalde

2002). Esta semántica puede verse, por ejemplo, como el resultado de decolorar redes coloreadas con semántica de infinitos servidores. Intuitivamente, en la red coloreada se distinguen los individuos, y las transiciones representan el “encuentro” de dos individuos; al decolorar se pasa a tener una población y las transiciones agrupan todos esos posibles encuentros. Por ejemplo, la red discreta coloreada de la Figura 16 representa uno de los modelos clásicos y más sencillos de presa/depredador de Volterra-Lotka. Los marcados $\mathbf{m}[z]$ y $\mathbf{m}[c]$ representan las poblaciones de presas y depredadores, donde cada individuo tiene un color/personalidad diferente. Si consideramos la transición t_3 , está sensibilizada de $\mathbf{m}[c] \cdot \mathbf{m}[z]$ formas/colores diferentes. En el modelo decolorado (discreto) (ver Figura 17(a)) la transición t_3 dispara con tasa $\lambda[t_3] \cdot \mathbf{m}[c] \cdot \mathbf{m}[z]$. Obsérvese que si este modelo se considera como continuo, el sistema ya no es un conjunto de sistemas lineales en conmutación, sino un único sistema no lineal. Al permitir el producto de variables en la definición del flujo de las transiciones, se pueden llegar a modelar comportamientos *caóticos*.

Si las tasas de nacimiento y muerte se definen como $\alpha_r = 0$, $\alpha_f = \alpha$, $\beta_r = 2$, $\beta_f = 0$, las ecuaciones que corresponden a la red continua decolorada son las clásicas de Volterra-Lotka:

$$\begin{aligned}\dot{\mathbf{m}}[c] &= \lambda[t_1]\mathbf{m}[c] - \lambda[t_3]\mathbf{m}[c]\mathbf{m}[z] \\ \dot{\mathbf{m}}[z] &= -\lambda[t_2]\mathbf{m}[z]f + (\alpha - 1)\lambda[t_3]\mathbf{m}[c]\mathbf{m}[z]\end{aligned}$$

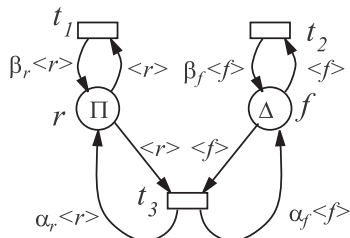


Figura 16. Modelo coloreado de un sistema presa/depredador.

El sistema de la Figura 17(a), visto como discreto es no limitado y no vivo, mientras que visto como continuo es vivo y limitado, debido a la interpretación determinista. Podemos transformar el sistema añadiendo cotas a las poblaciones, representados por los lugares $\sim c$ y $\sim z$ de la Figura 17(b). Si los marcados iniciales de estos lugares son lo bastante grandes, no suponen ninguna restricción al modelo continuo. Sin embargo, el proceso estocástico asociado al modelo discreto acabará cayendo en una de las dos situaciones de bloqueo, $\mathbf{m}[z] = 0$, con $\mathbf{m}[c] = 0$ o $\mathbf{m}[c] = k + 20$.

Como ejercicio, la Figura 18 muestra las trayectorias para el caso en que el número máximo de presas es $20 + k$. Como para este marcado inicial el

lugar $\sim f$ nunca restringe el disparo de t_3 , la ecuación de z en el estacionario es: $\lambda[t_3] \cdot \mathbf{m}[c] \cdot \mathbf{m}[z] - \lambda[t_2] \cdot \mathbf{m}[z] = 0$, y por tanto: $\mathbf{m}[c] = \lambda[t_2]/\lambda[t_3] = 80/3$. Este valor se alcanza realmente sólo en los casos en que k es pequeño. En los otros casos el comportamiento es de tipo oscilatorio, aunque estable.

La clase mono-T-semiflujo-reducible. Es la clase de RdP consistente y conservativa más general para la que la relación de flujo entre las transiciones en estado estacionario (*vector de ratios de visita* en terminología de redes de colas) depende únicamente de la estructura de la red y el vector de velocidades de disparo, pero no de \mathbf{m}_0 . Contiene por ejemplo a las redes de Conflictos Igualados vivas y limitadas (ver Figura 5).

Definición 8. (Júlvez *et al.* 2005) $\langle \mathcal{N}, \lambda \rangle$ es una red *mono-T-semiflujo-reducible* si es consistente, conservativa y el siguiente sistema tiene una única solución:

$$\begin{cases} \mathbf{C} \cdot \mathbf{x} = \mathbf{0} \\ \forall t_i, t_j \text{ en rel. de conflicto igualado continuo,} \\ \quad \forall p \in \bullet t_i, \frac{\mathbf{x}[t_i]}{\mathbf{x}[t_j]} = \frac{\mathbf{Pre}[p, t_i] \cdot \lambda[t_i]}{\mathbf{Pre}[p, t_j] \cdot \lambda[t_j]} \\ \mathbf{x}[t_1] = 1 \end{cases}$$

donde, t_i y t_j están en relación de *conflicto igualado continuo* si existe $q > 0$ tal que $\mathbf{Pre}[P, t_i] = q \cdot \mathbf{Pre}[P, t_j] \neq 0$.

Volviendo ahora a las semánticas básicas de finitos e infinitos servidores, aunque en general no se puede asegurar que una semántica sea mejor que la otra, en (Mahulea *et al.* 2006b) se demuestra que para un importante conjunto de redes pertenecientes a la clase mono-T-semiflujo-reducible, la semántica de infinitos servidores siempre aproxima mejor el flujo del sistema discreto en estado estacionario. Además, en (Alla y David 1998) como en diversos otros trabajos prácticos se ha observado que en muchos casos la semántica de infinitos servidores resulta una notoriamente mejor aproximación del sistema discreto.

Por ello, el resto del trabajo se centrará en la semántica de infinitos servidores.

Homotecias y no monotonías. Las propiedades de homotecia respecto al marcado (escala de amplitud) que aparecían en el modelo autónomo se extienden al temporizado, verificándose además para éste la existencia de propiedades de homotecia con respecto a la velocidad de las transiciones (escala de tiempos):

Propiedad 2. Sea $\langle \mathcal{N}, \mathbf{m}_0, \lambda \rangle$.

1. El sistema $\langle \mathcal{N}, \mathbf{m}_0', \lambda \rangle$ con $\mathbf{m}_0' = k \cdot \mathbf{m}_0$, verifica que para cualquier instante $\tau \geq 0$ $\mathbf{m}'(\tau) = k \cdot \mathbf{m}(\tau)$ y $\mathbf{f}'(\tau) = k \cdot \mathbf{f}(\tau)$ (es decir,

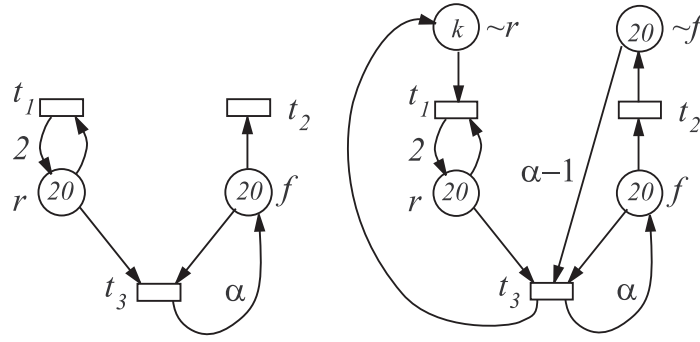


Figura 17. (a) Red decolorada asociada a la red de la Figura 16 (con $\alpha_c = 0, \alpha_z = \alpha, \beta_c = 2, \beta_z = 0$ y $|\Pi| = |\Delta| = 20$), (b) modelo con poblaciones acotadas.

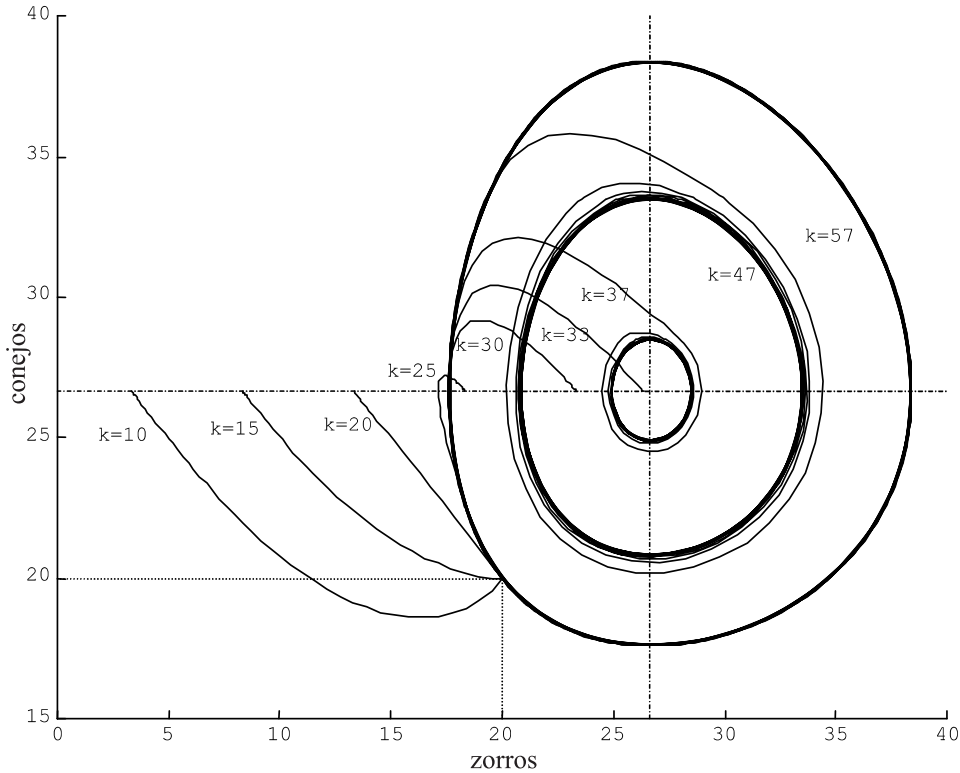


Figura 18. Trayectorias para $\lambda[t_1] = \lambda[t_2] = 20, \lambda[t_3] = 0,75, \alpha = 2, \mathbf{m}_0[c] = \mathbf{m}_0[z] = 20, \mathbf{m}_0[\sim z] = 40$ y $\mathbf{m}_0[\sim c] = k$.

el comportamiento se mantiene por escalado del marcado inicial).

2. El sistema $\langle \mathcal{N}, \mathbf{m}_0, \lambda' \rangle$ con $\lambda' = k \cdot \lambda$, verifica que para cualquier instante $\tau \geq 0$ $\mathbf{m}'(\tau/k) = \mathbf{m}(\tau)$ $\mathbf{f}'(\tau/k) = k \cdot \mathbf{f}(\tau)$ (es decir, el flujo y la escala de tiempos cambia al escalar la velocidad de las transiciones).

Podría parecer que, dado que la continuización relaja las restricciones del sistema, su rendimiento continuo debería ser al menos el del sistema discreto. ¡Esto no es así! El sistema continuo *no es* siempre una cota superior del discreto. Por ejemplo, en el sistema de la Figura 19, con $\lambda = [3 \ 1 \ 1 \ 10]$ y semántica de infinitos servidores, el rendimiento en régimen permanente es 0.801 como discreto mientras que es sólo 0.535 como continuo,

lo que está relacionado con no monotonías, comportamientos paradójicos (contraintuitivos) que se pueden describir con esta clase de modelos.

Como ocurre en las redes discretas, en las prestaciones del sistema continuo temporizado no existe en general ningún tipo de monotonía, ni con respecto al *marcado inicial*, ni con respecto a la *estructura* de la red, ni con respecto a las *velocidades de las transiciones*.

Por ejemplo, si en el sistema de la Figura 19 el marcado inicial de p_5 se aumenta a 5, el sistema se bloquea, es decir, ¡las prestaciones del régimen permanente se hacen 0! Mientras que si el marcado se reduce a 3 esas prestaciones se mejoran al pasar de 0.535 a 1.071. Observar que esta disminución de recursos es equivalente a añadir un lugar “en

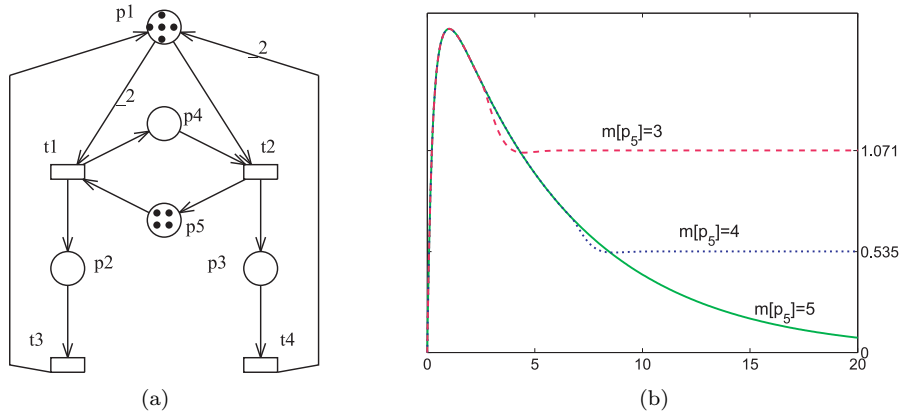


Figura 19. Fijando $\lambda = [3 \ 1 \ 1 \ 10]$, el rendimiento en régimen permanente del sistema continuo, no es cota superior del rendimiento que tiene como discreto. Además, la dinámica no es monótona con respecto a los recursos (marcado) empleado. Obsérvese que el comportamiento a altas frecuencias es sensiblemente idéntico en los tres casos.

paralelo” a p_5 (es decir, con un arco de entrada de t_2 y un arco de salida a t_1), con marcado 3. Por lo tanto, con respecto a la estructura de la red, ¡añadir restricciones puede incrementar las prestaciones!

Finalmente, aumentar la velocidad de una transición (por ejemplo, porque una máquina se reemplaza por otra más rápida) puede también producir una disminución de prestaciones. Por ejemplo, la Figura 20 muestra cómo cambia el rendimiento del sistema en régimen permanente, si la velocidad de t_1 varía entre 0 y 5, suponiendo que $\lambda[t_2] = 1$. Se observa que incluso aparece una *discontinuidad* en $\lambda[t_1] = 2$.

4.2 Análisis (semántica de infinitos servidores)

En (Haddad *et al.* 2006), se ha demostrado que un modelo “análogo” al de redes de Petri continuas con semántica de infinitos servidores es capaz de simular una máquina de Turing. Posteriormente, en (Recalde *et al.* 2007), se ha probado que son realmente dos modelos equivalentes, de donde la potencia descriptiva de esta clase de modelos es alta, aunque el precio pagado es que propiedades como la *alcanzabilidad* dejan de ser decidibles en general. Obviamente, en algunas subclases se encuentra no sólo decidibilidad, sino incluso polinomialidad computacional para determinadas propiedades.

Los valores propios de los sistemas lineales asociados a las configuraciones no dan información sobre estabilidad / tiempo de respuesta del sistema completo. En relación al estudio del estado estacionario, la primera observación es que no se puede garantizar su existencia. Por ejemplo, la red de la Figura 21 oscila indefinidamente (Jiménez *et al.* 2004). Observando las ecuaciones que se aplican inicialmente, el subsistema asociado a los

lugares p_1 y p_2 es un segundo orden marginalmente estable. Además, la solución siempre verifica que $\mathbf{m}[p_1] > 1$ y $\mathbf{m}[p_2] > \mathbf{m}[p_1]/20$, es decir no hay conmutación en las transiciones t_2 y t_4 . Sí que hay conmutación en t_5 . Obsérvese en la simulación que algunas veces $\mathbf{m}[p_3] < \mathbf{m}[p_4]$ y otras $\mathbf{m}[p_3] > \mathbf{m}[p_4]$. En realidad, como la alcanzabilidad, el problema de si el sistema se aproxima a un estado estacionario es indecidible en general (Recalde *et al.* 2007).

Sin embargo, este estado estacionario, si existe, tiene que cumplir ciertas condiciones: el flujo de transiciones asociado tiene que ser un T-semiflujo (ya que $\dot{\mathbf{m}} = \mathbf{C} \cdot \mathbf{f} = \mathbf{0}$), y tiene que verificar la ecuación de estado (necesario para que sea alcanzable). Es decir,

$$\begin{cases} \mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \boldsymbol{\sigma} \\ \mathbf{C} \cdot \mathbf{f} = \mathbf{0} \\ \mathbf{f}[t_j] = \lambda[t_j] \cdot \min_{p_i \in \bullet t_j} \left\{ \frac{\mathbf{m}[p_i]}{\text{Pre}[p_i, t_j]} \right\}, \forall t_j \in T \\ \mathbf{f}, \boldsymbol{\sigma} \geq \mathbf{0} \end{cases} \quad (3)$$

Si la red es consistente y conservativa la ecuación de estado se puede sustituir por $\mathbf{B}_y \cdot \mathbf{m} = \mathbf{B}_y \cdot \mathbf{m}_0$, $\mathbf{m} \geq \mathbf{0}$, con \mathbf{B}_y una base de P-(semi)flujos. En otras palabras, las soluciones de estas ecuaciones representan todas las formas posibles de distribuir las marcas de los P-(semi)flujos que hacen que el sistema se mantenga en ese estado. Sin embargo, puede ocurrir que varios marcados verifiquen las ecuaciones. Son los *marcados posibles de equilibrio*, y las configuraciones asociadas son las *configuraciones posibles de equilibrio*. Hay casos en los que sólo existe un marcado de equilibrio posible (redes Sin Sincronización, Figura 22), pero en otros pueden existir infinitos. Por otro lado, si hay varios, puede ser que todos estén en una única configuración (Figura 23), o en varias (Figura 24) (Mahulea *et al.* 2007). Como se ha visto previamente, también puede ser que no existan

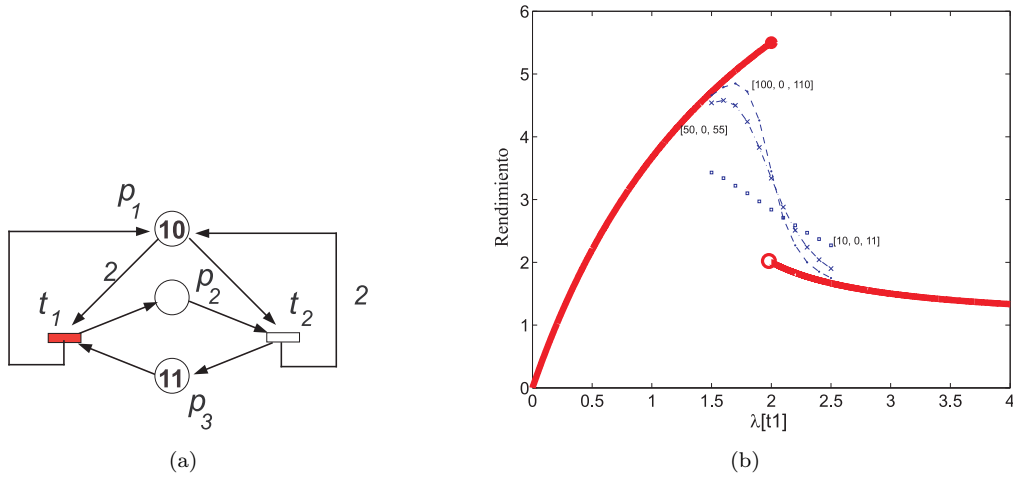


Figura 20. Aumentar la velocidad de t_1 no aumenta necesariamente el rendimiento del sistema completo. Con trazo discontinuo las prestaciones del sistema discreto para \mathbf{m}_0 , $5\mathbf{m}_0$ y $10\mathbf{m}_0$.

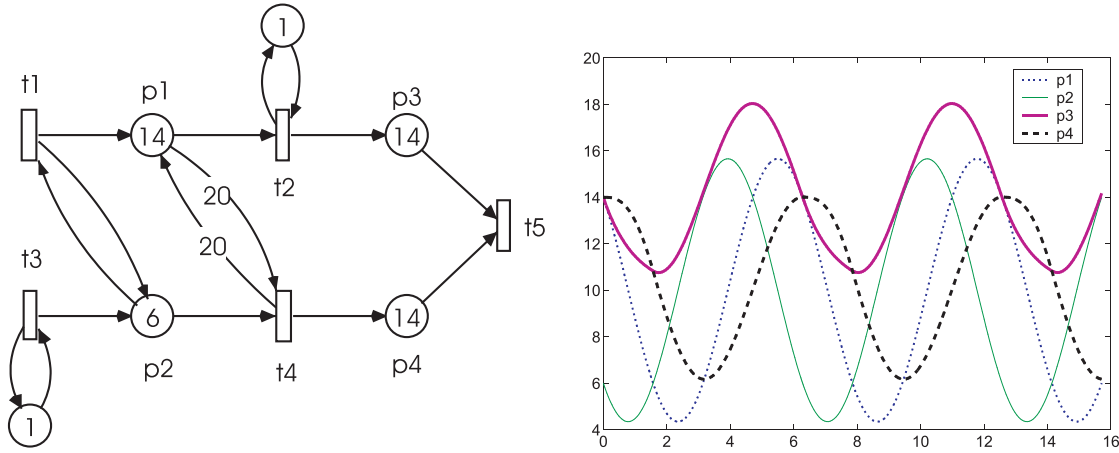


Figura 21. Una red continua y su evolución con velocidades $\lambda = [1 \ 10 \ 10 \ 20 \ 1]$.

estados de equilibrio, por ejemplo en el sistema de la Figura 21.

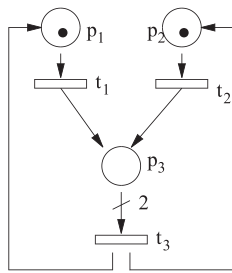


Figura 22. Con $\lambda = [1 \ 1 \ 1]$ este sistema tiene un único marcado de equilibrio posible, $\mathbf{m} = [0,66 \ 0,66 \ 0,66]$.

Teorema 7. (Mahulea *et al.* 2007) Sea un sistema continuo $\langle \mathcal{N}, \lambda, \mathbf{m}_0 \rangle$, consistente y en el que todas las transiciones son disparables. Si

$$\begin{aligned} \text{rango} \left[\begin{array}{c|c} \Lambda \cdot \Pi & \mathbf{B}_x \\ \mathbf{B}_y & 0 \end{array} \right] &= \\ &= \text{rango} \left[\begin{array}{c} \Pi \\ \mathbf{B}_y \end{array} \right] + |T| - \text{rango}(\mathbf{C}), \end{aligned}$$

con \mathbf{B}_x y \mathbf{B}_y bases de T-flujos y P-flujos, respectivamente, entonces todos los marcados posibles de equilibrio en la misma configuración tienen idéntico flujo.

Para las redes de Conflictos Igualados lim-vivas y limitadas, dada una carga para los P-flujos ($\mathbf{B}_y \cdot \mathbf{m} = \mathbf{B}_y \cdot \mathbf{m}_0$, con \mathbf{B}_y una base de P-flujos), aunque puedan existir infinitos marcados de régimen permanente, se puede demostrar que todos los puntos de equilibrio tienen el mismo flujo de transiciones. Este resultado se puede deducir trivialmente a partir de resultados en (Mahulea *et al.* 2007) y del teorema del rango para redes de Conflictos Igualados (Recalde *et al.* 1999).

Teorema 8. Sea \mathcal{N} una red de Conflictos Igualados consistente, conservativa y $\text{rango}(\mathbf{C}) = |T| - 1$. Sea $\mathbf{B}_y \mathbf{m}_0$ una distribución de cargas de los flujos que no deja vacío ninguno, \mathbf{B}_y una base de P-flujos.

Para todo marcado inicial \mathbf{m}_0' que verifique $\mathbf{B}_y \mathbf{m}_0' = \mathbf{B}_y \mathbf{m}_0$, existe al menos un marcado

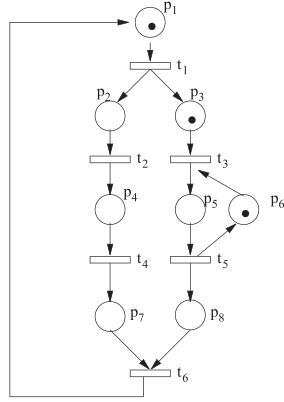


Figura 23. Con $\lambda = [1 \ 1 \ 1 \ 1 \ 1 \ 1]$ este sistema tiene muchos puntos de equilibrio posibles, todos en la misma configuración.

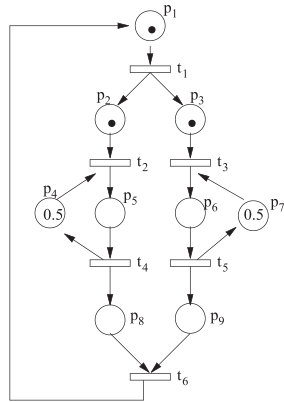


Figura 24. Con $\lambda = [1 \ 1 \ 1 \ 1 \ 1 \ 1]$ este sistema tiene muchos puntos de equilibrio posibles, en varias configuraciones. Existe conexidad pero no convexidad en los puntos (regiones) posibles de equilibrio.

posible de equilibrio. Si hay varios, todos tienen el mismo flujo, que sólo depende de la carga de los P-flujos ($\mathbf{B}_y \mathbf{m}_0$), y no de la distribución inicial concreta.

Las redes de las Figuras 23 y 24 pertenecen a la clase de Conflictos Igualados (y a la clase de redes Sin Elección). Por lo tanto todos los puntos de equilibrio tienen el mismo flujo. El siguiente teorema proporciona una condición suficiente (para redes de Conflictos Igualados) para que en una configuración haya como mucho un punto de equilibrio.

Teorema 9. (Mahulea *et al.* 2007) Sea un sistema de Conflictos Igualados $(\mathcal{N}, \lambda, \mathbf{m}_0)$, lim-vivo y limitado como continuo, y sea Π_k la matriz asociada a la configuración k -ésima. Si $\text{rank} \begin{bmatrix} \Pi_k \\ \mathbf{B}_y \end{bmatrix} = |P|$ existe a lo sumo un marcado de equilibrio en esta configuración \mathcal{C}_k .

Las redes Sin Sincronización tienen una única configuración, y si son consistentes y conservativas la condición anterior se cumple siempre.

Corolario 1. (Mahulea *et al.* 2007) Sea \mathcal{N} una red Sin Sincronización conservativa y consistente. Para cualquier marcado inicial, y cualquier vector λ , existe un único marcado de equilibrio.

El sistema de ecuaciones (3) puede utilizarse para obtener cotas superiores del flujo de una transición en estado estacionario.

$$\begin{aligned} \max : & \mathbf{f}[t_i] \\ \text{s.a. } & \mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \boldsymbol{\sigma} \end{aligned} \quad (4)$$

$$\mathbf{C} \cdot \mathbf{f} = \mathbf{0} \quad (5)$$

$$\mathbf{f}[t_j] = \lambda[t_j] \cdot \min_{p_i \in \bullet t_j} \left\{ \frac{\mathbf{m}[p_i]}{\mathbf{Pre}[p_i, t_j]} \right\}, \quad (6)$$

$$\forall t_j \in T$$

$$\mathbf{f}, \boldsymbol{\sigma} \geq \mathbf{0} \quad (7)$$

Este problema de programación es difícil de resolver debido a la función “min” que aparece en la ecuación 6, que lo hace no lineal. Esta ecuación puede *relajarse (linealizarse)*. Para ello es conveniente distinguir entre las transiciones que tienen un único lugar de entrada (para las cuales el mínimo puede obviarse) y las transiciones con varios lugares de entrada, para las cuales el mínimo se

puede sustituir por un conjunto de condiciones más débiles:

$$\begin{aligned} \text{si } p_i = \bullet t_j, \quad \mathbf{f}[t_j] &= \boldsymbol{\lambda}[t_j] \cdot \frac{\mathbf{m}[p_i]}{\mathbf{Pre}[p_i, t_j]} \\ \text{si } |\bullet t_j| > 1, \quad \forall p_i \in \bullet t_j & \\ \mathbf{f}[t_j] &\leq \boldsymbol{\lambda}[t_j] \cdot \frac{\mathbf{m}[p_i]}{\mathbf{Pre}[p_i, t_j]} \end{aligned} \quad (8)$$

Sin embargo, esta cota puede no ser alcanzable, debido a que el marcado que se obtiene puede no corresponder con un flujo permanente, en el sentido de que ninguno de los lugares de una sincronización esté limitándolo. Por ejemplo, maximizando $\mathbf{f}[\text{Out}]$ la solución para el sistema continuo de la Figura 6 es: $\mathbf{m}[\text{M1.Idle}] = \mathbf{m}[\text{B.2A}] = \mathbf{m}[\text{Max.A}] = \mathbf{m}[\text{B.2B}] = \mathbf{m}[\text{Max.B}] = \mathbf{m}[\text{B.3}] = \mathbf{m}[\text{M3.Idle}] = 0,111$, $\mathbf{m}[\text{M1.A}] = 0,333$, $\mathbf{m}[\text{Pallets.A}] = 9$, $\mathbf{m}[\text{B.1A}] = 9,111$, $\mathbf{m}[\text{M2.A}] = 0,444$, $\mathbf{m}[\text{M2.B}] = 0,333$, $\mathbf{m}[\text{Pallets.B}] = 4$, $\mathbf{m}[\text{B.1B}] = 9$, $\mathbf{m}[\text{M1.B}] = 0,555$, $\mathbf{m}[\text{M3.Work}] = 0,888$, $\mathbf{m}[\text{B.3.Empty}] = 9,888$, $\mathbf{m}[\text{M2.Idle}] = 0,222$, y $\mathbf{f}[\text{Out}] = 0,111$. Pero el flujo de las transiciones S_M2_A and E_M3 no se corresponde con el marcado de sus lugares de entrada (les “sobran” marcas).

Para obtener una cota alcanzable (para alguna distribución de las marcas), se puede realizar una búsqueda selectiva por regiones para resolver (4)-(7), utilizando por ejemplo una técnica de tipo ramificación y acotación. La idea es resolver primero el problema de programación lineal definido por el sistema de inecuaciones (8). Si el marcado que se obtiene no se corresponde con el flujo (es decir, hay al menos una transición tal que todos sus lugares de entrada tienen “demasiadas” marcas para el flujo calculado) se elige una de estas sincronizaciones y se resuelven los problemas de programación lineal que aparecen cuando cada uno de los lugares de entrada se usa para definir el mínimo. Estos subproblemas son los “hijos” del nodo raíz de la búsqueda. El algoritmo se aplica recursivamente generando un árbol de subproblemas. Los marcados óptimos que sean solución factible del problema (4)-(7) se pueden utilizar para podar el resto del árbol: si la solución del problema de programación lineal de un nodo tiene flujo menor, no es posible encontrar soluciones óptimas a partir de ese nodo. La búsqueda continúa hasta que todos los nodos han sido resueltos o descartados.

Por ejemplo, para el sistema anterior se crean dos problemas añadiendo en cada uno una ecuación para S_M2_A (ver Figura 25). Si se hace que el flujo esté definido por M2.Idle, el sistema lineal resulta no factible, y si se añade una restricción para B_1A el flujo que se obtiene es el mismo, pero con $\mathbf{m}[\text{Pallets.A}] = 18$, $\mathbf{m}[\text{B.1A}] = 0,111$, y $\mathbf{m}[\text{Max.A}] = 9,111$. Ahora el único problema es E_M3. Si se añade una ecuación para

B_3.Empty el sistema es no factible, y añadiendo una ecuación para M3_Work sólo se modifican $\mathbf{m}[\text{Pallets.A}] = 18,444$, $\mathbf{m}[\text{Pallets.B}] = 4,444$, $\mathbf{m}[\text{M3.Idle}] = 0,555$, y $\mathbf{m}[\text{M3.Work}] = 0,444$. Este marcado sí se corresponde con un flujo permanente, y no es posible mejorarlo, luego no es necesario igualar el resto de sincronizaciones.

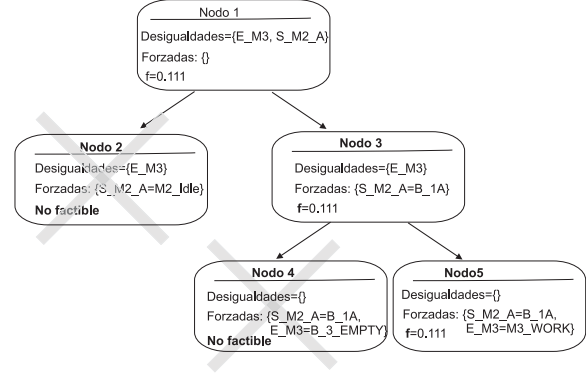


Figura 25. Ramificación y acotación para el ejemplo de la Figura 6.

Por propiedades de escala (Propiedad 2) si el vector $\boldsymbol{\lambda}$ se multiplica por una constante $k > 0$ entonces para cualquier marcado alcanzable el flujo de las transiciones también se multiplica por k . De la misma forma, si el marcado inicial de la red se multiplica por k , el sistema también será k veces más rápido. Sin embargo, ¿qué ocurre si sólo aumentan algunas componentes de $\boldsymbol{\lambda}$ o de \mathbf{m}_0 ? Se ha visto previamente que en general, como ocurre para las redes de Petri discretas, aumentar la velocidad de una transición, o el marcado inicial de un lugar, puede ralentizar el sistema. Sin embargo, este comportamiento no suele ser deseable. Por ejemplo, al reemplazar una máquina por otra más rápida, o al añadir más máquinas, el objetivo normalmente buscado es aumentar el rendimiento, no disminuirlo. Sin embargo, existen subclases de redes para las que se puede demostrar que cumplen propiedades de monotonía bajo condiciones bastante generales. Por ejemplo, las redes mono-T-semiflujo-reducibles.

Teorema 10. (Mahulea *et al.* 2006b) Sean los sistemas mono-T-semiflujo-reducibles $\langle \mathcal{N}, \boldsymbol{\lambda}_1, \mathbf{m}_0 \rangle$ y $\langle \mathcal{N}, \boldsymbol{\lambda}_2, \mathbf{m}_0 \rangle$, y $\boldsymbol{\lambda}_1 \leq \boldsymbol{\lambda}_2$ tales que imponen las mismas proporciones en los conflictos igualados (es decir, para cada pareja t_i, t_j en relación de conflicto igualado, $\boldsymbol{\lambda}_1[t_i]/\boldsymbol{\lambda}_1[t_j] = \boldsymbol{\lambda}_2[t_i]/\boldsymbol{\lambda}_2[t_j]$). Si estos dos sistemas alcanzan un estado estacionario, y las correspondientes configuraciones contienen el soporte de un P-semiflujo, entonces los flujos en estado estacionario verifican $\mathbf{f}_1 \leq \mathbf{f}_2$.

Este resultado permite obtener una interesante condición suficiente de monotonía de prestaciones con respecto a $\boldsymbol{\lambda} \in \mathcal{P}$, $\mathcal{P} \subseteq \mathbb{R}_{>0}^{|T|}$.

Teorema 11. (Mahulea *et al.* 2006b) Sea un sistema continuo mono-T-semiflujo-reducible definido por $\langle \mathcal{N}, \lambda, \mathbf{m}_0 \rangle$. Sea $\mathcal{P} \subseteq \mathbb{R}_{>0}^{|T|}$ tal que para cada pareja t_i, t_j en relación de conflicto igualado, $\lambda[t_i]/\lambda[t_j]$ es constante $\forall \lambda \in \mathcal{P}$. Si todas las posibles configuraciones de equilibrio contienen un P-semiflujo el sistema es monótono con respecto a λ en \mathcal{P} .

Teorema 12. (Mahulea *et al.* 2006b) Sea $\langle \mathcal{N}, \lambda \rangle$ una red temporizada continua mono-T-semiflujo-reducible, y sea $\mathbf{m}_1 \leq \mathbf{m}_2$. Si con ambos marcados iniciales se alcanza un estado estacionario, y las correspondientes configuraciones contienen el soporte de un P-semiflujo, entonces los flujos en estado estacionario verifican $\mathbf{f}_1 \leq \mathbf{f}_2$.

Ejemplo 3. Sea la red mono-T-semiflujo de la Figura 26 con $\lambda = [1 \ 1 \ 1]$. Tiene dos P-semiflujos mínimos, $[1 \ 1 \ 1 \ 0]$ y $[1 \ 0 \ 4 \ 1]$, y 4 configuraciones posibles: $\mathcal{C}_1 = \{p_1, p_2, p_3\}$, $\mathcal{C}_2 = \{p_1, p_4, p_3\}$, $\mathcal{C}_3 = \{p_4, p_2, p_3\}$ and $\mathcal{C}_4 = \{p_4, p_3\}$.

Sea $\mathcal{M}_1 = \{[0 \ 1 \ 1 \ z] \mid z > 0\}$. Dado un marcado inicial perteneciente a \mathcal{M}_1 , veamos que \mathcal{C}_4 no puede ser configuración de equilibrio. Si lo fuera, p_4 limitaría el flujo de t_1 y t_2 , y teniendo en cuenta que el único T-semiflujo mínimo es $[1 \ 1 \ 1]$, $\mathbf{f}[t_1] = \mathbf{f}[t_2] \rightarrow \mathbf{m}[p_4]/2 = \mathbf{m}[p_4]$. Luego, $\mathbf{m}[p_4] = 0$ y por el T-semiflujo, $\mathbf{m}[p_3] = 0$. Pero entonces no es posible verificar las ecuaciones asociadas a los P-semiflujos. Razonando de forma análoga se demuestra que \mathcal{C}_3 , tampoco puede ser configuración de equilibrio. Aplicando el Teorema 12, el flujo en estado estacionario es monótono con respecto a z .

Sea $\mathcal{M}_2 = \{[15 \ z \ 1 \ 0] \mid z > 0\}$. Ahora \mathcal{C}_3 y \mathcal{C}_4 pueden ser configuraciones de equilibrio y la monotonía se pierde, como puede comprobarse en la Figura 26.

Este resultado también se puede usar para obtener una condición suficiente de monotonía con respecto al marcado inicial en una región.

Teorema 13. Sea una red $\langle \mathcal{N}, \lambda \rangle$ mono-T-semiflujo-reducible, y sea $\mathcal{R} \subseteq \mathbb{R}_{>0}^{|P|}$. Si para todo $\mathbf{m}_0 \in \mathcal{R}$, todas las posibles configuraciones de equilibrio contienen un P-semiflujo, entonces el sistema es monótono con respecto a \mathbf{m}_0 en \mathcal{R} .

Obviamente, si todas las posibles configuraciones contienen un P-semiflujo, se tiene monotonía para cualquier temporización del sistema en todo $\mathbb{R}_{>0}^{|P|}$. Esta condición se puede verificar utilizando variables booleanas para seleccionar las configuraciones, análogamente a como se hizo anteriormente para ausencia de bloqueos.

5. OBSERVACIÓN Y CONTROL DE MODELOS TEMPORIZADOS

Para controlar un sistema dinámico, a menudo es necesario conocer su estado. Para obtener esta información, se pueden utilizar sensores distribuidos a lo largo del sistema. Sin embargo, puede ocurrir que algunas variables del sistema no se puedan medir directamente, o que la colocación de algunos de los sensores tenga un coste excesivo. Sin embargo, bajo ciertas condiciones algunas de estas variables no medidas se pueden estimar. Esta estimación constituye la observación.

5.1 Observación: apuntes

Cada lugar puede ser *medido* o *no medido*. Se dice que un lugar p_i es medido si existe algún tipo de sensor que permite conocer su nivel. Denotaremos por \mathbf{S} la matriz de los lugares medidos, es decir, para cada p_i medido existe una fila j en \mathbf{S} tal que $\mathbf{S}[j, i] \neq 0$ y $\mathbf{S}[j, k] = 0$ para todo $k \neq i$.

Definición 9. Sea $\langle \mathcal{N}, \lambda \rangle$ una red de Petri temporizada, y \mathcal{D} el conjunto de lugares medidos.

- Un lugar $p \in P$ es *observable* con \mathcal{D} si es posible calcular su marcado inicial $\mathbf{m}_0[p] = \mathbf{m}(\tau_0)[p]$ midiendo la evolución de \mathcal{D} . Si todo lugar es observable se dice que el sistema $\langle \mathcal{N}, \lambda \rangle$ es *observable* con \mathcal{D} .
- Un lugar $p \in P$ es *estructuralmente observable* con \mathcal{D} si es observable con \mathcal{D} para cualquier $\lambda > 0$. Si todo lugar es estructuralmente observable se dice que la red \mathcal{N} es *estructuralmente observable* con \mathcal{D} .

La observabilidad estructural puede interpretarse como una propiedad de observabilidad “robusta” ya que una red estructuralmente observable sigue siendo observable a pesar de los cambios que pueda sufrir el vector λ (supuesto conocido). Por ejemplo, supóngase que el único lugar medido del sistema de la Figura 27 es p_3 . La variación del marcado de un lugar viene dado por la diferencia entre los flujos de entrada y de salida. Para p_3 , se tiene que $\dot{\mathbf{m}}[p_3] = \mathbf{f}[t_2] - \mathbf{f}[t_3]$ donde: $\mathbf{f}[t_2] = \lambda[t_2] \cdot \mathbf{m}[p_2]$ y $\mathbf{f}[t_3] = \lambda[t_3] \cdot \mathbf{m}[p_3]$. Por lo tanto, $\mathbf{m}[p_2] = (\dot{\mathbf{m}}[p_3] + \lambda[t_3] \cdot \mathbf{m}[p_3])/\lambda[t_2]$. Es decir, conocido $\mathbf{m}[p_3]$ se puede calcular la evolución de $\mathbf{m}[p_2]$. Además, se verifica que $\dot{\mathbf{m}}[p_2] = \mathbf{f}[t_1] - \mathbf{f}[t_2]$ y $\mathbf{f}[t_1] = \lambda[t_1] \cdot \mathbf{m}[p_1]$, luego $\mathbf{m}[p_1]$ también se puede calcular. Este procedimiento se puede realizar para cualquier valor de λ , y por tanto la red es estructuralmente observable.

Proposition 1. (Júlvez *et al.* 2004) Sea \mathcal{N} una red continua Sin Sincronización y \mathcal{D} el conjunto de los lugares medidos. Sea p tal que existe un camino

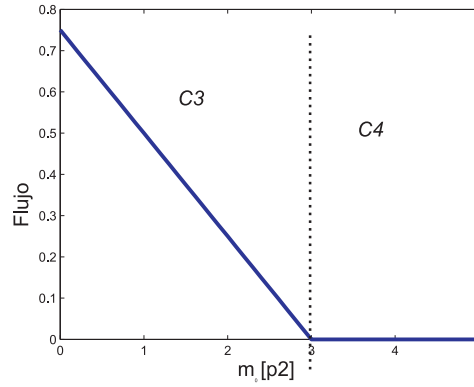
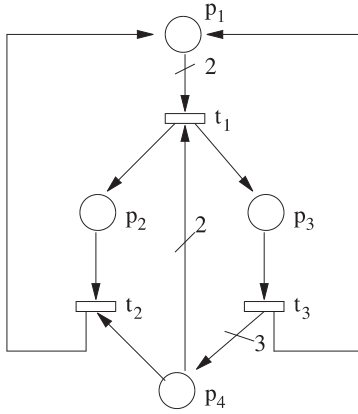


Figura 26. Flujo en estado estacionario del sistema representado si $\mathbf{m}_0 = [15, z, 1, 0]$ y $\lambda = [1, 1, 1]$.

de p a \mathcal{D} tal que cada lugar tiene una única transición de entrada. Entonces, p es estructuralmente observable.

Si hay lugares con varias transiciones de entrada, es posible que el sistema no sea observable para todos los valores de λ . Por ejemplo, el sistema de la Figura 22 con $\lambda[t_1] = \lambda[t_2]$ y p_3 como único lugar medido, no es observable. Como p_3 tiene dos transiciones de entrada, ni p_1 ni p_2 verifican la condición de la Proposición 1. De hecho, si se calcula la matriz de observabilidad del sistema (¡como no tiene sincronizaciones es un sistema lineal!) se deduce que p_1 y p_2 no pueden estimarse. Intuitivamente, si los flujos de las transiciones t_1 y t_2 tienen “igual velocidad”, no se puede decidir cuánto procede de cada lugar. Por lo tanto, la red no es estructuralmente observable, aunque sí es observable si $\lambda[t_1] \neq \lambda[t_2]$.

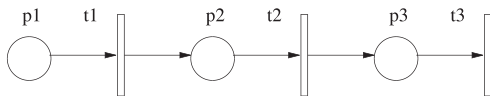


Figura 27. El marcado del sistema se puede deducir observando p_3 .

Aplicando la Propiedad 1 de forma iterativa en redes Sin Sincronización, se puede construir un algoritmo de punto fijo: a partir de los lugares medidos u observados en un paso, se añaden los lugares que se pueden deducir aplicándola, construyendo así un conjunto de lugares estructuralmente observables (Júlvez *et al.* 2004).

Esta propiedad no se puede aplicar si la red no es Sin Sincronización. Si una transición tiene más de un lugar de entrada, en un instante sólo se puede calcular el marcado del lugar que restringe el flujo en ese momento. Puede ocurrir que un lugar de entrada de la sincronización nunca sea el que limita el flujo de las transiciones de las que es entrada (es decir, sea implícito temporal). En ese caso la única forma de asegurar que se observa es midiéndolo directamente.

Si lo que se desea es estudiar la observabilidad para un λ concreto, se puede calcular un observador lineal para cada una de las configuraciones, y ejecutarlos en paralelo. Las estimaciones que se obtengan pueden utilizarse para desechar aquellas configuraciones que es seguro que no están limitando el sistema. Esto puede suceder por distintas causas. Se llamarán estimaciones *no factibles*, si no existe solución del sistema lineal para estas observaciones; se denominarán *incoherentes*, si la configuración que define las estimación no coincide con aquélla para la que se calculó; finalmente son *sospechosas*, las que pertenecen a varias regiones. Puede que después de eliminar estos estimadores, todavía queden otros varios válidos. Utilizaremos un *residuo* para elegir entre ellos una observación. El residuo en un instante dado, $r(\tau)$, es la distancia entre la salida del sistema y la salida de la estimación del observador. Para medir esta distancia en los ejemplos aquí utilizados se ha utilizado la norma L1.

Veamos un ejemplo utilizando observadores lineales de Luenberger. Sea el sistema de la Figura 28 y sea la salida del sistema el marcado de p_1 , es decir, $\mathbf{S} = (1 \ 0 \ 0)$. La red tiene dos configuraciones, en función de qué lugar limita a t_2 : $Z_1 = \{(p_1, t_1), (p_2, t_2), (p_3, t_3)\}$ y $Z_2 = \{(p_1, t_1), (p_2, t_2), (p_3, t_3)\}$. Por cuestiones puramente estructurales, para el sistema lineal definido por Z_1 sólo se puede estimar p_1 y p_3 . Sin embargo Z_2 permite estimar el marcado (estado) de todos los lugares. Sea $\lambda = ([0,9 \ 1 \ 1])$ y $\mathbf{m}_0 = [3 \ 0 \ 0]$. La evolución del marcado se muestra en la Figura 29.

Supóngase que el estado inicial del observador 1 (asociado a Z_1) es $\mathbf{e}_{01} = (1 \ 2)$ y que sus valores propios son $(-12 + 2 \cdot \sqrt{3} \cdot i, -12 - 2 \cdot \sqrt{3} \cdot i)$. Para el observador 2 (asociado a Z_2), sean $\mathbf{e}_{02} = (1 \ 0 \ 2)$ el estado inicial y $(-15, -12 + 2 \cdot \sqrt{3} \cdot i, -12 - 2 \cdot \sqrt{3} \cdot i)$ los valores propios. Las evoluciones de las estimaciones de los observadores están representadas en las Figuras 30 y 31.

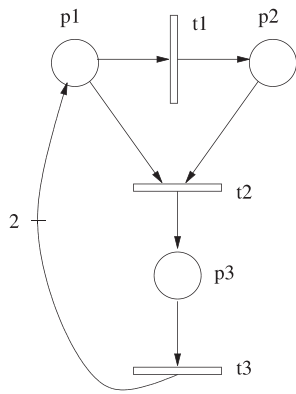


Figura 28. Sistema con dos configuraciones

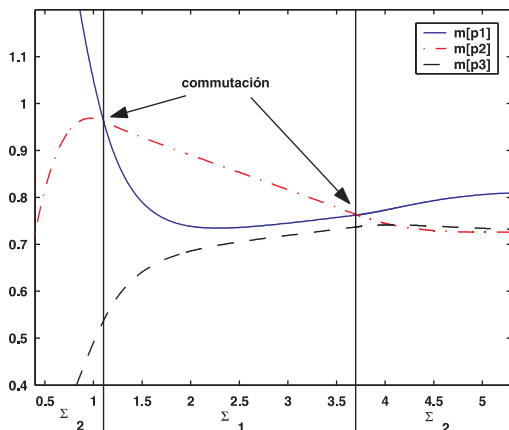


Figura 29. Evolución del marcado del sistema de la Figura 28 con $\mathbf{m}_0 = [3 \ 0 \ 0]$ y $\lambda = ([0,9 \ 1 \ 1])$.

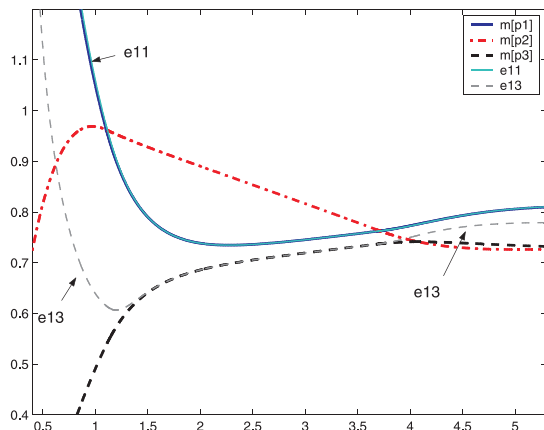


Figura 30. Evolución del observador 1, (e_{11} , e_{13}) es la estimación de $(\mathbf{m}[p_1], \mathbf{m}[p_3])$.

Inicialmente los dos observadores son válidos, pero el residuo del observador 2 es menor que el del observador 1 (coherente con que sea Z_2 la configuración en la que se encuentra el sistema). En $\tau = 1,1$ el sistema conmuta a Z_1 . El observador 2 deja de ser coherente, ya que su estimación predice un valor de $\mathbf{m}[p_2] > \mathbf{m}[p_1]$, en contra de la hipótesis de que se está en la configuración 2.

La estimación resultante puede mejorarse todavía utilizando información de que se dispone. Por

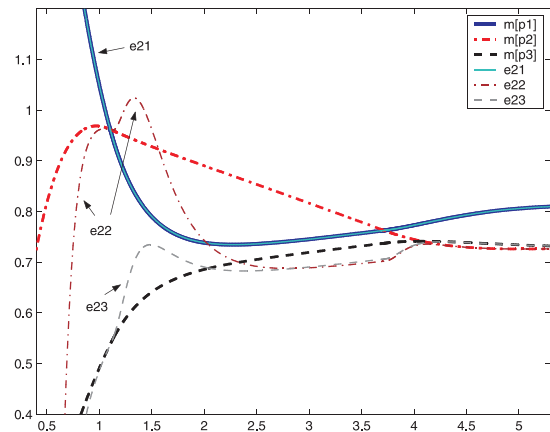


Figura 31. Evolución del observador 2, (e_{21} , e_{22} , e_{23}) es la estimación de $(\mathbf{m}[p_1], \mathbf{m}[p_2], \mathbf{m}[p_3])$.

ejemplo, en la Figura 32, cuando la primera conmutación ocurre, la estimación del observador 2 está muy próxima al marcado real. Al conmutar el observador se produce un empeoramiento inicial de la estimación. Para evitarlo, basta con actualizar el observador 1 con los valores que predice el observador 2 justo antes de conmutar. Por otra parte, cuando se utiliza el observador 1 no se dispone de estimación de $\mathbf{m}[p_2]$. Sin embargo, si se dispone de una estimación de todos los lugares con un error pequeño, sí que se puede realizar una simulación a partir de la estimación. La Figura 33 muestra los resultados de la observación aplicando estas sencillas mejoras.

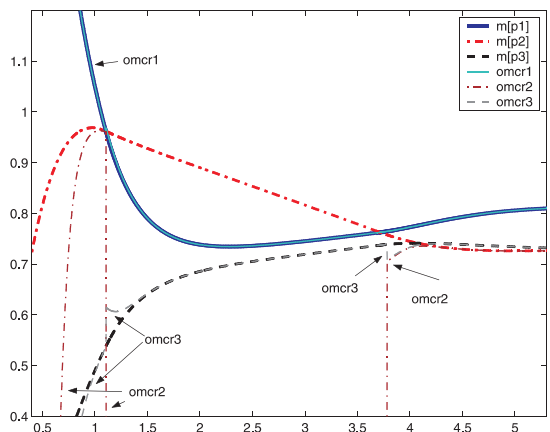


Figura 32. Observador coherente con mínimo residuo, ($omcr1$, $omcr2$, $omcr3$) es la estimación de $(\mathbf{m}[p_1], \mathbf{m}[p_2], \mathbf{m}[p_3])$.

5.2 Control en régimen permanente

Sobre los modelos de RdP se pueden plantear problemas de control del mismo tipo que en control clásico. Por ejemplo, se puede estudiar cuál es el control que hay que realizar sobre el sistema para que en régimen permanente el rendimiento

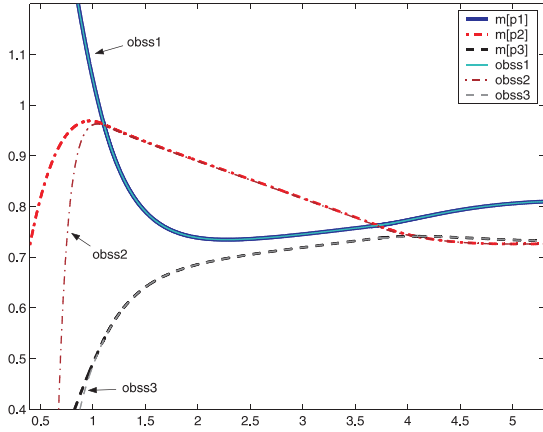


Figura 33. Mejora del observador con inicialización al conmutar y simulación, ($obs1$, $obs2$, $obs3$) es la estimación de ($\mathbf{m}[p_1]$, $\mathbf{m}[p_2]$, $\mathbf{m}[p_3]$).

sea óptimo, o cuál es la política de control que permite pasar de un marcado \mathbf{m}_0 a un marcado \mathbf{m} en tiempo mínimo.

De acuerdo con la semántica de los modelos, se supondrá que la única acción que se puede aplicar es *reducir* el flujo de disparo de transiciones. Pensando por ejemplo en sistemas de producción, resulta bastante lógico pensar en frenar las máquinas, pero resulta imposible hacer que vayan más rápido que su velocidad máxima. En otras palabras, sólo podemos actuar “cerrando válvulas”, frenando parcialmente. Como los sistemas no son monótonos, “frenar” puede hacer que el sistema vaya más rápido.

Si una transición se puede controlar (su flujo se puede reducir o incluso detener completamente), se dirá que es *controlable*. El flujo de una transición controlable t_i es $\mathbf{f}[t_i] - \mathbf{u}[t_i]$, donde $\mathbf{f}[t_i]$ es el flujo del sistema *no forzado* y $\mathbf{u}[t_i]$ es la acción de control y verifica $0 \leq \mathbf{u}[t_i] \leq \mathbf{f}[t_i]$. Obsérvese que esto significa que el valor máximo de la acción de control es variable, depende del marcado.

Por ejemplo, para la red de la Figura 19 con $\mathbf{m}_0 = [5 \ 0 \ 0 \ 0 \ 4]$, el flujo de las transiciones en régimen permanente si $\mathbf{u} = [0 \ 0 \ 0 \ 0]$ (i.e., si el sistema está *sin controlar*) es $\mathbf{f} = [0,535 \ 0,535 \ 0,535 \ 0,535]$. Sin embargo se puede llegar a conseguir que el flujo de cada transición sea 1.5625 si se usa como control $\mathbf{u} = [0,781 \ 0 \ 0 \ 0]$. ¡Es decir, frenando el flujo de t_1 se llega a casi multiplicar por tres el flujo global, y el de t_1 en particular!

En principio, se puede plantear si un subconjunto de transiciones es suficiente para controlar el sistema, o cuál es el conjunto mínimo que permite controlarlo. Sin embargo, este tipo de problemas no se van a abordar aquí. Por simplicidad se supondrá en lo sucesivo que todas las transicio-

nes son controlables. La ecuación de estado del sistema controlado es lineal a tramos y con acción acotada asimétricamente, superiormente por una función del marcado e inferiormente por cero:

$$\begin{cases} \dot{\mathbf{m}} = \mathbf{C} \cdot (\boldsymbol{\Lambda} \cdot \boldsymbol{\Pi}(\mathbf{m}) \cdot \mathbf{m} - \mathbf{u}) \\ 0 \leq \mathbf{u} \leq \boldsymbol{\Lambda} \cdot \boldsymbol{\Pi}(\mathbf{m}) \cdot \mathbf{m} \end{cases} \quad (9)$$

Controlando todas las transiciones, los marcados alcanzables en el sistema autónomo (no temporizado) pueden alcanzarse en el sistema temporizado, salvo quizás los que tienen algún lugar sin marcas (varios). En este caso, el marcado se alcanza en el límite (es como descargar un condensador en un circuito RC: teóricamente para descargarlo totalmente hace falta infinito tiempo). Sin embargo, los marcados que es interesante alcanzar en estado estacionario son normalmente marcados estrictamente positivos (si el marcado de un lugar es 0, el flujo de sus transiciones de salida es 0, lo que significa inactividad total) y éstos se alcanzan en tiempo finito.

Propiedad 3. (Mahulea *et al.* 2007) Sea un sistema $\langle \mathcal{N}, \boldsymbol{\lambda}, \mathbf{m}_0 \rangle$ tal que todas las transiciones son controlables. Si \mathbf{m} es alcanzable en el modelo autónomo, entonces \mathbf{m} es alcanzable en tiempo finito en el modelo temporizado controlado.

Obsérvese que bajo las condiciones del Teorema 2 \mathbf{m} es alcanzable en el modelo autónomo si y sólo si verifica la ecuación fundamental.

En control de producción, la función de beneficio a menudo depende de la producción, del trabajo en proceso y de la amortización de inversiones. Asumiendo que la función es lineal, y que las máquinas han sido fijadas (es decir $\boldsymbol{\lambda}$ está definido), la función objetivo tiene la siguiente forma: $\mathbf{w} \cdot \boldsymbol{\varphi} - \mathbf{z} \cdot \mathbf{m} - \mathbf{q} \cdot \mathbf{m}_0$, donde $\boldsymbol{\varphi}$ es el flujo de producción en régimen permanente, y \mathbf{w} es un vector que define el *beneficio* que produce cada producto; \mathbf{m} es el marcado medio, y \mathbf{z} es el coste unitario del inventario inmovilizado en proceso por unidad de tiempo; finalmente \mathbf{q} representa depreciación o amortización de las inversiones iniciales sobre \mathbf{m}_0 .

Sea el siguiente problema de programación lineal

$$\begin{aligned} \max: & \mathbf{w} \cdot \mathbf{f} - \mathbf{z} \cdot \mathbf{m} - \mathbf{q} \cdot \mathbf{m}_0 \\ \text{s.a.} & \mathbf{C} \cdot \mathbf{f} = 0, \mathbf{f} \geq 0 \end{aligned} \quad (10a)$$

$$\mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \boldsymbol{\sigma}, \mathbf{m}, \boldsymbol{\sigma} \geq 0 \quad (10b)$$

$$\begin{aligned} \mathbf{f}[t_i] &= \boldsymbol{\lambda}[t_i] \cdot \left(\frac{\mathbf{m}[p_j]}{\text{Pre}[p_j, t_i]} \right) - \mathbf{v}[p_j, t_i], \\ \forall p_j \in \bullet t_i, \mathbf{v}[p_j, t_i] &\geq 0 \end{aligned} \quad (10c)$$

donde $\mathbf{v}[p_j, t_i]$ son variables de *holgura*. Las ecuaciones representan: (a) \mathbf{f} es un T-semiflujo; (b) necesario para que \mathbf{m} sea alcanzable (ecuación fundamental); (c) regla de disparo de las transiciones controladas con semántica de infinitos servidores.

Teorema 14. (Mahulea *et al.* 2007) Sea un sistema $\langle \mathcal{N}, \lambda, \mathbf{m}_0 \rangle$ y $\langle \mathbf{f}, \mathbf{m}, \mathbf{v} \rangle$ una solución del problema de programación lineal (10). Para cada transición t_i , sea $\mathbf{u}[t_i] = \min_{p_j \in \bullet t_i} \mathbf{v}[p_j, t_i]$ su entrada de control en estado estacionario para \mathbf{m} . (Si $|\bullet t_i| = 1$ la variable de holgura coincide con la entrada de control.) Si para cada $\mathbf{u}[t_i] > 0$ la transición t_i es controlable, entonces \mathbf{u} es un control óptimo en estado estacionario.

Obsérvese que se trata de un problema de programación lineal, y por lo tanto tiene solución con complejidad polinomial.

5.3 Control del sistema dinámico

Una aproximación MPC. El sistema de (9) es lineal a tramos y tiene una restricción sobre la entrada de control \mathbf{u} que depende del estado del sistema \mathbf{m} . Sin embargo, definiendo $\mathbf{w} = \mathbf{f} - \mathbf{u}$, se puede reescribir como un sistema lineal, eso sí, con restricciones dinámicas adicionales sobre las entradas de control:

$$\begin{cases} \dot{\mathbf{m}}(\tau) = \mathbf{C} \cdot \mathbf{w}(\tau) \\ \mathbf{G} \cdot \begin{bmatrix} \mathbf{w}(\tau) \\ \mathbf{m}(\tau) \end{bmatrix} \leq \mathbf{0} \\ \mathbf{w}(\tau) \geq \mathbf{0} \end{cases} \quad (11)$$

donde $\mathbf{G} = [\mathbf{\Delta} \quad -\mathbf{\Gamma}]$ con $\mathbf{\Delta}$ ($q \times n$) y $\mathbf{\Gamma}$ ($q \times m$) y $q = \sum_{t \in T} |\bullet t|$, es decir el número de arcos “pre” en la red. Dado un arco “pre” (p_i, t_j) la fila correspondiente de $\mathbf{\Delta}$ es el vector

$$\left[\underbrace{0 \quad \dots \quad 0 \quad 1}_{j} \quad 0 \quad \dots \quad 0 \right],$$

mientras que la fila correspondiente de $\mathbf{\Gamma}$ es

$$\left[0 \quad \dots \quad 0 \quad \underbrace{\lambda_j \quad \text{Pre}[p_i, t_j]}_i \quad 0 \quad \dots \quad 0 \right].$$

En (Mahulea *et al.* 2006a) se ha planteado el problema de calcular la entrada que lleva al sistema desde un marcado inicial hasta un marcado final, optimizando cierta función de coste para sistemas modelados con redes de Petri continuas con semántica de infinitos servidores. Se han investigado dos soluciones, basadas en “Model Predictive Control” (MPC) (Bemporad *et al.* 2000) implícito y explícito.

Por simplicidad de cálculo, el problema se plantea sobre el sistema discretizado en el tiempo, sin embargo esto plantea una dificultad: la elección del periodo de muestreo. El muestreo debería preservar la información importante del modelo original (en particular la no negatividad de los marcados).

Propiedad 4. (Mahulea *et al.* 2006a) Si el periodo de muestreo verifica que

$$\forall p \in P, \sum_{t_j \in p^\bullet} \lambda[t_j] \Theta < 1, \quad (12)$$

el conjunto de marcados alcanzables en el sistema discretizado coincide con los marcados alcanzables en el sistema continuo temporizado.

En la práctica, la frecuencia de muestreo puede tener que ser mayor si se necesita reconstruir las señales (como en el teorema de muestreo de Nyquist-Shannon), pero ese es un tema que no se abordará aquí.

Si se pretende minimizar un índice de rendimiento cuadrático estándar que mida la distancia desde el estado actual (marcado y entrada) al estado que se desea alcanzar:

$$\begin{aligned} J(\mathbf{m}(k), N) = & (\mathbf{m}(k+N) - \mathbf{m}_f)' \cdot \mathbf{Z} \cdot (\mathbf{m}(k+N) - \mathbf{m}_f) + \\ & \sum_{j=0}^{N-1} [(\mathbf{m}(k+j) - \mathbf{m}_f)' \cdot \mathbf{Q} \cdot (\mathbf{m}(k+j) - \mathbf{m}_f) + \\ & (\mathbf{w}(k+j) - \mathbf{w}_f)' \cdot \mathbf{R} \cdot (\mathbf{w}(k+j) - \mathbf{w}_f)] \end{aligned} \quad (13)$$

donde \mathbf{Z} , \mathbf{Q} y \mathbf{R} han de ser matrices definidas positivas.

Añadiendo las restricciones del sistema, en cada paso el problema que hay que resolver es:

$$\begin{aligned} \min & J(\mathbf{m}(k), N) \\ \text{s.a. :} & \mathbf{m}(k+j+1) = \mathbf{m}(k+j) + \Theta \cdot \mathbf{C} \mathbf{w}(k+j), \\ & \mathbf{G} \cdot \begin{bmatrix} \mathbf{w}(k+j) \\ \mathbf{m}(k+j) \end{bmatrix} \leq \mathbf{0}, \\ & \mathbf{w}(k+j) \geq \mathbf{0}, \\ & \text{para } j = 0, \dots, N-1 \end{aligned} \quad (14)$$

El ejemplo (sistema no monótono) sobre el que se va a trabajar, es el representado en la Figura 26, siendo la velocidad de las transiciones $\lambda = [1 \ 1 \ 1]$. Supóngase que el objetivo es llevar al sistema desde el estado inicial $\mathbf{m}_0 = [3 \ 3 \ 1 \ 3]$ a un estado final $\mathbf{m}_f = [2,50 \ 3,25 \ 1,25 \ 2,50]$ con $\mathbf{w}_f = [1,25 \ 1,25 \ 1,25]$. Obsérvese que aunque el sistema tiene 4 lugares, es en realidad un sistema de orden 2 debido a la existencia de 2 P-semiflujos: $\mathbf{m}[p_1] + \mathbf{m}[p_2] + \mathbf{m}[p_3] = 7$ y $\mathbf{m}[p_1] + 4\mathbf{m}[p_3] + \mathbf{m}[p_4] = 10$. Sean $\mathbf{R} = 0,01\mathbf{I}$, $\mathbf{Q} = \mathbf{I}$ y $\mathbf{Z} = 100\mathbf{I}$.

El Cuadro 3 resume los resultados numéricos obtenidos en el caso de MPC implícito con periodos de muestreo $\Theta = 0,1, 0,05, 0,01$, todos satisfaciendo la desigualdad (12), y distintos valores de N . La columna \bar{J} representa el coste

$$\bar{J}(\mathbf{m}_0, \Theta) = \Theta \cdot \sum_{j=0}^{\infty} [(\mathbf{m}(j) - \mathbf{m}_f)' \cdot \mathbf{Q} \cdot (\mathbf{m}(j) - \mathbf{m}_f) + (\mathbf{w}(j) - \mathbf{w}_f)' \cdot \mathbf{R} \cdot (\mathbf{w}(j) - \mathbf{w}_f)]$$

Se indica el tiempo medio en segundos necesario para resolver cada paso en un Intel Pentium 4 a 3.20 GHz.

Se observa que el coste \bar{J} es prácticamente el mismo para $N = 10$ y $N = 20$, es decir, no es necesario incrementar mucho el horizonte. Por otra parte, aunque para $\Theta = 0,1$ todas las soluciones se pueden implementar en este ordenador, esto no es cierto para valores menores de Θ , ya que el tiempo necesario para resolver el problema en cada paso es mayor que el periodo de muestreo si N es grande.

Las simulaciones se han realizado también para el planteamiento del MPC explícito. La partición del espacio de estados que se obtiene para $\Theta = 0,1$ y $N = 10$ se ha representado en la Figura 34. En el Cuadro 3, las columnas 4, 8 y 12 resumen el número de regiones politopos n_P .

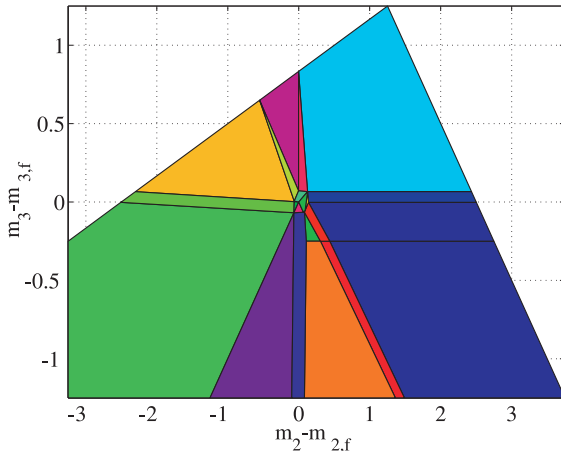


Figura 34. Partición del espacio de estados para el sistema de la Figura 26 con $N = 1$ y $\Theta = 0,05$.

En sistemas generales, no hay garantía de que el problema a resolver sea factible en todo momento, o en términos de MPC explícito, no hay garantía de que la partición resultante incluya todos los estados factibles. Sin embargo, debido a la forma particular de las restricciones en el caso de las redes de Petri continuas se puede demostrar que la factibilidad se mantiene (Mahulea *et al.* 2006a).

Aproximación baja-alta ganancia. En (Xu *et al.* 2006) se ha estudiado el seguimiento de sistemas modelados con redes Sin Sincronización utilizando técnicas “baja-y-alta ganancia” (Saberi *et al.* 1996). En la actualidad se termina su extensión a redes cualesquiera.

Figura 35. Red Sin Sincronización con $\lambda = [1 \ 1 \ 2 \ 1 \ 2]$.

El objetivo es conseguir que el marcado y la acción converjan a los valores deseados. Aunque existe abundante literatura sobre el uso de este tipo de técnicas (Lien 1998, Turner *et al.* 2000, Chen *et al.* 2003), una hipótesis habitual es que los límites superiores e inferiores de las acciones son constantes positivas y negativas. Sin embargo, en los sistemas modelados con redes de Petri, la acción no puede ser negativa, y la cota superior no es constante, sino que depende del estado, como se ha visto.

En primer lugar, se utiliza la teoría LQ para diseñar un controlador de “baja-ganancia”, de manera que se satisfagan las restricciones. Debido a las restricciones sobre la acción, no se puede garantizar en general la convergencia global asintótica para referencias en escalón. Para poder asegurar la convergencia, evitando al mismo tiempo cambios bruscos en la trayectoria, se ha optado por utilizar como referencia un primer escalón, seguido de una rampa hasta llegar al valor deseado. Así, el problema se divide en dos partes, primero la generación de la trayectoria, y segundo su seguimiento.

A continuación, se añade una parte de “alta-ganancia” para utilizar mejor las acciones disponibles, y hacer que la convergencia sea más rápida. Se ha demostrado la convergencia asintótica global del algoritmo de control para redes Sin Sincronización conservativas y fuertemente conexas.

Por ejemplo, sea la red de la Figura 35 con marcado inicial $\mathbf{m}_0 = [2 \ 3 \ 5 \ 6 \ 9]$ y marcado objetivo $\mathbf{m}_d = [3 \ 5 \ 2 \ 7 \ 8]$. Para maximizar el flujo en el estado final, la acción de control objetivo es $\mathbf{u}_d = [1, \ 3 \ 0 \ 5 \ 6]$.

La Figura 36 ilustra la convergencia de los marcados con el control diseñado, y la Figura 37 muestra las entradas de control. Obsérvese que la acción calculada siempre respeta las restricciones.

Tabla 3. Resultados del control MPC implícito para el sistema de la Figura 26

$\Theta = 0,1$				$\Theta = 0,05$				$\Theta = 0,01$			
N	\bar{J}	comput. time [sec]	n_P	N	\bar{J}	comput. time [sec]	n_P	N	\bar{J}	comput. time [sec]	n_P
1	0,3409	0,0363	26	1	0,2359	0,0364	26	1	0,0787	0,0372	12
2	0,1794	0,0384	-	2	0,1795	0,0380	98	2	0,0787	0,0380	41
10	0,0851	0,0469	888	10	0,0822	0,0468	1436	10	0,0784	0,0449	651
20	0,0846	0,0614	2635	20	0,0810	0,0616	3202	20	0,0782	0,0562	2489

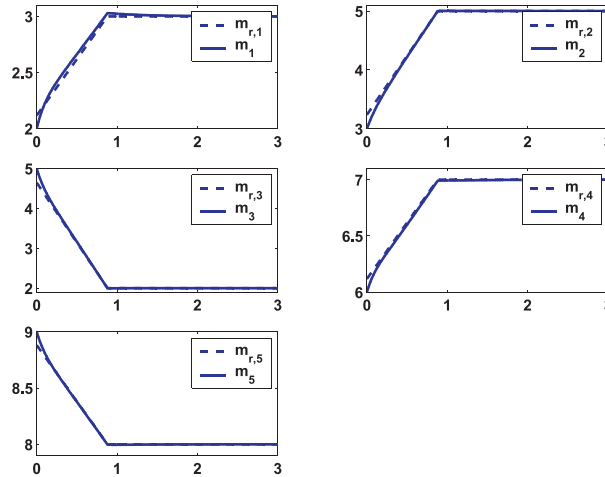


Figura 36. Convergencia de los marcados con el control definido.

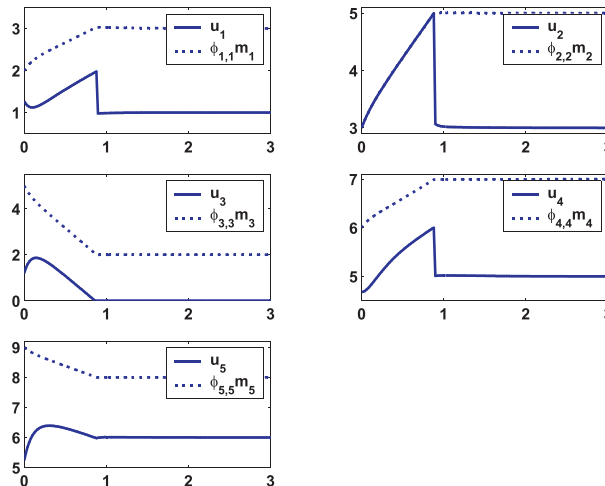


Figura 37. Las señales de control que se obtienen, y las cotas definidas por el marcado en cada instante.

En la actualidad la aproximación se ha generalizado para redes con sincronizaciones, al tiempo que se sabe cómo planificar trayectorias poligonales que reduzcan el tiempo para llegar a \mathbf{m}_d . En el futuro se abordará la comparación entre las técnicas de “baja-y-alta ganancia” y MPC.

6. TOMANDO DISTANCIA

Una comparación recapitulativa de lo aquí presentado con la aproximación hoy más clásica basada en los Diagramas de Forrester (DdF) permite situar nuestra contribución basada en redes de Petri.

El primer punto a evocar es una coincidencia epistemológica importante: ambas clases de formalismos son singularmente apropiadas para modelar sistemas de producción/consumo; de donde resulta “natural” que sean bipartidos, alternándose elementos de almacenamiento (depósitos ó lugares) con los de establecimiento de flujos (válvulas ó transiciones).

A partir de lo dicho, las diferencias son sustanciales. Por un lado los DdF definen “directamente” modelos fluidificados (o continuos) y temporizados, separando flujos de materia y de control que, sea dicho, se expresan con funciones de una generalidad absoluta. Con las redes de Petri se parte de una indiferenciación entre los flujos de materia y control, y funciones de flujo relativamente simples

(las que corresponden a las semánticas mencionadas de finitos e infinitos servidores); recientemente se ha demostrado que la diferenciación de los flujos de control y materia en las RdP continuas no añade potencia descriptiva, aunque a veces pueda facilitar la “lectura” de algunos modelos (Recalde *et al.* 2007).

En el marco petrista es posible la existencia de modelos (siempre “visiones” de las realidades que se consideran) no temporizados, denominados *autónomos*, lo que permite su consideración bajo hipótesis de comportamiento *no determinista*, entendido éste en su más amplio sentido, más allá de asumir comportamientos estocásticos: todo comportamiento posible, por inverosímil que pueda parecer, se puede producir. Esta hipótesis, frecuente en informática, sobre todo en sistemas distribuidos, es una posición “de humildad” frente a la impredecibilidad de los comportamientos reales en sistemas complejos. Por otro lado, en el marco del paradigma de RdP se pueden —al menos en parte— contrastar propiedades de modelos fluidificados frente a las “visiones” (de eventos) discretas. Así, merece la pena recordar que los modelos clásicos de Volterra-Lotka en dinámica de poblaciones evolucionan según una órbita, cuando se sabe que con certeza, bajo hipótesis estocásticas generales, que los sistemas discretos subyacentes tienden a cero depredadores e infinitas presas, o cero presas y, por consiguiente, también cero depredadores (Nicolis y Prigogine 1971). En este caso, la RdP autónoma (discreta o continua) es modelo con-bloqueos y no-limitado, mientras que la relajación continua y temporizada es sin-bloqueos y sí-limitada. La aproximación de Forrester se limita al modelo fluidificado y temporizado, mientras que la aquí adoptada tiene siempre otros posibles puntos de vista en el mismo paradigma de modelado. Los formalismos de red continua y temporizada, bajo las semánticas de finitos e infinitos servidores son técnicamente híbridos (sistemas lineales conmutados); además la primera de estas semánticas corresponde a sistemas conceptualmente híbridos (los “pocos” servidores son considerados de forma discreta, mientras que los “muchos” clientes son fluidificados). Si se adopta la semántica con productos de poblaciones se obtienen sistemas no-lineales que pueden simular fenómenos caóticos. La semántica de infinitos servidores conduce a modelos que permiten simular máquinas de Turing, con lo que ello acarrea de indecidibilidades. Lo aquí presentado representa la urdimbre de la aproximación, donde hay que reconocer la existencia de tres tramas complementarias: la visión autónoma (no-determinista) de los sistemas, de raíz informática; la evaluación de prestaciones de los modelos temporizados no forzados, que releva en gran parte de la investigación de operaciones; y las cuestiones relativas a la

observación y control, propias de la teoría de control. En todos los casos, nuestro interés básico es la comprensión de las propiedades de los sistemas, para lo que el énfasis se hace en el estudio de las relaciones estructura-comportamiento, más que la exploración por simulación de casos particulares. En otros términos, nuestra aproximación es epistemológicamente próxima, pero técnicamente aún muy distante, de la “no-ortodoxa” en Dinámica de Sistemas “a la Forrester” de (Mosekilde *et al.* 1988), al estudiar la dinámica de sistemas no lineales.

Ante la posible impresión de encontrarnos frente a una teoría bastante “hecha”, hemos de reconocer que, “el camino se hace al andar” y aún estamos muy lejos de lo deseado. Valga la enumeración de algunos problemas abiertos para transmitir esa certeza, al tiempo que esperamos se consideren las evocaciones que siguen como una invitación a trabajar en estas líneas, con las múltiples variantes que evidentemente se pueden establecer:

- Dado un sistema (de eventos) discreto concreto. La descripción con RdP, ¿es “razonablemente” fluidificable?
- ¿Qué política de servicio es la más adecuada: semántica de finitos o de infinitos servidores? ¿Otra semántica alternativa?
- Dada una semántica (temporizada) de servicios: ¿existe un régimen permanente? (como se ha dicho, este problema es en general indecidible para semántica de infinitos servidores (Haddad *et al.* 2006, Recalde *et al.* 2007).
- Se dispone de resultados razonables para decidir la alcanzabilidad en redes autónomas (no temporizadas). Cuando el modelo se temporiza, la cuestión se torna más peliaguda: ¿Cómo mejorar la caracterización de los marcados alcanzables en régimen permanente?
- Incluso si para problemas de diseño fuera-de-línea existen resultados interesantes, las cuestiones de control dinámico requieren de esfuerzos adicionales, particularmente en el entrelazamiento de observación y control.
- En sistemas lineales conmutados, se sabe que varios sistemas estables (con valores propios con parte real negativa) pueden dar un comportamiento inestable. Sin haber probado su imposibilidad en el caso de redes con semántica de infinitos servidores, no hemos sabido encontrar un ejemplo. Hace falta, pues una demostración o un contraejemplo que tenga sentido como sistema de producción/consumo.
- Asumiendo la existencia de una ley para controlar eficientemente el sistema continuo, y si ello fuese necesario, ¿cómo volver al ámbito de los modelos discretos?

AGRADECIMIENTOS

Han contribuido con sus tesis doctorales a esta línea de investigación Jorge Júlvez (actualmente en la UPC, Barcelona) y Cristian Mahulea. Además, han participado en el desarrollo de puntos precisos de la presente aproximación: Alessandro Giua y Carla Seatzu (U. de Cagliari), Serge Haddad (U. de Paris IX), Emilio Jiménez (U. de La Rioja), Antonio Ramírez Treviño (Cinvestav, Guadalajara, MX), Jing Xu (U. de Singapur) y Enrique Teruel. Por razones de espacio, en la necesaria selección de temas realizada no se mencionan las importantes colaboraciones con Alberto Bemporad (U. de Siena) y René Boel (U. de Gante). A todos, nuestra gratitud y afecto.

REFERENCIAS

- Ajmone Marsan, M., G. Balbo, G. Conte, S. Donatelli y G. Franceschinis (1995). *Modelling with Generalized Stochastic Petri Nets*. Wiley.
- Alla, H. y R. David (1998). Continuous y hybrid Petri nets. *Journal of Circuits, Systems, and Computers* **8**(1), 159–188.
- Bemporad, A., F.D. Torrisi y M. Morari (2000). Performance analysis of piecewise linear systems and model predictive control systems. In: *Proc. of the 39th IEEE Conference on Decision and Control*. Sydney, Australia. pp. 4957–4962.
- Brams, G. W. (1983). *Réseaux de Petri: Théorie et Pratique*. Masson.
- Champagnat, R., P. Esteban, H. Pingaud y R. Valette (1998). Modeling and simulation of a hybrid system through Pr/Tr PN-DAE model. In: *Proc. of ADPM'98*. Reims, France. pp. 131–137.
- Champagnat, R., R. Valette, J.C. Hochon y H. Pingaud (2001). Modeling, simulation and analysis of batch production systems. *Discrete Event Dynamic Systems: Theory and Application* **11**(1/2), 119–136.
- Chen, B.M., T.H. Lee, K. Peng y V. Venkataraman (2003). Composite nonlinear feedback control for linear systems with input saturation: theory and an application. *IEEE Trans. on Automatic Control* **48**(3), 427–439.
- Chen, H. y D.D. Yao (2001). *Fundamentals of Queueing Networks. Performance, Asymptotics and Optimization*. Vol. 46 of *Applications of Mathematics. Stochastic Modelling and Applied Probability*. Springer.
- David, R. y H. Alla (1992). *Petri Nets and Graphet*. Prentice-Hall.
- David, R. y H. Alla (2004). *Discrete, Continuous, and Hybrid Petri Nets*. Springer-Verlag.
- Demongodin, I. y N. T. Koussoulas (1998). Differential Petri nets: Representing continuous systems in a discrete-event world. *IEEE Trans. on Automatic Control (Special Issue)* **43**(4), 573–579.
- Dicesare, F., G. Harhalakis, J. M. Proth, M. Silva y F. B. Vernadat (1993). *Practice of Petri Nets in Manufacturing*. Chapman & Hall.
- Haddad, S., L. Recalde y M. Silva (2006). On the computational power of Timed Differentiable Petri nets. In: *Formal Modeling and Analysis of Timed Systems, 4th Int. Conf. FORMATS 2006* (E. Asarin y P. Bouyer, Eds.). Vol. 4202 of *LNCS*. Springer. Paris. pp. 230–244.
- Horton, G., V.G. Kulkarni, D.M. Nicol y K.S. Trivedi (1998). Fluid stochastic Petri nets: Theory, applications, and solution techniques. *European Journal of Operational Research* **105**, 184–201.
- Jensen, K. (1997). *Coloured Petri Nets: Basic Concepts, Analysis Methods, and Practical Use*. EATCS Monographs on Theoretical Computer Science. Springer.
- Jiménez, E., J. Júlvez, L. Recalde y M. Silva (2004). Relaxed continuous views of discrete event systems: considerations in Forrester diagrams and Petri nets. In: *Proc. of the Int. Conf. on Systems, Man and Cybernetics (SMC 2004)*. The Hague, The Netherlands. pp. 4897–4904.
- Júlvez, J., E. Jiménez, L. Recalde y M. Silva (2004). On observability in timed continuous Petri net systems. In: *Proc. of the 1st Int. Conf. on the Quantitative Evaluation of Systems (QEST 2004)* (J.P. Katoen G. Franceschinis and M. Woodside, Eds.). IEEE Computer Society Press. Enschede, The Netherlands. pp. 60–69.
- Júlvez, J., L. Recalde y M. Silva (2003). On reachability in autonomous continuous Petri net systems. In: *Proc. of the 24th Int. Conf. on Application and Theory of Petri Nets (ICATPN 2003)* (W. van der Aalst y E. Best, Eds.). Vol. 2679. pp. 221–240. Springer. Eindhoven, The Netherlands.
- Júlvez, J., L. Recalde y M. Silva (2005). Steady-state performance evaluation of continuous mono-t-semiflow Petri nets. *Automatica* **41**(4), 605–616.
- Júlvez, J., L. Recalde y M. Silva (2006). Deadlock-freeness analysis of continuous mono-t-semiflow petri nets. *IEEE Trans. on Automatic Control* **51**(9), 1472–1481.
- Lien, Y. E. (1998). Toward improvement of tracking performance - nonlinear feedback for linear systems. *International Journal of Control* **70**(1), 1–11.
- Mahulea, C., A. Giua, L. Recalde, C. Seatzu y M. Silva (2006a). On sampling continuous timed Petri nets: reachability “equivalence” under infinite servers semantics. In: *2nd IFAC Conference on Analysis and Design of Hybrid Systems, ADHS06*. Alghero, Italy. pp. 37–43.

- Mahulea, C., A. Ramirez-Trevino, L. Recalde y M. Silva (2007). Steady state control reference and token conservation laws in continuous Petri net systems. *IEEE Trans. on Automation Science and Engineering*. Aceptado para su publicación.
- Mahulea, C., L. Recalde y M. Silva (2006b). On performance monotonicity and basic servers semantics of continuous Petri nets. In: *8th Int. Workshop on Discrete Event Systems WODES06*. IEEE Computer Society Press. Ann Arbor, USA. pp. 345–351.
- Mandelbau, A. y H. Chen (1991). Discrete flow networks: Bottleneck analysis and fluid approximations. *Mathematical Operations Research* **16**, 408–446.
- Molloy, M. K. (1982). Performance analysis using stochastic Petri nets. *IEEE Trans. on Computers* **31**(9), 913–917.
- Mosekilde, E., J. Aracil y P.M. Allen (1988). Instabilities and chaos in nonlinear dynamic systems. *Systems Dynamics Review* **4**(1-2), 14–55.
- Murty, K. G. (1983). *Linear Programming*. Wiley and Sons.
- Nicolis, G. y I. Prigogine (1971). Fluctuations in nonequilibrium systems. *Proc. Nat. Acad. Sci. U.S.A.* **68**, 2102–2107.
- Peterson, J. L. (1981). *Petri Net Theory and the Modeling of Systems*. Prentice-Hall.
- Recalde, L. y M. Silva (2001). Petri Nets fluidification revisited: Semantics and steady state. *European Journal of Automation APII-JESA* **35**(4), 435–449.
- Recalde, L., E. Teruel y M. Silva (1999). Autonomous continuous P/T systems. In: *Application and Theory of Petri Nets 1999* (S. Donatelli y J. Kleijn, Eds.). Vol. 1639 of *LNCS*. Springer. pp. 107–126.
- Recalde, L., S. Haddad y M. Silva (2007). Continuous Petri nets: Expressive power and decidability issues. Research report. Dep. Informática e Ingeniería de Sistemas, Universidad de Zaragoza. María de Luna, 1, 50018 Zaragoza, Spain.
- Saberi, A., Z. Lin y A.R. Teel (1996). Control of linear system with saturating actuators. *IEEE Trans. on Automatic Control* **41**(3), 368–378.
- Silva, M. (1993). Introducing Petri nets. pp. 1–62. in *Practice of Petri Nets in Manufacturing* (Dicesare et al. 1993).
- Silva, M. (2002). *Las Redes de Petri: en la Automática y la Informática*. AC/Thomson.
- Silva, M. (2006a). Discrete event systems and life cycle: A Petri nets-based view. In: *Procs. of the Int. Mediterranean Multiconference, I3M2006* (A. Bruzzone, A. Guasch, M. A. Piera y J. Rozenblit, Eds.). Barcelona. pp. 5–12.
- Silva, M. (2006b). Terminología. *Revista Iberoamericana de Automática e Informática Industrial* **3**(2), 122–123.
- Silva, M. y E. Teruel (1996). A systems theory perspective of discrete event dynamic systems: The Petri net paradigm. In: *Symposium on Discrete Events and Manufacturing Systems. CESA '96 IMACS Multiconference* (P. Borne, J. C. Gentina, E. Craye y S. El Khattabi, Eds.). Lille, France. pp. 1–12.
- Silva, M. y E. Teruel (1998). DEDS along their life-cycle: interpreted extension of Petri nets. In: *IEEE International Conference on Systems, Man and Cybernetics (SMC'98)*. IEEE. SAn Diego, USA.
- Silva, M. y L. Recalde (2002). Petri nets and integrality relaxations: A view of continuous Petri nets. *IEEE Trans. on Systems, Man, and Cybernetics* **32**(4), 314–327.
- Silva, M. y L. Recalde (2003). Unforced continuous Petri nets and positive systems. In: *Positive Systems. Proceedings of the First Multidisciplinary International Symposium on Positive Systems: Theory and Applications (POSTA 2003)* (Luca Benvenuti, Alberto De Santis and Lorenzo Farina, Eds.). Vol. 294 of *LNCIS*. Springer. Rome, Italy. pp. 55–62.
- Silva, M. y L. Recalde (2004). On fluidification of Petri net models: from discrete to hybrid and continuous models. *Annual Reviews in Control* **28**(2), 253–266.
- Silva, M., E. Teruel y J. M. Colom (1998). Linear algebraic and linear programming techniques for the analysis of net systems. In: *Lectures in Petri Nets. I: Basic Models* (G. Rozenberg y W. Reisig, Eds.). Vol. 1491 of *LNCS*. pp. 309–373. Springer.
- Sussmann, H. J. y J. C. Willems (1997). 300 years of optimal control: from the brachystochrone to the maximum principle. *IEEE Control Systems Magazine* **17**(3), 32–44.
- Turner, M.C., I. Postlethwaite y D.J. Walker (2000). Non-linear tracking control for multivariable constrained input linear systems. *International Journal of Control* **73**(12), 1160–1172.
- Valk, R. y Girault, C., Eds. (2003). *Petri Nets for Systems Engineering. A Guide to Modeling, Verification, and Applications*. Springer.
- Xu, J., L. Recalde y M. Silva (2006). Tracking control of join-free timed continuous Petri net systems. In: *2nd IFAC Conf. on Analysis and Design of Hybrid Systems, ADHS06*. Alghero, Italy. pp. 30–36.