

UNIVERSIDAD NACIONAL DE ROSARIO
FACULTAD DE CIENCIAS EXACTAS, INGENIERIA Y
AGRIMENSURA

I.S.S.N. 03260690

CUADERNOS

DEL

INSTITUTO DE MATEMATICA "BEPPO LEVI"

**ANALISIS Y CONTROL
DE SISTEMAS DE
EVENTOS DISCRETOS:
DE REDES DE PETRI
TEMPORIZADAS AL ALGEBRA**

Guy COHEN

ENPC & INRIA (Francia)

29

Rosario - República Argentina

- 2001 -

UNIVERSIDAD NACIONAL DE ROSARIO
FACULTAD DE CIENCIAS EXACTAS, INGENIERIA Y AGRIMENSURA

I.S.S.N. 03260690

CUADERNOS

DEL

INSTITUTO DE MATEMATICA "BEPPPO LEVI"

ANALISIS Y CONTROL DE SISTEMAS DE EVENTOS DISCRETOS: DE REDES DE PETRI TEMPORIZADAS AL ALGEBRA

Guy COHEN

École Nationale des Ponts et Chaussées

et

Institut National de Recherche en Informatique et en Automatique

Profesor visitante - Rosario - Abril 2000

Lecciones redactadas por:

Lic. Ricardo KATZ

Ing. Ernesto KOFMAN

Rosario - República Argentina

- 2001 -

Prólogo

Este curso intensivo dictado por el Prof. Guy Cohen en la Facultad de Ciencias Exactas, Ingeniería y Agrimensura de la Universidad Nacional de Rosario (FCEIA-UNR) en abril de 2000 fue precedido por el que dictara en noviembre de 1988 el Prof. Jean-Pierre Quadrat sobre “*Semi-anillos en Matemática Aplicada*”, que fuera publicado como el No. 28 de esta Serie CUADERNOS.

Ambos forman parte del ciclo de cuatro cursos intensivos previstos en el marco del Programa de Capacitación e Iniciación en la Investigación sobre “Sistemas de Eventos Discretos y Redes de Comunicación” que vincula al INRIA (Institut National de Recherche en Informatique et en Automatique) de Francia con el Instituto de Matemática “Beppo Levi” de la FCEIA-UNR. Es con satisfacción que señalamos que jóvenes investigadores que han participado en varias actividades del Programa intervienen desde comienzos de 1999 en un proyecto de investigación y desarrollo que el equipo METALAU del INRIA tiene a su cargo, tarea que incluye el estudio de casos concretos en sistemas de transporte, lo que ya ha dado lugar a la publicación de varios artículos en revistas especializadas. Por otra parte miembros de dicho equipo han asumido la dirección de una tesis de doctorado en matemáticas actualmente en curso en nuestra Facultad.

Deseamos finalmente agradecer a todos quienes han permitido con su dedicación y apoyo la materialización de la etapas ya cumplidas, como aquéllas en curso, del Programa y, en particular, al CONICET, a la Fundación Antorchas, al Ministère des Affaires Etrangères de Francia, al ENPC, al INRIA y al Proyecto Alapedes del Programa Europeo TMR.

Consejo Científico del Instituto de Matemática “Beppo Levi”

Ce mémoire est une introduction à la théorie des systèmes à événements discrets que l’on peut modéliser comme des systèmes dynamiques linéaires grâce à l’utilisation d’une nouvelle classe d’algèbres, dont l’algèbre max-plus est un prototype simple. Ces systèmes constituent en fait une sous-classe de réseaux de Petri temporisés désignés comme des “graphes d’événements” et dont le mécanisme dynamique de base est constitué de phénomènes de synchronisation entre processus concurrents. Pour moi, et pour mon ami Jean-Pierre Quadrat, auteur du cours précédent dans la série de cours autour de ces thèmes donnés à l’Université de Rosario, la découverte de ces systèmes et de ces outils algébriques remonte au début des années 80, et ces thèmes n’ont cessé de nous occuper depuis lors. À

ces travaux ont été associés depuis le début, et au fur et à mesure, de nombreux collègues qu'il serait sans doute trop long de mentionner tous ici.¹ Je ferai une exception pour Stéphane Gaubert qui s'est joint à nous depuis de nombreuses années. Les contributions de tous ces collègues font aujourd'hui partie intégrante d'une théorie qui a fini par acquérir ses lettres de noblesse et qui a rejoint, au plan mondial, d'autres courants de pensée et d'autres sources d'inspiration qui ont tous en commun l'utilisation de ces outils algébriques marqués par l'idempotence de l'addition.

Jean-Pierre, Stéphane et moi-même avons été souvent sollicités pour donner, dans un bref exposé ou dans une série de cours, un aperçu de cette théorie et de ses ramifications. Nous l'avons toujours fait avec le plaisir que l'on peut trouver à enseigner des choses simples mais ayant, dans leur simplicité, une certaine esthétique. Ce fut en particulier le cas pour moi à Rosario pour cette raison, mais aussi pour les raisons suivantes :

- le climat particulièrement studieux et chaleureux créé par ce public avec lequel j'ai été en contact là-bas pendant deux semaines ;
- les à-côtés touristiques qui ont entouré ce voyage (et notamment une semaine qui a précédé le cours lui-même, semaine consacrée entièrement à la découverte du Nord-Ouest de l'Argentine) qui ont gravé dans la mémoire de ma compagne et dans la mienne (et aussi sur divers supports vidéo et informatiques) des images inoubliables de ce pays attachant ;
- et enfin, *last but not least*, la présence constante d'Edmundo Rofman, qui, avant, pendant, et après ce voyage et ce cours, et jusqu'à sa matérialisation dans ce mémoire final, nous a entourés de toute son attention bienveillante.

Je dois aussi remercier les rédacteurs de ce mémoire, Ricardo Katz et Ernesto Kofman qui ont pallié ma méconnaissance quasi-complète de la langue hispanique et ont ajouté leur touche personnelle à ce mémoire, ainsi qu'Elina Mancinelli et Pablo Lotito qui ont eux-mêmes pris part au cours oral en l'illustrant de leurs propres contributions.

Guy Cohen

CERMICS
École Nationale des Ponts et Chaussées
6–8, avenue Blaise Pascal
Cité Descartes, Champs-sur-Marne
77455 MARNE-LA-VALLÉE cedex 2

Institut National de Recherche
en Informatique et en Automatique
Domaine de Voluceau, Rocquencourt
B.P. 105
78153 LE CHESNAY cedex

`guy.cohen@mail.enpc.fr`

¹Mais Jean-Pierre l'a quand même fait dans le texte qu'il a placé au début du Cuaderno No. 28 et leurs noms apparaissent évidemment dans les listes des publications.

Indice General

1. Introducción a las redes de Petri	1
1.1. Componentes básicos	1
1.1.1. Punto de vista de la teoría de grafos	1
1.1.2. Punto de vista del álgebra lineal	2
1.1.3. Semántica de las redes de Petri	4
1.2. Ventajas y desventajas de la representación “lineal” clásica	4
1.2.1. Desventajas	4
1.2.2. Invariantes.	5
1.2.2.1. Invariantes de lugares	5
1.2.2.2. Invariantes de transiciones	5
1.2.3. Otros conceptos y propiedades	6
1.2.3.1. Conflictos	6
1.2.3.2. Paralelismo	6
1.2.3.3. Vivacidad	7
1.2.3.4. Acotación, seguridad	7
1.2.3.5. Marcas alcanzables	8
1.3. Introducción del tiempo y comportamiento sistemático	8
1.3.1. Tiempos asociados a lugares y transiciones	8
1.3.2. Transiciones de entrada y salida	10
1.3.3. Reglas de funcionamiento	10
1.4. Estructuras especiales	11
1.4.1. Competencia y sincronización	11
1.4.1.1. Competencia por proveer: unión en un lugar	11
1.4.1.2. Competencia por consumir: bifurcación en un lugar	11
1.4.1.3. Sincronización en la provisión: bifurcación en una transición	12
1.4.1.4. Sincronización en el consumo: unión en una transición	12
1.4.2. Algunos mecanismos de control	12
1.4.2.1. Previendo los disparos simultáneos de una transición	12
1.4.2.2. Control del flujo	13
1.4.2.3. Lugares con capacidad finita	14
1.4.2.4. Sincronización del disparo de transiciones	15
1.5. Redes de Petri especiales	15
1.5.1. Sin sincronización: máquinas de estado, autómatas	15
1.5.1.1. Características	15
1.5.1.2. Invariantes	16

1.5.2.	Sin conflictos: grafos de eventos	17
1.5.2.1.	Características	17
1.5.2.2.	Invariantes	17
1.5.3.	Separación de conflictos y sincronización: redes con libre elección	19
2.	Semi-anillos idempotentes, dioide	21
2.1.	Definición y ejemplos	21
2.1.1.	Semi-anillos	21
2.1.2.	Idempotencia y orden	23
2.1.3.	Ejemplos de semi-anillos idempotentes	25
2.1.4.	Subdioide	26
2.2.	Homomorfismos, isomorfismos y congruencias	26
2.2.1.	Homomorfismos	26
2.2.2.	Isomorfismos	27
2.2.3.	Congruencias	28
2.3.	Más acerca de la relación de orden en un dioide	29
2.3.1.	Dioides completos	29
2.3.1.1.	Completitud y completación	29
2.3.1.2.	Máxima cota inferior	30
2.3.1.3.	Relación entre \otimes y \wedge	31
2.3.2.	Dioide arquimediano	32
2.3.3.	Dioide distributivo	33
2.4.	Dioides derivados de un dioide “escalar”	34
2.4.1.	Dioide de matrices	34
2.4.1.1.	Desde los escalares hacia las matrices	34
2.4.1.2.	Grafos asociados a las matrices	35
2.4.2.	Dioides de polinomios y series de potencias	37
2.5.	La ecuación implícita $x = ax \oplus b$	39
3.	Grafos de eventos temporizados como sistemas lineales	41
3.1.	Punto de vista del “dater”	41
3.1.1.	¿ Qué es necesario para comenzar una simulación ?	41
3.1.2.	Ecuaciones de daters	42
3.1.3.	Inicialización	42
3.1.4.	Forma matricial	43
3.1.5.	Forma canónica	44
3.1.5.1.	Eliminación de la parte implícita	44
3.1.5.2.	Ecuaciones de estados	45
3.1.6.	Resolución de un ejemplo	46
3.1.6.1.	Obtención de las ecuaciones	46
3.1.6.2.	Eliminación de la parte implícita	46
3.1.6.3.	Eliminación de las variables de estados “espúreas”	47
3.1.6.4.	Confinación de los retardos a la parte interna	47
3.2.	Matrices transferencia via la transformada γ	49
3.2.1.	La transformada γ	49

3.2.2.	Obtención de la matriz transferencia de un sistema	50
3.2.3.	Matriz transferencia y respuesta al impulso	52
3.2.4.	Resolución de un ejemplo (continuación)	52
3.2.5.	Simplificación del sistema por el cálculo de la matriz transferencia	54
3.2.6.	Imposición de condiciones iniciales no canónicas	54
3.2.7.	Una controversia sobre las condiciones iniciales	56
4.	Teoría de residuación	59
4.1.	Funciones monótonas y continuidad	59
4.1.1.	Funciones monótonas	59
4.1.2.	Continuidad	61
4.2.	Elementos de la teoría de residuación	65
4.2.1.	Resultados básicos	65
4.2.2.	Resultados adicionales	69
4.2.3.	Residuación restringida	71
4.2.4.	¿ Cuándo la ecuación $\Pi(x) = b$ admite solución ?	72
4.3.	Funciones clausura y clausura de funciones	72
4.3.1.	Funciones clausura y clausura dual	72
4.3.2.	Mejor función clausura y clausura dual	76
4.4.	Residuación de la suma y la multiplicación	78
4.4.1.	Fórmulas principales	78
4.4.2.	Aplicaciones y ejemplos	80
4.4.3.	Algunas ecuaciones de punto fijo	85
4.5.	Residuación de matrices e inversibilidad	86
4.5.1.	Residuación de matrices	86
4.5.2.	Una observación importante	87
4.5.3.	Inversibilidad	88
5.	Otras representaciones de grafos de eventos temporizados	89
5.1.	Punto de vista del “counter”	90
5.1.1.	Introducción a los counters	90
5.1.2.	Ecuación de counters	92
5.1.3.	La transformada δ y las matrices transferencia de counters	94
5.2.	Más sobre los enfoques de daters y counters	95
5.2.1.	Orden del sistema	95
5.2.2.	Monotonía de las trayectorias	95
5.2.3.	Simplificaciones	96
5.2.4.	Cambios de base	97
5.3.	Construcción del álgebra $\mathcal{M}_{in}^{dx}[\gamma, \delta]$	99
5.3.1.	Observaciones previas	100
5.3.1.1.	Un isomorfismo entre subconjuntos y series de potencias	100
5.3.1.2.	Dioide de semirectas infinitas	100
5.3.2.	Filtrado de los daters decrecientes	101
5.3.3.	El álgebra $\mathcal{M}_{in}^{dx}[\gamma, \delta]$	103
5.3.3.1.	Diferentes caminos hacia $\mathcal{M}_{in}^{dx}[\gamma, \delta]$	103

5.3.3.2.	Punto de vista práctico de $\mathcal{M}_{in}^{\delta x}[\gamma, \delta]$	104
5.3.3.3.	Interpretación informal de $\mathcal{M}_{in}^{\delta x}[\gamma, \delta]$	106
6.	Comportamiento asintótico de grafos de eventos temporizados	107
6.1.	Motivación: modelización de Sistemas de Manufactura Flexible (SMF)	107
6.1.1.	Ejemplo de jobshop	107
6.1.2.	Ecuaciones de lazo cerrado (con daters), sistema autónomo	108
6.2.	Comportamiento a lazo cerrado (o autónomo)	109
6.2.1.	Interpretación de un autovalor	109
6.2.2.	Preguntas	110
6.2.3.	El caso del autovalor nulo	110
6.3.	El caso irreducible	110
6.3.1.	Unicidad del autovalor	110
6.3.2.	Autovectores	110
6.3.3.	Comportamientos transitorio y asintótico	111
6.3.4.	Ejemplo de un caso patológico	111
6.4.	El caso reducible	111
6.4.1.	Repaso	111
6.4.2.	Aplicación: respuesta en frecuencia de GET	112
6.5.	Autovalores generalizados	113
6.6.	Algoritmos para calcular autovalores y/o autovectores	114
6.7.	“Periodicidad” de la respuesta al impulso, racionalidad, realizabilidad	114
6.8.	Estabilizabilidad y optimización de recursos	114
6.8.1.	Estabilidad y control feedback	114
6.8.2.	Aplicación a SMF	116
6.8.3.	Ejemplo	116
	Epílogo	119
	Bibliografía	121

Capítulo 1

Introducción a las redes de Petri

Las redes de Petri son una herramienta gráfica, debida a *Carl Adam Petri*, que se utiliza para representar fenómenos tales como la dinámica de eventos, la evolución en paralelo, la dependencia condicional (como la sincronización), la competencia por recursos, etc. Estos fenómenos aparecen frecuentemente en sistemas de producción, protocolos de comunicaciones, computadoras y redes de computadoras, software de tiempo real, sistemas de transporte, etc. Todos estos sistemas son conocidos en la actualidad como sistemas de eventos discretos.

1.1. Componentes básicos

1.1.1. Punto de vista de la teoría de grafos

Una red de Petri es un par (\mathcal{G}, M) compuesto por un grafo dirigido bipartito $\mathcal{G} = (E, V)$ y una marca inicial M . El conjunto de los nodos V está dividido en dos subconjuntos disjuntos P y T . Los elementos de P reciben el nombre de *lugares* mientras que los elementos de T reciben el nombre de *transiciones*. Los lugares serán representados por: P_i , $i = 1, \dots, |P|$ y las transiciones por: T_j , $j = 1, \dots, |T|$. Los arcos de E van de los lugares a las transiciones o bien de las transiciones a los lugares. Como el grafo es bipartito los arcos no pueden ir de un lugar a un lugar ni de una transición a una transición. Gráficamente los lugares serán representados mediante círculos y las transiciones mediante barras (ver Figura 1.1). A los arcos se les asigna un peso el cual está dado por un número entero (en ausencia de indicación se considera que el peso es 1). A fin de completar la definición formal de una red de Petri, debe introducirse una marca inicial. La marca inicial asigna un entero no negativo M_i a cada lugar P_i . Gráficamente, M_i puntos o marcas serán colocados en el círculo que representa a P_i . El vector columna M , cuyas componentes son las M_i , recibe el nombre de marca inicial de la red de Petri. Diremos que un lugar P_i es *anterior* a una transición T_j , si hay un arco que va de P_i a T_j . Análogamente, diremos que un lugar P_i es *posterior* a la transición T_j , si hay un arco que va de T_j a P_i .

Habitualmente los lugares representan condiciones y las transiciones representan eventos. Una transición (es decir, un evento) posee un cierto número de lugares anteriores y posteriores que representan las pre-condiciones y post-condiciones de dicho evento. La presencia de una marca en un lugar (cuando el peso de todos los

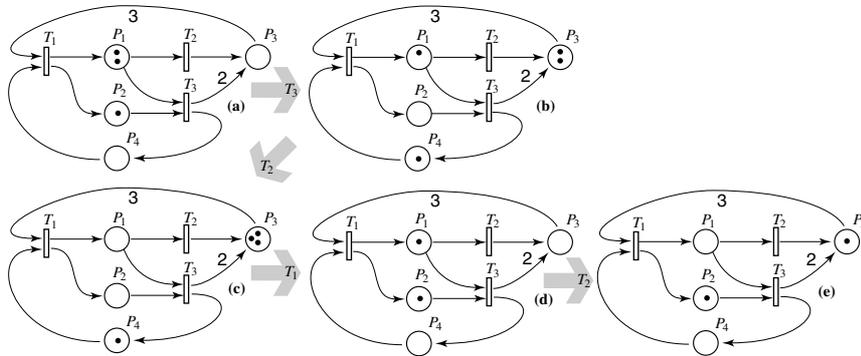


Figura 1.1: Evolución de una red de Petri

arcos es 1) puede interpretarse como que la condición asociada a dicho lugar se verifica. Otra interpretación posible es la siguiente: M_i marcas son colocadas en un lugar para indicar que hay M_i recursos disponibles.

Dentro del enfoque clásico de las redes de Petri, la marca de la red de Petri es identificada como su estado. Los cambios de estados ocurren según las siguientes reglas de evolución:

- una transición T_i puede *dispararse* o *activarse* (en dicho caso se dice que está *habilitada*) si cada lugar anterior a T_i contiene al menos tantas marcas como el peso del arco que los une;
- cuando una transición T_i es *disparada* o *activada* se elimina de cada lugar anterior a dicha transición tantas marcas como el peso del arco que los une. También se agrega a cada lugar posterior a T_i tantas marcas como el peso del arco que los une (ver Figura 1.1).

Observación 1.1. En lugar de asociar pesos a los arcos, podemos suponer que todos los arcos tienen peso uno; pero permitir varios arcos “paralelos” entre lugares y transiciones (ver Figura 1.2).

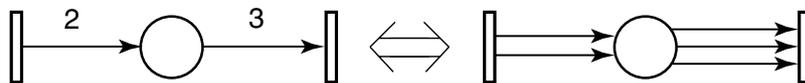


Figura 1.2: Arcos paralelos en lugar de pesos en los arcos

1.1.2. Punto de vista del álgebra lineal

En este caso el análisis se centra en el estudio de las marcas en los lugares. Consideremos como en la sección anterior el vector (columna) M de las marcas,

es decir que M_i es el número de marcas que hay en el lugar P_i . Sea Ant la matriz de dimensión $|P| \times |T|$, en donde Ant_{ij} es el peso del arco que va de P_i a T_j (si dicho arco no existe se toma $\text{Ant}_{ij} = 0$). Análogamente, sea Post la matriz de dimensión $|P| \times |T|$, en donde Post_{ij} es el peso del arco que va de T_j a P_i (si dicho arco no existe se toma $\text{Post}_{ij} = 0$). A partir de esta definición puede verse que la transición T_j está habilitada para dispararse si y sólo si $M \geq \text{Ant}_{.j}$. En dicho caso, el disparo de la transición T_j conduce a la nueva marca \tilde{M} , la cual verifica la siguiente ecuación:

$$\tilde{M} = M + \text{Post}_{.j} - \text{Ant}_{.j} .$$

Si definimos la matriz $A \stackrel{\text{def}}{=} \text{Post} - \text{Ant}$, entonces se verifica que:

$$\tilde{M} = M + Au$$

donde u es el vector columna definido por: $u_j = 1$ y $u_i = 0$ si $i \neq j$.

Ejemplo 1.2. Para la red de Petri de la Figura 1.1 obtenemos:

$$\text{Ant} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 3 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{Post} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -1 \\ -3 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} .$$

Entonces las sucesivas marcas que se obtienen por la secuencia de disparos factibles T_3, T_2, T_1, T_2 , están dadas por:

$$\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} + A \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} 0 \\ 0 \\ 3 \\ 1 \end{pmatrix} + A \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + A \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} .$$

Notemos que la marca final \tilde{M} de esta secuencia de disparos (factibles) también puede obtenerse por:

$$\tilde{M} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + A \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + A \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + A \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + A \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} .$$

En general, la marca \tilde{M} que se obtiene después de una secuencia de disparos factibles puede obtenerse como:

$$\tilde{M} = M + Au ,$$

donde u es el vector columna cuya componente i -ésima es el número de veces que aparece la transición T_i en la secuencia de disparos dada (es decir que u es la suma de los vectores que representan a los disparos individuales de la secuencia). Esto sugiere cierta linealidad en el enfoque considerado; pero como veremos más adelante, no es así.

1.1.3. Semántica de las redes de Petri

Los diferentes componentes de una red de Petri tienen habitualmente las siguientes interpretaciones:

- las marcas representan recursos en el amplio sentido de la palabra. Pueden ser tanto recursos físicos, como recursos no materiales, tales como: información, mensajes, etc.;
- los lugares es donde los recursos pueden esperar o almacenarse;
- las transiciones representan acciones que consumen recursos para luego transformarlos, o bien producir nuevos recursos;
- los pesos de los arcos que van de un lugar a una transición, representan el número mínimo de recursos de la clase almacenada en dicho lugar que son necesarios para llevar a cabo la acción representada por la transición;
- los pesos de los arcos que van de una transición a un lugar representan el número de recursos de la clase que es almacenada en dicho lugar que son producidos al llevarse a cabo la acción que representa la transición;
- el número total de marcas en una red de Petri no necesariamente debe conservarse pues, por ejemplo, una transición puede representar la operación de ensamblaje de una parte compleja a partir de partes elementales, o bien inversamente, puede representar el desguace de una parte compleja en partes elementales. También los mensajes pueden combinarse para producir un nuevo mensaje (por ejemplo, sumar dos números) o un mismo mensaje puede enviarse a varios lugares.

1.2. Ventajas y desventajas de la representación “lineal” clásica

1.2.1. Desventajas

Una marca \tilde{M} se dice que es alcanzable desde la marca M , si existe una secuencia de disparos factibles que convierten a la marca M en \tilde{M} . En dicho caso, hemos visto que si u es el vector representante de dicha secuencia de disparos (es decir, $u_i = n$ si la transición T_i aparece n veces en la secuencia), entonces se verifica que:

$$\tilde{M} = M + Au .$$

Sin embargo, la existencia de un vector no negativo u tal que $\tilde{M} = M + Au$, no implica que la marca \tilde{M} sea alcanzable desde M , ya que la condición $M' \geq \text{Ant}.i$ debe verificarse en cada paso en el que se pasa de la marca M' a la M'' por el disparo de la transición T_i . Además, el vector u no dice nada acerca del orden en que las transiciones deben dispararse, el cual es importante.

Debido a estas restricciones, el sistema no es realmente lineal y en realidad el principio de superposición no se verifica.

1.2.2. Invariantes.

1.2.2.1. Invariantes de lugares

Supongamos que v es un vector fila de dimensión $|P|$ el cual verifica que $vA = 0$. Entonces vM (lo cual puede interpretarse como la suma ponderada del número de marcas en los lugares con los pesos dados por las componentes del vector v) se mantiene constante cualquiera sea la secuencia de disparos, pues si \tilde{M} es una marca alcanzable desde M , entonces resulta que:

$$\tilde{M} = M + Au \Rightarrow v\tilde{M} = vM + vAu = vM .$$

Ejemplo 1.3. Para la red de Petri de la Figura 1.1 tenemos que:

$$(0 \ 1 \ 0 \ 1)A = (0 \ 1 \ 0 \ 1) \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -1 \\ -3 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} = (0 \ 0 \ 0 \ 0) .$$

Por lo tanto, el número total de marcas en los lugares P_2 y P_4 se mantiene constante independientemente de los disparos que se realicen.

1.2.2.2. Invariantes de transiciones

Supongamos ahora que u es un vector columna de dimensión $|T|$ y con componentes no negativas el cual verifica que: $Au = 0$. Entonces, cualquier secuencia de disparos factible que tenga al vector u como representante conserva la marca inicial, pues

$$\tilde{M} = M + Au = M .$$

Sin embargo, como fue indicado anteriormente, dicha secuencia de disparos factible puede no existir.

Ejemplo 1.4. Para la red de Petri de la Figura 1.3 obtenemos:

$$A = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -1 \\ -2 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} , \quad A \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

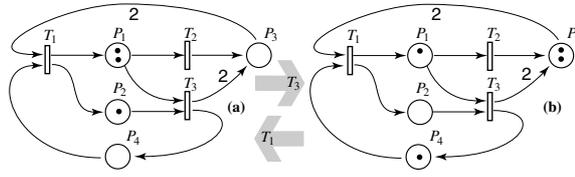


Figura 1.3: Conservación de la marca inicial por la secuencia $\{T_3, T_1\}$

En este caso podemos ver que la secuencia de disparos T_3, T_1 es factible (y conserva la marca inicial); pero la secuencia T_1, T_3 no lo es. Ambas tienen el mismo vector representante $u = (1 \ 0 \ 1)^t$.

1.2.3. Otros conceptos y propiedades

1.2.3.1. Conflictos

Definición 1.5. Se dice que dos transiciones T_i y T_j están en un *conflicto estructural* cuando:

$$\exists k : \text{Ant}_{ki} \times \text{Ant}_{kj} \neq 0 ,$$

lo cual significa que P_k es un lugar anterior a ambas transiciones T_i y T_j .

Se dice que un conflicto estructural es *efectivo* para la marca M si además ambas transiciones T_i y T_j pueden dispararse, es decir, $M \geq \text{Ant}_{.i}$ y $M \geq \text{Ant}_{.j}$.

Ejemplo 1.6. En la red de Petri de la Figura 1.3, las transiciones T_2 y T_3 están en un conflicto estructural (P_1 es un lugar anterior a ambas transiciones) y dicho conflicto es efectivo en el caso de la Figura 1.3-(a); pero no lo es en el caso de la Figura 1.3-(b).

El término conflicto proviene del hecho de que si T_i y T_j están en un conflicto estructural efectivo, entonces puede suceder que solamente una de estas transiciones pueda dispararse (y no ambas a la vez porque no hay “suficientes” marcas en los lugares anteriores a las mismas como para que esto suceda), y entonces es necesario decidir cual de las dos transiciones se disparará. En cierta forma, dos transiciones que están en un conflicto estructural compiten por los recursos almacenados en el lugar anterior que ambas comparten.

1.2.3.2. Paralelismo

Definición 1.7. Se dice que dos transiciones T_i y T_j están *estructuralmente en paralelo* cuando:

$$(\text{Ant}_{.i})^t \cdot \text{Ant}_{.j} = 0 ,$$

lo cual significa que T_i y T_j no tienen lugares anteriores en común.

Se dice que el paralelismo es *efectivo* para una marca M si, además, ambas transiciones T_i y T_j pueden dispararse (están habilitadas), es decir, $M \geq \text{Ant}_{.i}$ y $M \geq \text{Ant}_{.j}$.

Ejemplo 1.8. En la red de Petri de la Figura 1.3, las transiciones T_1 y T_2 están estructuralmente en paralelo. Dicho paralelismo es efectivo en el caso de la Figura 1.3-(b); pero no lo es en el caso de la Figura 1.3-(a).

1.2.3.3. Vivacidad

Definición 1.9. Una transición T_i se dice *viva*, cuando desde cualquier marca alcanzable desde la marca inicial, existe una secuencia de disparos factibles que conduce a una marca en la cual la transición T_i está habilitada para dispararse.

Una red de Petri se dice *viva* si todas sus transiciones son vivas.

Notemos que si una transición es viva, entonces puede dispararse indefinidamente (es decir que existe una sucesión de disparos factibles en donde dicha transición aparece infinitas veces). Si una transición no es viva, entonces se la podría sacar de la red de Petri pasado cierto tiempo de funcionamiento de la misma.

Ejemplo 1.10. La red de Petri de la Figura 1.1 no es viva pues la secuencia de disparos considerada en dicha figura conduce a una marca en la cual ninguna transición está habilitada para dispararse (esta situación recibe el nombre de “dead-lock”).

1.2.3.4. Acotación, seguridad

Definición 1.11. Un lugar P_i se dice *k-acotado* si su marca no es mayor que k para todas las marcas alcanzables. Un lugar se dice *seguro* si es 1-acotado. Una red de Petri se dice *segura* si todos sus lugares son seguros.

Notemos que si un lugar P_i de una red de Petri representa a un lugar de almacenamiento con cierta capacidad finita k (un depósito por ejemplo), entonces el mismo debería ser *k-acotado* si la modelización es correcta. Más adelante veremos un método sencillo para asegurar dicha acotación.

Ejemplo 1.12. La secuencia de disparos T_3, T_1 de la Figura 1.4 conduce a una marca que es igual a la marca inicial, salvo por el hecho de que el lugar P_3 tiene una marca más. Por lo tanto, repitiendo dicha secuencia de disparos k veces obtenemos k marcas en P_3 , y entonces el lugar P_3 no es acotado.

Matemáticamente tenemos que (u es el vector representante de la secuencia T_3, T_1):

$$Au = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -1 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y entonces

$$\tilde{M} = M + Au > M .$$

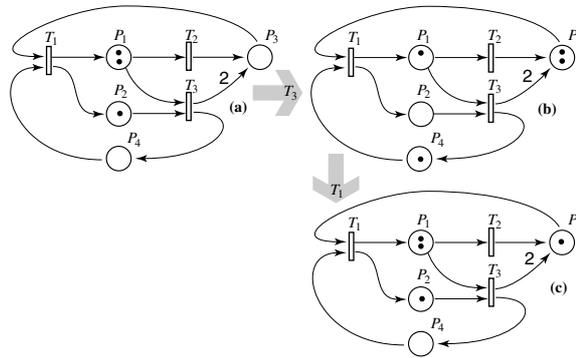


Figura 1.4: Red de Petri no acotada

Como $\tilde{M} > M$, la secuencia T_3, T_1 sigue siendo factible, y entonces es posible repetir este procedimiento para hacer crecer la marca de P_3 (pues $(Au)_3 > 0$) indefinidamente.

A partir de este ejemplo podemos ver que cuando una red de Petri admite una secuencia de disparos factibles cuyo vector representante u verifica:

$$Au > 0,$$

entonces dicha red de Petri no es acotada (se aplica exactamente el mismo razonamiento que en el ejemplo anterior para llegar a la conclusión de que la marca no es acotada en los lugares P_i para los cuales se verifica que: $(Au)_i > 0$).

1.2.3.5. Marcas alcanzables

La mayoría de las propiedades anteriores podrían verificarse si se conociese el conjunto de las marcas alcanzables desde la marca inicial. Sin embargo, el cálculo de todas las marcas alcanzables desde la marca inicial es en general una tarea muy complicada. En la Figura 1.5 puede observarse dicho cálculo en un caso sencillo.

1.3. Introducción del tiempo y comportamiento sistemático

1.3.1. Tiempos asociados a lugares y transiciones

La teoría original de las redes de Petri trata del ordenamiento de eventos, y en la misma, la dinámica de la red ha sido considerada como una cadena de eventos (los disparos de las transiciones) restringida únicamente por condiciones lógicas (una transición puede dispararse sólo si está habilitada). Entonces preguntas tales como: ¿cuándo un evento se produce? no son consideradas. Por lo tanto, para poder responder a preguntas relacionadas con la evaluación de la performance de una red (como por ejemplo, ¿cuán rápido puede una red producir?), es necesario

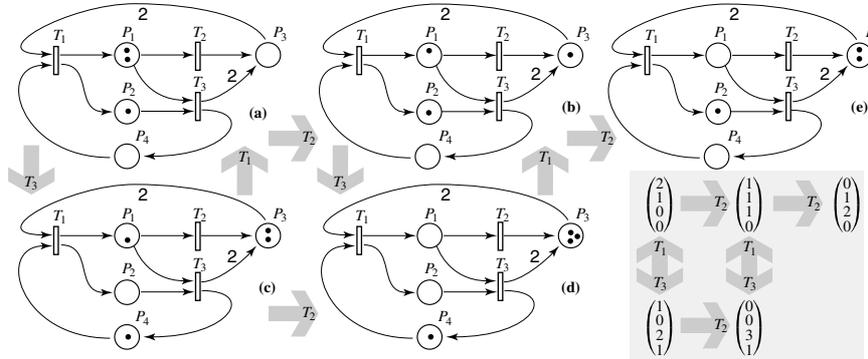


Figura 1.5: Conjunto de marcas alcanzables

introducir el tiempo. Esto puede hacerse asociándole tiempos a las transiciones y a los lugares:

- tiempos asociados a los lugares: estos son los tiempos mínimos que las marcas deben permanecer en un lugar antes de contribuir a habilitar el disparo de una transición posterior a dicho lugar. Los mismos reciben el nombre de *tiempos de espera*;
- tiempos asociados a las transiciones: estos son los tiempos que separan el comienzo (el consumo de marcas de los lugares anteriores) y la finalización (la producción de marcas hacia los lugares posteriores) del disparo de una transición. Los mismos reciben el nombre de *tiempos de disparo*.

Los tiempos de disparo pueden utilizarse, por ejemplo, para representar tiempos de producción en el caso de los sistemas de producción (en donde las transiciones representan habitualmente a máquinas). Los tiempos de espera, pueden utilizarse para representar tiempos de transporte (es el caso en el que el lugar representa a una ruta o canal que comunica dos procesos); o bien, tiempos de almacenamiento mínimos (como por ejemplo en el caso en que una pieza debe enfriarse antes de que se pueda aplicarle el siguiente proceso).

Los tiempos de disparo y de espera pueden ser constantes en el tiempo, o pueden variar según cual sea el número de la marca en llegar a un lugar o en disparar una transición, o según el tiempo, o aleatoriamente.

Observación 1.13. En realidad, no hay pérdida de generalidad en suponer que todos los disparos son instantáneos (es decir que todos los tiempos de disparo son nulos), pues aquellas transiciones con tiempo de disparo no nulo pueden dividirse en dos transiciones instantáneas (que representan el comienzo y la finalización del disparo) separadas por un lugar que tiene por tiempo de espera al tiempo de disparo de la transición original (ver Figura 1.6).

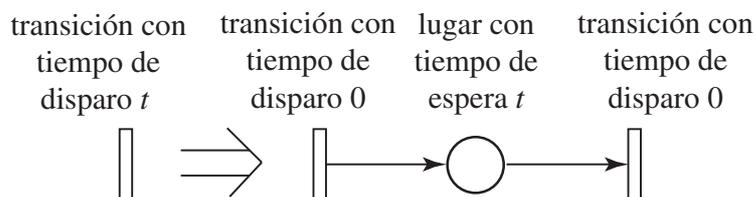


Figura 1.6: Transformación que se utiliza para que los disparos sean instantáneos

1.3.2. Transiciones de entrada y salida

Las transiciones que no poseen lugares anteriores reciben el nombre de *transiciones de entrada* (o fuentes). Los disparos de las mismas se deben a decisiones externas (son “controladas” desde el exterior). Las transiciones que no poseen lugares posteriores se llaman *transiciones de salida* (o sumideros). Los disparos de las mismas nos indican cuando se producen marcas desde la red hacia el exterior.

Las mismas definiciones pueden hacerse para los lugares (los lugares de entrada deben ser provistos de marcas desde el exterior). En realidad, como veremos más adelante, de acuerdo a cual sea la clase particular de red de Petri bajo consideración, puede resultar más apropiado “terminar” la red con transiciones (sin lugares de entrada ni de salida) o con lugares (sin transiciones de entrada ni de salida).

1.3.3. Reglas de funcionamiento

Hasta aquí únicamente restricciones lógicas fueron impuestas al disparo de las transiciones; pero nada se ha dicho acerca de cuando una transición debe dispararse. Habiendo introducido el tiempo, es posible definir ahora un comportamiento sistemático mediante la siguiente regla.

- **Tiempo más rápido de funcionamiento:** las transiciones son disparadas *tan pronto como sea posible*, es decir, en el primer instante en el cual resultan disponibles todas las marcas que son necesarias para llevar a cabo el disparo (es decir que han estado en el lugar correspondiente por lo menos el tiempo de espera de dicho lugar).

Siguiendo la regla anterior resulta que si además son conocidas.

- **Reglas de prioridad:** estas son reglas que sirven para arbitrar en los lugares que tienen un conflicto (las mismas indican que transición debe dispararse cuando se produce un conflicto, ver §1.2.3.1).
- **Trayectorias de entrada:** son funciones $u_i : \mathbb{N} \rightarrow \mathbb{R}^+$ (una para cada transición de entrada U_i) donde $u_i(n)$ es el instante en que la transición de entrada U_i se enciende por n -ésima vez.

entonces es posible determinar todos los instantes en los cuales se producen los eventos en la red de Petri, como por ejemplo, los sucesivos disparos, el arribo y la

partida de marcas de un lugar, etc. En particular, es posible determinar los instantes en que se producen los disparos de las transiciones de salida, los cuales conforman las *trayectorias de salida*.

1.4. Estructuras especiales

1.4.1. Competencia y sincronización

1.4.1.1. Competencia por proveer: unión en un lugar

Esta situación se da cuando un lugar tiene varias transiciones anteriores (ver Figura 1.7). En este caso existen varias “fuentes” que proveen de marcas a dicho lugar.

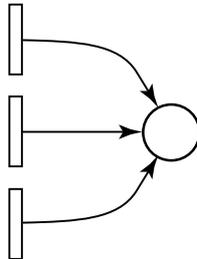


Figura 1.7: Unión en un lugar

Como ejemplo de una situación de este tipo podemos citar el caso en el cual el lugar representa a un depósito de cierto producto que es fabricado por varias máquinas diferentes (las transiciones representan a dichas máquinas).

1.4.1.2. Competencia por consumir: bifurcación en un lugar

Esta situación se da cuando un lugar tiene varias transiciones posteriores (ver Figura 1.8). Estas transiciones “compiten” por las marcas de dicho lugar. Se trata de un conflicto estructural el cual ya fue estudiado anteriormente (ver §1.2.3.1).

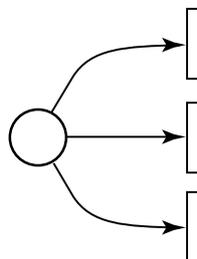


Figura 1.8: Bifurcación en un lugar

1.4.1.3. Sincronización en la provisión: bifurcación en una transición

Esta situación se da cuando una transición tiene varios lugares posteriores (ver Figura 1.9). En este caso las marcas (las cuales pueden representar recursos, partes, mensajes, etc.) son enviadas en forma simultánea hacia los consumidores posteriores.

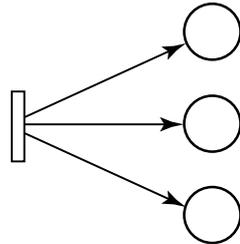


Figura 1.9: Bifurcación en una transición

Por ejemplo, este caso se presentaría cuando la transición representa a una operación en la cual una pieza compleja es dividida en varias partes más elementales (desguace).

1.4.1.4. Sincronización en el consumo: unión en una transición

Esta situación se da cuando una transición tiene varios lugares anteriores (ver Figura 1.10). Las marcas esperan en estos lugares hasta el instante en el cual resulta disponible la última marca que es necesaria para habilitar el disparo de la transición (es decir que ha estado en el lugar correspondiente al menos el tiempo de espera de dicho lugar). En dicho instante son consumidas simultáneamente todas las marcas que son necesarias para el disparo de la transición.

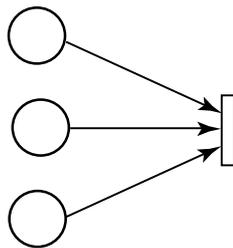


Figura 1.10: Unión en una transición

1.4.2. Algunos mecanismos de control

1.4.2.1. Previendo los disparos simultáneos de una transición

Debido a las definiciones que hemos realizado, nada impide que una transición realice varios disparos simultáneamente (si el disparo de una transición no es

instantáneo, entonces puede suceder que dicha transición sea disparada por segunda vez, sin que el primer disparo se haya completado). Si se desea evitar dicho fenómeno (como por ejemplo, cuando una transición representa a una máquina que realiza cierta tarea sobre una pieza; pero que no puede trabajar sobre dos o más piezas a la vez) se debe agregar un lugar extra asociado a dicha transición. Este lugar extra debe tener como única transición anterior y posterior a la transición considerada. Esto se representa en la Figura 1.11 con o sin la transformación que se utiliza para obtener que el disparo de todas las transiciones sea instantáneo (ver Figura 1.6).

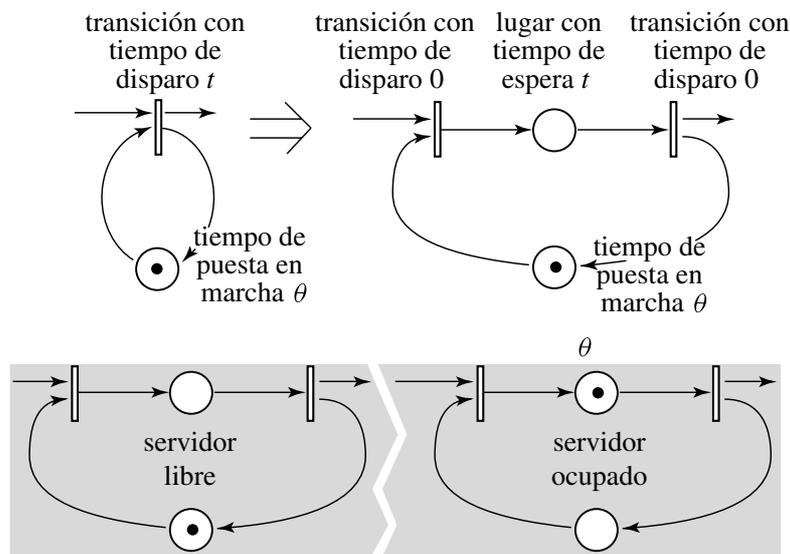


Figura 1.11: Reciclado de una transición

El agregado de este "bucle" recibe el nombre de *reciclado* de la transición. Además se podría asignar un tiempo de espera positivo al "lugar de reciclado" para forzar un tiempo mínimo entre la finalización de un disparo y el inicio del siguiente (esto puede interpretarse como un tiempo de "puesta en marcha"). Notemos que la posición de la marca en el "bucle" indica si la transición está "ocupada" o "desocupada".

1.4.2.2. Control del flujo

Una modificación similar al reciclado permite acotar el flujo de marcas a través de una transición con tiempo de disparo nulo (ver Figura 1.12). Notemos que si la marca inicial del lugar extra asociado a la transición (el cual al igual que en el reciclado debe tener como única transición anterior y posterior a la transición considerada) es m , y si su tiempo de espera es t , entonces el flujo máximo de marcas a través de la transición considerada es de m marcas por cada t unidades de tiempo.

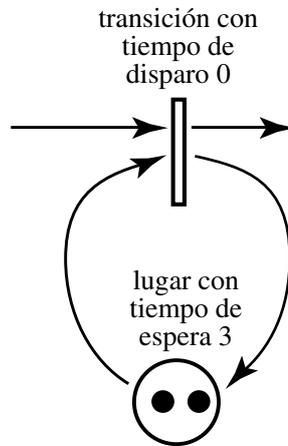


Figura 1.12: Flujo máximo de 2 marcas por cada 3 unidades de tiempo

1.4.2.3. Lugares con capacidad finita

En la modelización de muchos sistemas físicos resulta natural encontrar que algunos lugares poseen cierta cota superior en el número de marcas que pueden contener. Por ejemplo, este sería el caso cuando un lugar representa a un depósito que posee una capacidad finita. Esta acotación puede lograrse de la siguiente manera. Supongamos que deseamos que el lugar P posea una capacidad máxima de K marcas. Consideremos el caso en el cual P tiene a T_i como su única transición anterior y a T_j como su única transición posterior (esta construcción puede extenderse en forma directa a casos más generales). Entonces se agrega un lugar extra P' de tal manera que posea a T_j como su única transición anterior y a T_i como su única transición posterior. Si la marca inicial de P es μ ($\mu \leq K$), entonces a P' se le asigna la marca inicial $K - \mu$ (ver Figura 1.13). Como el número total de marcas en P y P' se mantiene constante (notemos que cada vez que se saca una marca de uno de estos lugares se agrega una marca al otro), resulta que el número máximo de marcas en P es K .

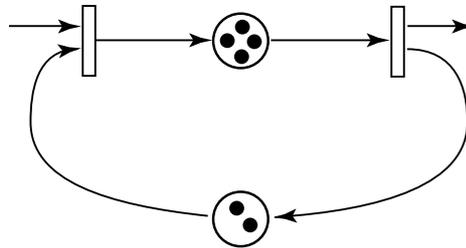


Figura 1.13: Depósito con capacidad máxima 6 (el depósito contiene 4 elementos y hay dos lugares libres)

1.4.2.4. Sincronización del disparo de transiciones

A veces puede suceder que dos o más transiciones de una red de Petri representen la misma operación física. En ese caso dichas transiciones deberían *sincronizarse* para que se enciendan siempre simultáneamente. Esto puede lograrse en al menos dos formas (una de las cuales no es del todo aceptable desde el punto de vista de la teoría clásica de redes de Petri; pero como veremos más adelante, es matemáticamente correcta en el aspecto de expresar simultaneidad) las cuales equivalen a “unir” a las transiciones consideradas (ver Figura 1.14).

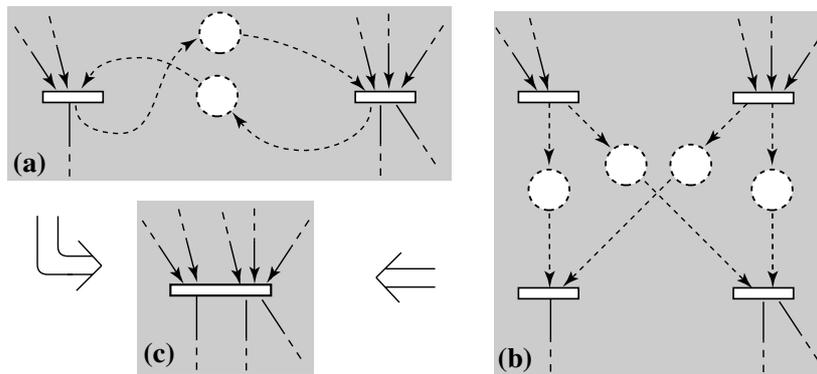


Figura 1.14: Dos mecanismos para sincronizar dos transiciones

1.5. Redes de Petri especiales

1.5.1. Sin sincronización: máquinas de estado, autómatas

1.5.1.1. Características

Las redes de Petri sin sincronización son aquellas en las cuales todas las transiciones tienen a lo sumo un lugar anterior y a lo sumo un lugar posterior (ver Figura 1.15). En dichas redes no resulta apropiado tener transiciones de entrada o de salida, y por lo tanto, resulta conveniente “terminarlas” con lugares (en dicho caso cada transición posee exactamente un lugar anterior y un lugar posterior). Las redes de Petri en las cuales cada transición posee exactamente un lugar anterior y un lugar posterior reciben el nombre de *máquinas de estado*. En las máquinas de estado la asignación de tiempo a las transiciones y lugares no es importante (su único efecto es el de retardar los disparos de las transiciones) ya que la cuestión principal es lógica (marcas alcanzables, eliminación de bloqueo, etc.). En las mismas, el control principal es el disparo de las transiciones. Cuando el número total de marcas es uno, podemos pensar que la única marca indica cual es el estado del sistema (los lugares representan los posibles estados del sistema), y la red que obtenemos puede interpretarse como un autómata. Si además cada lugar posee exactamente una transición posterior, entonces dicho autómata resulta ser determinístico. Si este no

es el caso, el autómata no es determinístico (como en la Figura 1.15) y entonces entre cada par de estados es posible tener varias trayectorias. Si en el caso no determinístico se asignan probabilidades a los arcos que salen de un lugar, entonces obtenemos que el autómata es estocástico.

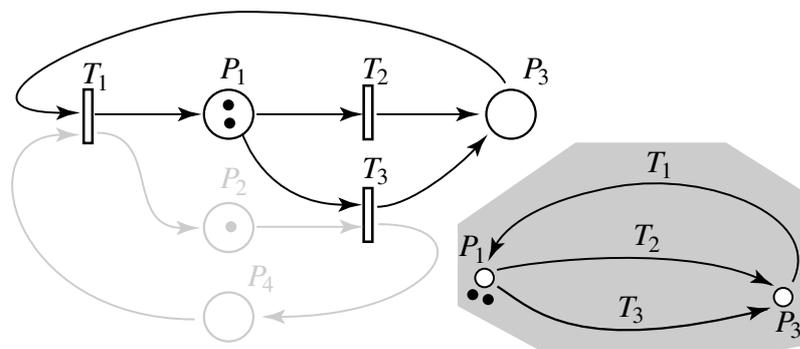


Figura 1.15: Máquina de estados que se obtiene al eliminar la sincronización en la red de Petri de la Figura 1.1

1.5.1.2. Invariantes

Como en una máquina de estados cada transición tiene exactamente un lugar anterior y un lugar posterior, la matriz $A = \text{Post} - \text{Ant}$ tiene exactamente un -1 y un 1 por columna (supondremos que todos los arcos tienen peso uno). En realidad, A puede pensarse como la matriz de incidencia nodo-arco¹ del grafo dirigido que se obtiene cuando se reemplaza a cada transición por un arco que va del lugar anterior al lugar posterior a dicha transición (los nodos de este grafo son los lugares de la red de Petri inicial). A partir de esta observación, y de resultados simples de la teoría de grafos, obtenemos las siguientes consecuencias.

- **Invariante de lugares:** como A tiene exactamente un -1 y un 1 por columna resulta que

$$(1 \quad \dots \quad 1) A = 0.$$

Por lo tanto, el número total de marcas en una máquina de estados permanece constante.

Para que una red de Petri sea viva su marca inicial debe ser no nula. Para una máquina de estados esta condición es también suficiente si la misma es fuertemente conexa.

¹La componente A_{ij} de la matriz de incidencia nodo-arco del grafo dirigido $\mathcal{G} = (E = \{e_1, \dots, e_k\}, V = \{v_1, \dots, v_r\})$ se define como: 1 si $e_j = (v_l, v_i)$ para algún l , -1 si $e_j = (v_i, v_l)$ para algún l , 0 en cualquier otro caso.

- **Invariante de transiciones:** supongamos que u es el vector (columna) característico de un circuito,² es decir que las componentes de u correspondientes a los arcos (transiciones) del circuito valen 1, y las restantes valen 0. Entonces se verifica que:

$$Au = 0 .$$

En consecuencia, cuando en una máquina de estados todas las transiciones de un circuito son disparadas una sola vez, entonces se obtiene nuevamente la marca inicial (ver por ejemplo los circuitos $\{T_2, T_1\}$ y $\{T_3, T_1\}$ en la Figura 1.15).

1.5.2. Sin conflictos: grafos de eventos

1.5.2.1. Características

Las redes de Petri sin competencia son aquellas en las cuales todos los lugares tienen a lo sumo una transición anterior y a lo sumo una transición posterior (ver Figura 1.16). Para dichas redes de Petri no resulta apropiado tener lugares de entrada o de salida, y por lo tanto, resulta conveniente “terminarlas” con transiciones (logrando de esta manera que cada lugar posea exactamente una transición anterior y una transición posterior). Las redes de Petri en las que cada lugar posee exactamente una transición anterior y una transición posterior reciben el nombre de *Grafos de Eventos*. Un grafo de eventos no puede modelizar conflictos (la marca de un lugar puede ser “consumida” por una única transición predeterminada); pero si puede modelizar sincronización. Por lo tanto, la asignación de tiempo a los lugares y a las transiciones juega un rol importante en la dinámica de dichos grafos. En los Grafos de Eventos los lugares juegan el rol de arcos (que tienen marcas y que van desde su única transición anterior a su única transición posterior) y las transiciones el rol de nodos.

1.5.2.2. Invariantes

Como en un Grafo de Eventos cada lugar tiene exactamente una transición anterior y una posterior, la matriz $A = \text{Post} - \text{Ant}$ tiene exactamente un -1 y un 1 por fila (supondremos que todos los arcos tienen peso uno). La matriz A puede pensarse como la matriz de incidencia arco-nodo³ de un grafo (aquel en el cual se reemplaza a los lugares por arcos), de donde obtenemos las siguientes consecuencias.

- **Invariante de transiciones:** como

$$A \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0 ,$$

²Recordemos que en una máquina de estados pensamos a las transiciones como arcos.

³La matriz de incidencia arco-nodo es igual a la transpuesta de la matriz de incidencia nodo-arco que ya fue definida.

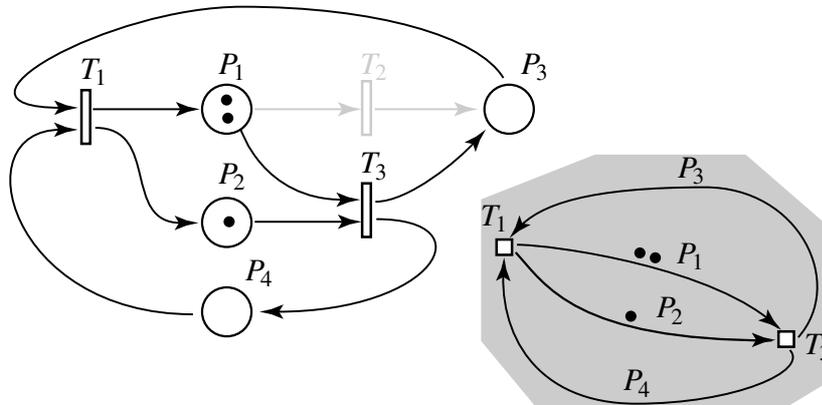


Figura 1.16: Grafo de eventos que se obtiene al eliminar los conflictos en la red de Petri de la Figura 1.1

resulta que si en un Grafo de Eventos todas las transiciones son disparadas una sola vez, entonces se obtiene nuevamente la marca inicial.

- **Invariante de lugares:** si u es un vector (fila) característico de un circuito (pensando a los lugares como arcos), es decir que las componentes de u correspondientes a los arcos (lugares) del circuito valen 1 y las restantes valen 0, entonces se verifica que

$$uA = 0 .$$

Por lo tanto, el número total de marcas en cualquier circuito de un Grafo de Eventos es constante.

Un Grafo de Eventos es *vivo* solamente cuando existe al menos una marca en cada uno de sus circuitos (pues por la propiedad anterior una transición de un circuito sin marcas nunca puede dispararse). Se puede demostrar que esta condición también es suficiente si el Grafo de Eventos es autónomo (es decir, sin transiciones de entrada).

Utilizando la transformación representada en la Figura 1.6, podemos suponer que solamente hay tiempos de espera asignados a los lugares (es decir que el disparo de todas las transiciones es instantáneo). Cuando estos tiempos de espera son constantes (que además consideraremos enteros definiendo apropiadamente la unidad de tiempo), los mismos son representados mediante barras colocadas en los lugares correspondientes. Por ejemplo, en la figura siguiente se representa a un lugar con una sola marca y un tiempo de espera de dos unidades de tiempo.



1.5.3. Separación de conflictos y sincronización: redes con libre elección

Supongamos que una red de Petri verifica la siguiente propiedad: si un lugar P_j es anterior a la transición T_i , entonces P_j es el único lugar anterior a T_i . Entonces en dicha red de Petri, una bifurcación en un lugar no puede ser seguida inmediatamente por una unión en una transición (ver Figura 1.17). En este caso la elección de la transición posterior al lugar P_j que consumirá una marca del mismo es libre, y no está condicionada por una restricción de sincronización sobre dicha transición (si además del lugar P_j la transición T_i tuviese otro lugar anterior P_k sin marcas, entonces esta transición no estaría habilitada para dispararse, y por lo tanto no podría consumir una marca de P_j , con lo cual la elección dejaría de ser libre). En este caso los conflictos y las sincronizaciones están “separados”.

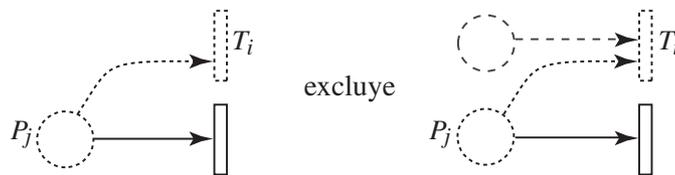


Figura 1.17: Libre elección, primer definición

Otra condición menos restrictiva que asegura esta propiedad de libre elección es la siguiente: si dos transiciones T_i y T_k tienen un lugar anterior en común, entonces tienen el mismo conjunto de lugares anteriores (ver Figura 1.18). En este caso, dos transiciones posteriores a un lugar están sujetas a las mismas restricciones de sincronización, y por lo tanto, la elección de la transición que consumirá una marca de dicho lugar no resulta afectada por restricciones de sincronización.

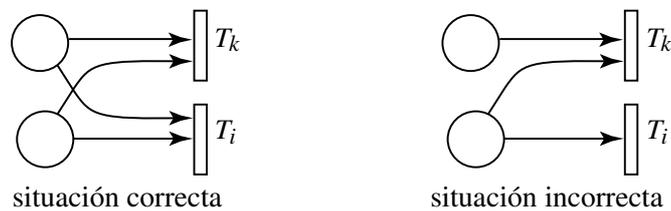


Figura 1.18: Libre elección, segunda definición

Existen otras condiciones equivalentes como por ejemplo: si dos lugares P_i y P_k tienen una transición posterior en común, entonces tienen el mismo conjunto de transiciones posteriores.

La clase de las redes de Petri que poseen esta propiedad de libre elección parece ser la clase menos restrictiva en la cual resultan tratables cuestiones tales como vivacidad, acotación, etc. Muchos resultados se basan en la descomposición de una red en subredes que sean Grafos de Eventos (conectadas mediante lugares con elección), o bien, en subredes que sean máquinas de estados (conectadas mediante transiciones con sincronización).

Capítulo 2

Semi-anillos idempotentes, dioide

En este capítulo estudiaremos ciertas estructuras algebraicas que están motivadas por el deseo de obtener modelos “lineales” para una clase de redes de Petri temporalizadas (grafos de eventos).

Comenzaremos con la estructura de semi-anillo, la cual comparte varias características con el álgebra lineal convencional. En estas estructuras, la suma puede o no simetrizarse, es decir, se puede o no extender el conjunto de manera tal que cada elemento posea un elemento opuesto.

Luego consideraremos los semi-anillos idempotentes o dioides que son semi-anillos en los cuales la suma es idempotente. Una consecuencia del hecho de que la suma sea idempotente es que no es posible definir el elemento opuesto. Sin embargo, una relación de orden queda naturalmente definida como consecuencia de que la suma es idempotente, y entonces, un semi-anillo idempotente posee características en común tanto con el álgebra lineal convencional como con los lattices o semilattices. Esta relación de orden juega un rol muy importante, tanto es así, que se puede decir que lo que se pierde por la ausencia de elemento opuesto se recupera en cierta forma gracias a la relación de orden.

2.1. Definición y ejemplos

2.1.1. Semi-anillos

Definición 2.1. Un *semi-anillo* es un conjunto S en el cual están definidas dos operaciones llamadas *suma* (representada por \oplus) y *multiplicación* (representada por \otimes) que verifican los siguientes axiomas:

- asociatividad de la suma

$$(a \oplus b) \oplus c = a \oplus (b \oplus c), \quad \forall a, b, c \in S;$$

- conmutatividad de la suma

$$a \oplus b = b \oplus a, \quad \forall a, b \in S;$$

- asociatividad de la multiplicación

$$(a \otimes b) \otimes c = a \otimes (b \otimes c), \quad \forall a, b, c \in S;$$

- la multiplicación es distributiva con respecto a la suma

$$\begin{aligned} a \otimes (b \oplus c) &= (a \otimes b) \oplus (a \otimes c) , \\ (b \oplus c) \otimes a &= (b \otimes a) \oplus (c \otimes a) , \quad \forall a, b, c \in S ; \end{aligned}$$

es decir, la multiplicación es distributiva con respecto a la suma tanto a derecha como a izquierda (notemos que una afirmación no implica a la otra porque la multiplicación no es necesariamente conmutativa);

- existe un elemento nulo

$$\exists \varepsilon \in S : a \oplus \varepsilon = a , \quad \forall a \in S ;$$

- existe un elemento unidad

$$\exists e \in S : a \otimes e = e \otimes a = a , \quad \forall a \in S ;$$

- el elemento nulo es absorbente

$$\varepsilon \otimes a = a \otimes \varepsilon = \varepsilon , \quad \forall a \in S .$$

Ejemplo 2.2. Como primer ejemplo de semi-anillo tomemos $S = \mathbb{R} \cup \{-\infty\}$ (lo mismo vale para $S = \mathbb{Q} \cup \{-\infty\}$ o $S = \mathbb{Z} \cup \{-\infty\}$) y definamos

$$a \oplus b = \max\{a, b\} , \quad a \otimes b = a + b , \quad \forall a, b \in S .$$

Entonces se puede verificar fácilmente que S con estas operaciones verifica todos los axiomas anteriores (notemos que es necesario agregar a \mathbb{R} el $-\infty$, pues en caso contrario, la suma no tendría elemento nulo). En este caso tenemos que $\varepsilon = -\infty$ y $e = 0$. Observemos también que la suma es idempotente, es decir, $a \oplus a = a, \forall a \in S$. Este semi-anillo será representado por \mathbb{R}_{\max} .

Ejemplo 2.3. Otro ejemplo de semi-anillo es el conjunto de los números naturales \mathbb{N} con la suma y multiplicación habituales.

Observación 2.4. Muchas veces el símbolo \otimes será omitido como sucede habitualmente en el álgebra convencional. Desde aquí en adelante

$$\underbrace{a \otimes \cdots \otimes a}_{k \text{ veces}}$$

será representado por a^k y a^0 será igual a e .

2.1.2. Idempotencia y orden

Definición 2.5. Si en un semi-anillo S la suma \oplus es idempotente, es decir si

$$a \oplus a = a, \quad \forall a \in S,$$

entonces diremos que S es un *semi-anillo idempotente* o un *dioide*.

Cuando la suma \oplus es idempotente, queda inmediatamente asociada a ella una relación de orden, como lo indica el siguiente teorema.

Teorema 2.6. En un dioide S se verifica la siguiente equivalencia:

$$a = a \oplus b \iff \exists c : a = b \oplus c, \quad \forall a, b \in S.$$

Estas afirmaciones equivalentes definen una relación de orden (parcial) \succeq en S que está definida por:

$$a \succeq b \iff a = a \oplus b.$$

Esta relación de orden es compatible con la suma, es decir

$$a \succeq b \Rightarrow a \oplus c \succeq b \oplus c, \quad \forall c \in S,$$

y con la multiplicación, esto es

$$a \succeq b \Rightarrow ac \succeq bc, \quad \forall c \in S$$

(lo mismo para la multiplicación a izquierda). Además todo par de elementos a y b de S posee una menor cota superior (que llamaremos simplemente cota superior) la cual está dada por $a \oplus b$ y ε es el menor elemento de S .

Demostración.

- Si $a = a \oplus b$, entonces tomando $c = a$, resulta que $b \oplus c = b \oplus a = a$. Inversamente, si $a = b \oplus c$, entonces sumando b a ambos lados de esta igualdad obtenemos: $a \oplus b = (b \oplus c) \oplus b = (b \oplus b) \oplus c = b \oplus c = a$.
- Para mostrar que \succeq es una relación de orden es necesario probar que es reflexiva, antisimétrica y transitiva.

1. Reflexiva ($a \succeq a, \forall a \in S$):

$$a \oplus a = a \Rightarrow a \succeq a, \quad \forall a \in S.$$

2. Antisimétrica ($a \succeq b, b \succeq a \Rightarrow a = b$):

$$\left. \begin{array}{l} a \succeq b \Rightarrow a = a \oplus b \\ b \succeq a \Rightarrow b = b \oplus a \end{array} \right\} \Rightarrow a = a \oplus b = b \oplus a = b,$$

3. Transitiva ($a \succeq b, b \succeq c \Rightarrow a \succeq c$):

$$\begin{aligned} \left\{ \begin{array}{l} a \succeq b \\ b \succeq c \end{array} \right\} &\Rightarrow \left\{ \begin{array}{l} a = a \oplus b \\ b = b \oplus c \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} a \oplus c = a \oplus b \oplus c \\ a = a \oplus b \\ b = b \oplus c \end{array} \right\} \Rightarrow \\ &\Rightarrow \left\{ \begin{array}{l} a \oplus c = a \oplus b \\ a = a \oplus b \end{array} \right\} \Rightarrow \{a = a \oplus c\} \Rightarrow \{a \succeq c\} . \end{aligned}$$

- Como $a \succeq b \Rightarrow a = a \oplus b$, resulta que $a \oplus c = a \oplus b \oplus c = a \oplus b \oplus c \oplus c = (a \oplus c) \oplus (b \oplus c)$, y por lo tanto $a \oplus c \succeq b \oplus c$.
- Como $a \succeq b \Rightarrow a = a \oplus b$, resulta que $ca = c(a \oplus b) = ca \oplus cb$, y entonces $ca \succeq cb$.
- Claramente $a \oplus b \succeq a$ y $a \oplus b \succeq b$ pues

$$(a \oplus b) \oplus a = (a \oplus a) \oplus b = a \oplus b ,$$

y análogamente, $(a \oplus b) \oplus b = a \oplus (b \oplus b) = a \oplus b .$

Además si $c \succeq a$ y $c \succeq b$, resulta que $c = c \oplus a$ y $c = c \oplus b$, y entonces

$$c = c \oplus c = c \oplus a \oplus c \oplus b = c \oplus a \oplus b ,$$

lo cual implica que $c \succeq a \oplus b$. Por lo tanto, $a \oplus b$ es la menor de las cotas superiores de a y b .

- Finalmente como $a \oplus \varepsilon = a$, $\forall a \in S$, resulta que $a \succeq \varepsilon$, $\forall a \in S$, es decir que ε es el menor elemento de S (como ε es el elemento neutro de la suma, de esta última propiedad se podría concluir que todos los elementos de S son mayores que el “cero” y por lo tanto “positivos”). \square

Notemos que la relación de orden \succeq inducida por \oplus es un orden *total*, es decir

$$a \succeq b \quad \text{o bien} \quad b \succeq a , \quad \forall a, b \in S$$

si y solamente si

$$a \oplus b = a \quad \text{o} \quad a \oplus b = b , \quad \forall a, b \in S .$$

El teorema anterior nos dice que toda suma idempotente induce una estructura de *sup-semilattice*, es decir, induce una relación de orden parcial de tal manera que todo par de elementos posee una cota superior (recordemos que llamamos simplemente cota superior a la menor de las cota superiores). Podríamos pensar también en sentido inverso. Supongamos que D es un *sup-semilattice* (con elemento mínimo ε) y representemos con $a \vee b$ a la cota superior de a y b . Es fácil verificar que

si definimos la suma de dos elementos como su cota superior, entonces la operación suma resultante es idempotente (además de asociativa y conmutativa). Ahora la cuestión es la introducción de otra operación \otimes que verifique todos los axiomas de la definición de semi-anillos, en particular, debe ser distributiva respecto a la suma \vee .

Observación 2.7. Una consecuencia importante del hecho de que la suma \oplus sea idempotente es que un elemento a no posee opuesto (es decir, no existe b tal que $a \oplus b = \varepsilon$) salvo en el caso en que a es ε , ya que

$$a \oplus b = \varepsilon \Rightarrow \varepsilon \succeq a ;$$

pero como $a \succeq \varepsilon$, resulta entonces que $a = \varepsilon$.

También la suma \oplus no es cancelativa, es decir, $a \oplus b = a \oplus c$ no implica en general que $b = c$. Por ejemplo en $\mathbb{R}_{\text{máx}}$

$$5 \oplus 3 = \text{máx}\{5, 3\} = 5 = \text{máx}\{5, 1\} = 5 \oplus 1.$$

2.1.3. Ejemplos de semi-anillos idempotentes

En la siguiente tabla (extracta de [10]) pueden encontrarse los semi-anillos idempotentes más comunes así como su área de aplicación.

S	\oplus	\otimes	ε	e	Aplicación	Nombre
$\mathbb{R} \cup \{+\infty\}$	mín	+	$+\infty$	0	camino más corto	$\mathbb{R}_{\text{mín}}$
$\mathbb{R} \cup \{+\infty, -\infty\}$	mín	+	$+\infty$	0	camino más corto	$\mathbb{R}_{\text{mín}}$
$\mathbb{R} \cup \{-\infty\}$	máx	+	$-\infty$	0	camino más largo	$\mathbb{R}_{\text{máx}}$
$\mathbb{R}^+ \cup \{+\infty\}$	máx	mín	0	$+\infty$	capacidad máxima	$\mathbb{R}_{\text{máx,mín}}^+$
$[0, 1]$	máx	\times	0	1		
\mathbb{R}^+	máx	\times	0	1		$\mathbb{R}_{\text{máx},\times}^+$
$\{0, 1\}$	máx	mín	0	1	lógica	\mathbb{B}
$\mathcal{P}(\Sigma^*)$	\cup	prod. lat.	\emptyset	-	lenguajes	\mathbb{L}

Notemos que el orden inducido por la suma \oplus en $\mathbb{R}_{\text{mín}}$ es el orden inverso al habitual, es decir, \succeq es \leq (por ejemplo, como $1 \oplus 3 = \text{mín}\{1, 3\} = 1$, resulta que $1 \succeq 3$).

Ejemplo 2.8. Sea $2^{\mathbb{R}^2}$ el conjunto formado por todos los subconjuntos del plano \mathbb{R}^2 , incluyendo a \emptyset y a \mathbb{R}^2 . Definamos a la suma \oplus como \cup y a la multiplicación \otimes como $+$ (suma vectorial), es decir

$$A \otimes B = A + B = \{x \in \mathbb{R}^2 \mid x = y + z, y \in A, z \in B\} .$$

Entonces $2^{\mathbb{R}^2}$ con estas operaciones es un dioide en donde la relación de orden \succeq es \supseteq . Este es un ejemplo en donde la relación de orden es parcial.

Ejemplo 2.9. Un ejemplo similar en dimensión uno se obtiene considerando el subconjunto L de $2^{\mathbb{R}}$ compuesto por todos los intervalos de la forma $(-\infty, x]$ donde $x \in \mathbb{R}$, más el conjunto vacío, y con las mismas operaciones que en el ejemplo anterior. Notemos que este dioide es isomorfo (la definición exacta será dada más adelante) a $\mathbb{R}_{\text{máx}}$ por la biyección:

$$\mathbb{R} \cup \{-\infty\} \rightarrow L : x \mapsto \begin{cases} \emptyset & \text{si } x = -\infty; \\ (-\infty, x] & \text{en caso contrario.} \end{cases}$$

2.1.4. Subdioide

Definición 2.10. Un subconjunto D de un dioide S se dice que es un *subdioide* si

- $\varepsilon \in D$ y $e \in D$;
- D es cerrado con respecto a la suma \oplus y a la multiplicación \otimes , es decir, $a \oplus b \in D$, $a \otimes b \in D$, $\forall a, b \in D$.

Notemos que la primera condición es importante. Por ejemplo, \mathbb{B} no es un subdioide de $\mathbb{R}_{\text{máx}, \text{mín}}^+$ (ver la tabla anterior) pues $+\infty \notin \{0, 1\}$. El dioide $(\mathbb{Z} \cup \{-\infty\}, \text{máx}, +)$ es un subdioide de $\mathbb{R}_{\text{máx}}$.

2.2. Homomorfismos, isomorfismos y congruencias

2.2.1. Homomorfismos

Definición 2.11. Una función Π de un semi-anillo D en un semi-anillo S se dice que es un *homomorfismo* si

$$\Pi(a \oplus b) = \Pi(a) \oplus \Pi(b) \quad \text{y} \quad \Pi(\varepsilon) = \varepsilon, \quad \forall a, b \in D, \quad (2.1a)$$

$$\Pi(a \otimes b) = \Pi(a) \otimes \Pi(b) \quad \text{y} \quad \Pi(e) = e, \quad \forall a, b \in D. \quad (2.1b)$$

Por supuesto que las operaciones que aparecen en el lado izquierdo (respectivamente en el lado derecho) son las operaciones de D (respectivamente de S).

Supongamos que Π es sobreyectiva. Veamos entonces que la primera parte de (2.1a) (respectivamente (2.1b)) implica a la segunda parte, que por lo tanto resulta ser redundante. Debemos mostrar que $\Pi(\varepsilon)$ es un elemento nulo de S (y por lo tanto es igual a ε ya que en un semi-anillo el elemento nulo es único). Sea $c \in S$. Como Π es sobreyectiva sabemos que existe $a \in D$ tal que $\Pi(a) = c$. Entonces a partir de la primera parte de (2.1a) obtenemos que

$$c \oplus \Pi(\varepsilon) = \Pi(\varepsilon) \oplus c = \Pi(a) \oplus \Pi(\varepsilon) = \Pi(a \oplus \varepsilon) = \Pi(a) = c.$$

Como c es arbitrario, resulta entonces que $\Pi(\varepsilon) = \varepsilon$.

Una función que verifica únicamente la primera parte de (2.1a) (respectivamente la primera parte de (2.1b)) recibe el nombre de \oplus -*morfismo* (respectivamente \otimes -*morfismo*).

Definición 2.12. Sean D y S dos conjuntos ordenados.¹ Entonces se dice que una función Π de D en S es monótona si

$$a \succeq b \Rightarrow \Pi(a) \succeq \Pi(b), \quad \forall a, b \in D.$$

Si D y S son semi-anillos idempotentes, entonces todo \oplus -morfismo Π es una función monótona, ya que

$$a \succeq b \Rightarrow a = a \oplus b \Rightarrow \Pi(a) = \Pi(a \oplus b) = \Pi(a) \oplus \Pi(b) \Rightarrow \Pi(a) \succeq \Pi(b).$$

Sin embargo, una función monótona no es necesariamente un \oplus -morfismo. Una función monótona solamente verifica que

$$\Pi(a \oplus b) \succeq \Pi(a) \oplus \Pi(b), \quad \forall a, b \in D$$

pues

$$\left. \begin{array}{l} a \oplus b \succeq a \Rightarrow \Pi(a \oplus b) \succeq \Pi(a) \\ a \oplus b \succeq b \Rightarrow \Pi(a \oplus b) \succeq \Pi(b) \end{array} \right\} \Rightarrow \Pi(a \oplus b) \succeq \Pi(a) \oplus \Pi(b)$$

donde la última implicación se debe a que $\Pi(a) \oplus \Pi(b)$ es la menor de las cotas superiores de $\Pi(a)$ y $\Pi(b)$.

2.2.2. Isomorfismos

Definición 2.13. Una función Π de un semi-anillo D en un semi-anillo S se dice que es un *isomorfismo* si Π^{-1} existe y tanto Π como Π^{-1} son homomorfismos.

Lema 2.14. Si Π es un homomorfismo biyectivo de D en S , entonces Π es un isomorfismo.

Demostración. Bastaría probar que Π^{-1} verifica (2.1a). Sean $c, d \in S$. Entonces como

$$\Pi(\Pi^{-1}(c) \oplus \Pi^{-1}(d)) = \Pi(\Pi^{-1}(c)) \oplus \Pi(\Pi^{-1}(d)) = c \oplus d,$$

resulta que

$$\Pi^{-1}(c \oplus d) = \Pi^{-1}(c) \oplus \Pi^{-1}(d),$$

y por lo tanto Π^{-1} es un \oplus -morfismo. Análogamente se puede probar (2.1b). \square

¹Es decir que tanto en D como en S está definida una relación de orden (posiblemente parcial).

2.2.3. Congruencias

Definición 2.15. Una *congruencia* en un semi-anillo S es una relación de equivalencia \mathcal{R} en S que es compatible con la suma \oplus y con la multiplicación \otimes , es decir

$$a \mathcal{R} b \Rightarrow (a \oplus c) \mathcal{R} (b \oplus c), \quad (a \otimes c) \mathcal{R} (b \otimes c), \quad \forall a, b, c \in S.$$

(lo mismo para la multiplicación a izquierda por c).

Lema 2.16. *El espacio cociente de S por una congruencia \mathcal{R} (es decir, el conjunto de las clases de equivalencias S/\mathcal{R}) es un semi-anillo con la suma y la multiplicación inducidas por las operaciones de S .*

Demostración. La principal dificultad radica en mostrar que están bien definidas la suma y la multiplicación inducidas en S/\mathcal{R} ,² esto es, que no dependen de los representantes de las clases que se tomen. Supongamos que $a \mathcal{R} c$ y $b \mathcal{R} d$. Entonces debemos mostrar que $(a \oplus b) \mathcal{R} (c \oplus d)$ (con esto quedaría demostrado que la suma está bien definida). Pero

$$\left. \begin{array}{l} a \mathcal{R} c \Rightarrow (a \oplus b) \mathcal{R} (c \oplus b) \\ b \mathcal{R} d \Rightarrow (c \oplus b) \mathcal{R} (c \oplus d) \end{array} \right\} \Rightarrow (a \oplus b) \mathcal{R} (c \oplus d).$$

Un razonamiento análogo demuestra que la multiplicación está bien definida. La verificación de los axiomas de semi-anillos es directa a partir del hecho de que S es un semi-anillo. \square

Una clase especial de congruencia, que jugará un rol importante más adelante, está dada por el siguiente lema.

Lema 2.17. *Sea Π un homomorfismo del semi-anillo S en el semi-anillo D . Entonces la relación de equivalencia \mathcal{R}_Π definida por*

$$a \mathcal{R}_\Pi b \Leftrightarrow \Pi(a) = \Pi(b), \quad \forall a, b \in S,$$

es una congruencia.

Demostración. La verificación de que \mathcal{R}_Π es una relación de equivalencia es directa. Veamos que es una congruencia. Supongamos que $a \mathcal{R}_\Pi b$ y sea $c \in S$. Entonces como

$$\Pi(a \oplus c) = \Pi(a) \oplus \Pi(c) = \Pi(b) \oplus \Pi(c) = \Pi(b \oplus c),$$

resulta que $(a \oplus c) \mathcal{R}_\Pi (b \oplus c)$ y por lo tanto \mathcal{R}_Π es compatible con la suma. Un razonamiento análogo demuestra que \mathcal{R}_Π es compatible con la multiplicación. \square

²Estas operaciones están definidas por:

$$[a] \oplus [b] = [a \oplus b] \text{ y } [a] \otimes [b] = [a \otimes b],$$

donde $[a]$ es la clase de equivalencia de a , es decir $\{b \in S \mid b \mathcal{R} a\}$.

Notemos que el espacio cociente S/\mathcal{R}_Π es isomorfo a $\Pi(S)$ (el cual también es un semi-anillo) por la biyección:

$$S/\mathcal{R}_\Pi \rightarrow \Pi(S) : [a] \mapsto \Pi(a) .$$

Además S/\mathcal{R}_Π es isomorfo a S si Π es inyectiva.

2.3. Más acerca de la relación de orden en un dioide

Como ya hemos visto, un dioide tiene asociada naturalmente una relación de orden que le confiere una estructura de sup-semilattice. También hemos visto que un dioide posee un elemento *base o mínimo* (ε) que es menor que cualquier otro elemento. En esta sección estudiaremos más propiedades de los semilattices y lattices, así como su relación con la multiplicación \otimes .

2.3.1. Diodes completos

2.3.1.1. Completitud y completación

Definición 2.18. Un *sup-semilattice completo* es un conjunto ordenado en el cual cada subconjunto finito o infinito de elementos posee una menor cota superior. Análogamente se define un *inf-semilattice completo*. Un *lattice completo* es un inf-semilattice y sup-semilattice completo.

A partir de esta definición resulta natural definir:

Definición 2.19. Un dioide es *completo* si

- es un sup-semilattice completo (entonces se define la suma de un conjunto infinito de elementos como la menor cota superior de ese conjunto),
- la propiedad distributiva de la multiplicación con respecto a la suma se extiende a sumas infinitas (como veremos más adelante, esta condición puede pensarse como la propiedad de semicontinuidad inferior de la multiplicación).

Todo dioide completo S posee un elemento máximo \top ($\top \succeq a$, $\forall a \in S$) el cual está dado por la suma de todos los elementos de S , es decir

$$\top = \bigoplus_{a \in S} a .$$

Notemos que \top es absorbente para la suma pues $\top \oplus a = \top$, $\forall a \in S$ y además

$$\top \otimes \varepsilon = \left(\bigoplus_{a \in S} a \right) \otimes \varepsilon = \bigoplus_{a \in S} (a \otimes \varepsilon) = \varepsilon .$$

Ejemplo 2.20. $\mathbb{R}_{\text{máx}}$ no es un dioide completo ya que, por ejemplo, el conjunto \mathbb{N} no posee cotas superiores. Para que $\mathbb{R}_{\text{máx}}$ sea completo es necesario agregar el elemento $\top = +\infty$. El dioide completo que se obtiene de esta manera se representa por $\overline{\mathbb{R}}_{\text{máx}}$. Entonces $\top \otimes \varepsilon = \varepsilon$ se traduce en la regla $(-\infty) + (+\infty) = -\infty$. Observemos que $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ con la operación \otimes (la suma habitual) ya no tiene estructura de grupo pues no existe \top^{-1} .

Observación 2.21. Un subdioide de un dioide completo no es necesariamente completo. Por ejemplo, $(\mathbb{Q} \cup \{-\infty\} \cup \{+\infty\}, \text{máx}, +)$ es un subdioide de $\overline{\mathbb{R}}_{\text{máx}}$; pero no es completo.

2.3.1.2. Máxima cota inferior

Teorema 2.22. *Un sup-semilattice completo que posee un elemento mínimo o base es un lattice completo.*

Demostración. Sea C un subconjunto de un sup-semilattice completo S . Debemos probar que C tiene una máxima cota inferior (es decir, mayor que cualquier otra cota inferior de C). Sea T el subconjunto de todas las cotas inferiores de C . Este subconjunto es distinto del vacío pues contiene al menos al elemento mínimo o base. Sea c la mínima cota superior de T , la cual sabemos que existe pues S es un sup-semilattice completo. Veamos que c es la máxima cota inferior de C . Primero veamos que c es cota inferior de C (y por lo tanto pertenece a T). En realidad T está acotado superiormente por cada $b \in C$ (por definición de T). Como c es menor o igual que cualquier cota superior de T (por definición de mínima cota superior), resulta que $b \succeq c, \forall b \in C$, y por lo tanto $c \in T$. Entonces, c es menor o igual que todos los elementos de C y mayor o igual que cualquier cota inferior de C (los elementos de T). Por lo tanto, c es la máxima cota inferior de C . \square

Como un dioide completo S es un sup-semilattice completo y además posee un elemento mínimo ε , por el teorema anterior sabemos que todo subconjunto de S posee una máxima cota inferior (que desde ahora en adelante llamaremos simplemente cota inferior). La cota inferior de a y b será representada por $a \wedge b$. Entonces se verifica la siguiente equivalencia:

$$a \succeq b \Leftrightarrow a = a \oplus b \Leftrightarrow b = a \wedge b, \quad \forall a, b \in S. \quad (2.2)$$

Notemos que la operación \wedge es asociativa, conmutativa, idempotente y tiene elemento neutro \top . Además se verifica que

$$a \wedge (a \oplus b) = a \oplus (a \wedge b) = a, \quad \forall a, b \in S,$$

pues

$$(a \oplus b) \succeq a \Rightarrow a \wedge (a \oplus b) = a$$

$$\text{y } a \succeq (a \wedge b) \Rightarrow a \oplus (a \wedge b) = a, \quad \forall a, b \in S.$$

Esta última propiedad recibe el nombre de *ley de absorción*.

Ejemplo 2.23. En el dioide $(2^{\mathbb{R}^2}, \cup, +)$ la cota inferior \wedge es simplemente la \cap .

2.3.1.3. Relación entre \otimes y \wedge

La equivalencia (2.2) puede dejar la impresión de que la suma \oplus y la cota inferior \wedge juegan roles simétricos en un dioide completo. Esto es verdad desde el punto de vista de la estructura de lattice; pero no es verdad en lo que concierne a la relación con la segunda operación de un dioide, es decir, la multiplicación \otimes . Como la multiplicación a izquierda por c (análogamente para la multiplicación a derecha) es una función monótona (ver Teorema 2.6) se verifica que

$$(ca) \wedge (cb) \succeq c(a \wedge b), \quad \forall a, b, c \in S$$

pues

$$\left. \begin{array}{l} a \succeq a \wedge b \Rightarrow (ca) \succeq c(a \wedge b) \\ b \succeq a \wedge b \Rightarrow (cb) \succeq c(a \wedge b) \end{array} \right\} \Rightarrow (ca) \wedge (cb) \succeq c(a \wedge b), \quad \forall a, b, c \in S.$$

Esta propiedad recibe el nombre de sub-distributividad de \otimes con respecto a \wedge . Cuando un dioide es totalmente ordenado se puede demostrar la distributividad de \otimes con respecto a \wedge ; pero esto no es cierto en general para un dioide con una relación de orden parcial, como lo demuestra el siguiente ejemplo.

Ejemplo 2.24. Consideremos el dioide $(2^{\mathbb{R}^2}, \cup, +)$. Si tomamos

$$a = \{(1, 0)\}, \quad b = \{(0, 1)\} \quad \text{y} \quad c = [-1, 1] \times [-1, 1]$$

entonces resulta que

$$\begin{aligned} c(a \wedge b) &= c + (a \cap b) = c + \emptyset = \emptyset, \\ (ca) \wedge (cb) &= (c + a) \cap (c + b) \\ &= ([0, 2] \times [-1, 1]) \cap ([-1, 1] \times [0, 2]) \\ &= [0, 1] \times [0, 1], \end{aligned}$$

y por lo tanto $(ca) \wedge (cb) \succeq c(a \wedge b)$; pero $(ca) \wedge (cb) \neq c(a \wedge b)$.

Existen situaciones en las cuales la distributividad de \otimes con respecto a \wedge se verifica para ciertos elementos. En el siguiente lema se presenta uno de estos casos.

Lema 2.25. Si a admite una inversa a izquierda b y una inversa a derecha c , entonces

- $b = c$ y a esta única inversa la representaremos por a^{-1} ;
- $a(x \wedge y) = ax \wedge ay, \quad \forall x, y.$

Lo mismo se verifica para la multiplicación a derecha por a y la multiplicación tanto a izquierda como a derecha por a^{-1} .

Demostración.

- Como $b = be = b(ac) = (ba)c = ec = c$, la inversa a izquierda y a derecha es única.
- Sean $\xi = ax$, $\eta = ay$ (lo cual es equivalente a que $x = a^{-1}\xi$, $y = a^{-1}\eta$). Entonces

$$\xi \wedge \eta = aa^{-1}(\xi \wedge \eta) \preceq a[a^{-1}\xi \wedge a^{-1}\eta] = a[x \wedge y] \preceq ax \wedge ay = \xi \wedge \eta,$$

y por lo tanto $a(x \wedge y) = ax \wedge ay$. \square

2.3.2. Dioide arquimediano

Supongamos que S es un dioide completo. Nos preguntamos ahora cuando el elemento máximo \top es absorbente para la multiplicación, es decir, cuando

$$\top \otimes a = a \otimes \top = \top, \quad \forall a \in S \setminus \{\varepsilon\}. \quad (2.3)$$

Esta propiedad puede demostrarse para los dioides arquimediano que definimos a continuación.

Definición 2.26. Un dioide S se dice *arquimediano* si

$$\forall a \in S \setminus \{\varepsilon\}, \forall b \in S, \exists c \text{ y } d \in S: a \otimes c \succeq b \text{ y } d \otimes a \succeq b.$$

Ejemplo 2.27. Todos los dioides que presentamos son arquimediano salvo por el dioide $\mathbb{R}_{\text{máx,mín}}^+$.

Teorema 2.28. Si S es un dioide completo y arquimediano, entonces

$$\top \otimes a = a \otimes \top = \top, \quad \forall a \in S \setminus \{\varepsilon\}.$$

Demostración. Probaremos esta propiedad para la multiplicación a izquierda únicamente (la demostración para la multiplicación a derecha es análoga). Por la definición de dioide arquimediano sabemos que dado $a \in S \setminus \{\varepsilon\}$, para cada $b \in S$ existe $c_b \in S$ tal que $a \otimes c_b \succeq b$. Entonces

$$a \otimes \top = a \otimes \left(\bigoplus_{b \in S} b \right) \succeq a \otimes \left(\bigoplus_{b \in S} c_b \right) = \bigoplus_{b \in S} (a \otimes c_b) \succeq \bigoplus_{b \in S} b = \top,$$

y por lo tanto $a \otimes \top = \top$. \square

Teorema 2.29. Si un dioide S es completo, arquimediano, y si la multiplicación es cancelativa, entonces es isomorfo al álgebra de Boole \mathbb{B} .

Demostración. Como se verifica (2.3) y como la multiplicación es cancelativa, resulta que todo elemento diferente de ε es igual a \top . Por lo tanto, S se reduce a $\{\varepsilon, \top\}$. \square

2.3.3. Dioide distributivo

Estamos interesados ahora en la relación entre las operaciones \oplus y \wedge . La cota inferior \wedge no es necesariamente distributiva con respecto a la cota superior \oplus , ni tampoco la cota superior \oplus es necesariamente distributiva con respecto a la cota inferior \wedge (la distributividad sólo puede asegurarse en el caso de cadenas, es decir, cuando la relación de orden es total). Como las funciones $x \mapsto x \oplus c$ y $x \mapsto x \wedge c$ son monótonas para cada c , tenemos que se verifican las siguientes desigualdades:

$$(a \wedge b) \oplus c \leq (a \oplus c) \wedge (b \oplus c)$$

$$\text{y } (a \oplus b) \wedge c \geq (a \wedge c) \oplus (b \wedge c), \quad \forall a, b, c \in S,$$

lo que significa que \oplus es *subdistributiva* con respecto a \wedge , y que \wedge es *superdistributiva* con respecto a \oplus .

Definición 2.30. Un lattice³ S se dice distributivo si

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \quad \forall a, b, c \in S.$$

Se puede demostrar que la condición anterior es equivalente a la condición:

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \quad \forall a, b, c \in S.$$

Ejemplo 2.31. El siguiente es un ejemplo de un lattice completo que no es distributivo. Consideremos el conjunto de todos los intervalos de \mathbb{R} (incluyendo a \emptyset y a \mathbb{R}) con la relación de orden \geq definida por \supseteq . Entonces la cota superior \vee (o \oplus) de cualquier familia (finita o infinita) de intervalos es el menor intervalo que contiene a toda la familia, es decir, es la envolvente convexa de la unión de todos los intervalos de la familia. La cota inferior \wedge es simplemente la \cap . Entonces, tomando $a = [-3, -2]$, $b = [2, 3]$ y $c = [-1, 1]$ obtenemos que $(a \vee b) \wedge c = c$, mientras que $(a \wedge c) \vee (b \wedge c) = \emptyset$.

Definición 2.32. Un dioide S se dice *distributivo* si es completo y además se verifica que

$$\left(\bigwedge_{c \in C} c \right) \oplus a = \bigwedge_{c \in C} (c \oplus a),$$

$$\left(\bigoplus_{c \in C} c \right) \wedge a = \bigoplus_{c \in C} (c \wedge a),$$

para cada subconjunto C de S y $a \in S$.

³Un lattice es un sup y un inf-semilattice, es decir, un conjunto ordenado en el cual todo par de elementos posee una (máxima) cota inferior y una (mínima) cota superior. Representaremos con $a \vee b$ (respectivamente con $a \wedge b$) a la cota superior (respectivamente inferior) de a y b .

Notemos que en esta definición se pide que la propiedad distributiva se verifique también para subconjuntos infinitos. En este caso ambas condiciones deben verificarse, ya que para subconjuntos infinitos, las dos condiciones no son más equivalentes como se afirmaba en la definición de lattice distributivo.

Ejemplo 2.33. El ejemplo anterior nos proporciona un ejemplo de un dioide no distributivo si definimos la multiplicación \otimes como la suma (aritmética convencional) $+$ de intervalos (lo único que es necesario verificar es la propiedad distributiva de \otimes con respecto a \oplus).

Observación 2.34. Un dioide distributivo tiene también una estructura de dioide con las operaciones $\tilde{\oplus} \stackrel{\text{def}}{=} \oplus$ y $\tilde{\otimes} \stackrel{\text{def}}{=} \wedge$. Lo mismo sucede con $\tilde{\oplus} \stackrel{\text{def}}{=} \wedge$ y $\tilde{\otimes} \stackrel{\text{def}}{=} \oplus$. Las características especiales de este dioide es que $\tilde{\top}$ coincide con \tilde{e} y que además la multiplicación es conmutativa e idempotente.

2.4. Dioides derivados de un dioide “escalar”

2.4.1. Dioide de matrices

2.4.1.1. Desde los escalares hacia las matrices

Sea S un dioide “escalar” arbitrario. Consideremos el conjunto $S^{n \times n}$ formado por todas las matrices cuadradas $n \times n$ con elementos en S . Definamos sobre este conjunto la suma y la multiplicación convencionales a partir de las operaciones suma \oplus y multiplicación \otimes de S , es decir, definamos

$$[A \oplus B]_{ij} = A_{ij} \oplus B_{ij}, \quad [A \otimes B]_{ij} = \bigoplus_k (A_{ik} \otimes B_{kj}), \quad \forall i, j = 1, \dots, n, \\ \forall A, B \in S^{n \times n}.$$

Entonces el conjunto $S^{n \times n}$ con estas operaciones adquiere una estructura de dioide. Notemos que la matriz identidad (es decir, el elemento neutro de la multiplicación) es la matriz que tiene en la diagonal a e (el elemento neutro de la multiplicación de S) y fuera de la diagonal a ε (el elemento neutro de la suma de S). La matriz identidad será representada por e . El elemento neutro de la suma en $S^{n \times n}$ es la matriz con todas sus componentes iguales a ε , la cual también será representada por ε .

Observación 2.35. Hemos considerado matrices cuadradas únicamente pues la multiplicación de dos matrices de $S^{n \times n}$ es una matriz de $S^{n \times n}$ (la multiplicación es una operación cerrada en $S^{n \times n}$) y entonces también es posible darle una estructura de dioide a $S^{n \times n}$. Sin embargo, desde el punto de vista práctico y para la mayoría de los temas que consideraremos más adelante, en particular para los sistemas de ecuaciones lineales, nosotros trabajaremos con matrices que no son cuadradas, y en especial con vectores fila y columna. Entonces es necesario completar las matrices que no son cuadradas con vectores fila o columna con componentes iguales a ε de forma tal de convertirlas en matrices cuadradas, y luego se debe verificar, que para

el problema considerado, esta parte adicional no interfiere en el problema, y por lo tanto, la misma sólo agrega una parte trivial al problema.

Notemos que aún cuando S sea un dioide *conmutativo*,⁴ $S^{n \times n}$ generalmente no lo es.

Como la relación de orden inducida en $S^{n \times n}$ es

$$A \succeq B \text{ en } S^{n \times n} \Leftrightarrow A_{ij} \succeq B_{ij} \text{ en } S, \quad \forall i, j = 1, \dots, n,$$

(esto es consecuencia de que la suma en $S^{n \times n}$ está definida componente a componente) resulta que por más que S sea totalmente ordenado, $S^{n \times n}$ no lo es. Además, si S es completo, $S^{n \times n}$ también lo es y se verifica que:

$$(A \wedge B)_{ij} = A_{ij} \wedge B_{ij}, \quad \forall i, j = 1, \dots, n, \quad \forall A, B \in S^{n \times n}.$$

Si S es distributivo, $S^{n \times n}$ también es distributivo. Aún cuando S sea arquimediano, $S^{n \times n}$ no es arquimediano como lo muestra el siguiente ejemplo.

Ejemplo 2.36. Tomemos $n = 2$ y consideremos las matrices

$$A = \begin{pmatrix} a & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & b \end{pmatrix}$$

donde $a \neq \varepsilon, b \neq \varepsilon$. Observemos entonces que no existe una matriz C tal que $AC \succeq B$, pues siempre $(AC)_{22} = \varepsilon$; pero $b \neq \varepsilon$. Por lo tanto, $S^{2 \times 2}$ no es arquimediano cualquiera sea el dioide S .

Observación 2.37. Las matrices difícilmente tienen inversas o son sobreyectivas o inyectivas aún en el caso en el cual la multiplicación “escalar” posee una estructura de grupo como en $\mathbb{R}_{\text{máx}}$. Por ejemplo, la matriz

$$\begin{pmatrix} e & e \\ \varepsilon & e \end{pmatrix}$$

(cualquiera sea el dioide) tiene rango 2 en el sentido de que ni sus columnas ni sus filas son proporcionales. Sin embargo, esta matriz no es sobreyectiva ya que por ejemplo el vector $(\varepsilon \ e)^t$ no pertenece a su imagen (notemos que en la imagen de esta matriz, la primera componente de cada vector es mayor o igual que la segunda componente, pues esto se verifica para los vectores columna de la matriz).

2.4.1.2. Grafos asociados a las matrices

Existen dos tipos de grafos dirigidos con pesos asociados a sus arcos que pueden asociarse a una matriz con componentes en un dioide.

Definición 2.38. El *grafo de transición* asociado a una matriz A de dimensión $n \times p$, es un grafo bipartito dirigido con $n + p$ vértices, en el cual existe el arco que va del vértice $j \in \{1, \dots, p\}$ al vértice $p + i$ con $i \in \{1, \dots, n\}$ si $A_{ij} \neq \varepsilon$.

⁴Un dioide se dice conmutativo cuando la multiplicación es conmutativa.

En dicho caso, el peso asociado a este arco es A_{ij} (ver Figura 2.1).

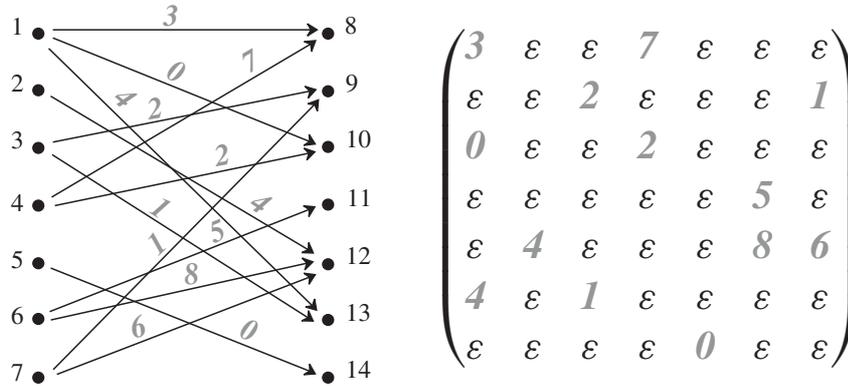


Figura 2.1: Grafo de transición de una matriz 7×7

Definición 2.39. El *grafo de precedencia*⁵ asociado a una matriz A de dimensión $n \times n$, es un grafo dirigido con n vértices, en el cual existe el arco que va del vértice $j \in \{1, \dots, n\}$ al vértice $i \in \{1, \dots, n\}$ si $A_{ij} \neq \varepsilon$. En dicho caso, el peso asociado a este arco es A_{ij} (ver Figura 2.2).

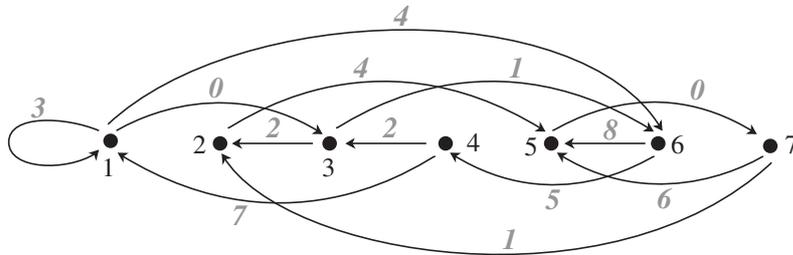


Figura 2.2: Grafo de precedencia asociado a la matriz de la Figura 2.1

Observemos que el grafo de transición de una matriz cuadrada $n \times n$, el cual posee $2n$ vértices, puede transformarse en el grafo de precedencia de dicha matriz (con n vértices) de la siguiente manera: basta con unir el vértices i con el vértice $n+i$ del grafo de transición para obtener el vértice i del grafo de precedencia, donde $i = 1, \dots, n$. Por ejemplo, observemos que es posible utilizar este procedimiento para obtener el grafo de la Figura 2.2 a partir del grafo de la Figura 2.1.

La suma de dos matrices A y B de igual dimensión, corresponde a la composición en paralelo de los respectivos grafos de transición o de precedencia, es decir, existe un arco entre dos vértices si existe un arco entre dichos vértices en al menos

⁵Esta definición se aplica únicamente a matrices cuadradas.

uno de los grafos asociados a las matrices y su peso asociado es la suma (\oplus) de los pesos asociados a dicho arco en los dos grafos.

El producto $C = A \otimes B$ de A por B (A de dimensión $m \times n$ y B de dimensión $n \times p$) corresponde a la composición en serie de los respectivos grafos de transición como se representa en la Figura 2.3. En este caso, el peso C_{ij} del arco que va del vértice j al vértice i en el grafo de transición asociado a la matriz C , se obtiene “sumando” los pesos de todos los caminos paralelos⁶ (de longitud dos) que van del vértice j al vértice i en el grafo que se obtiene al concatenar las grafos de transición de las matrices A y B . Notemos que si el dioide considerado es totalmente ordenado, entonces esto corresponde a quedarse con el camino de máximo peso que va de j a i .

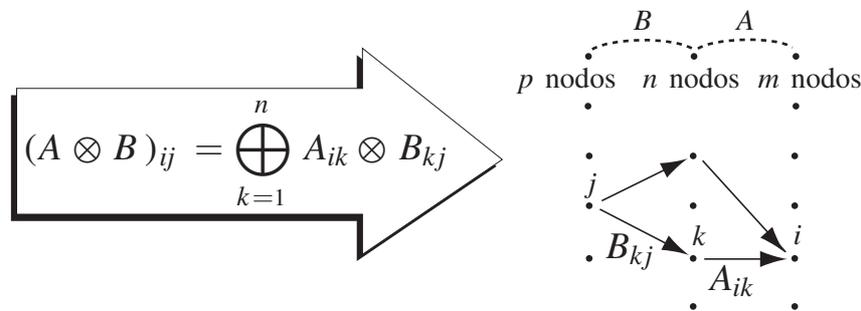


Figura 2.3: Producto de matrices y composición en serie

Como consecuencia de esta interpretación, la componente (i, j) de la matriz A^k resulta ser el peso máximo entre los pesos de los caminos de longitud k que van del vértice j al vértice i en el grafo de precedencia de A . En realidad, si el dioide no es totalmente ordenado, es la cota superior del conjunto de los pesos de estos caminos. Análogamente, la componente (i, j) de la matriz $A \oplus A \oplus \dots \oplus A^k$ es el peso máximo entre los pesos de todos los caminos de longitud menor o igual que k que van del vértice j al vértice i

2.4.2. Dioides de polinomios y series de potencias

Comenzando con un dioide de “escalares” S , podemos considerar el conjunto de las series de potencias o polinomios formales en una o varias variables con coeficientes en S . En el caso de varias variables (por ejemplo, z_1 y z_2), nosotros solamente consideraremos variables conmutativas (es decir que por ejemplo, $z_1 z_2$ y $z_2 z_1$ son el mismo objeto). Los exponentes k_i de z_i pueden considerarse tanto en \mathbb{N} como en \mathbb{Z} ; en este ultimo caso hablaremos de *series de Laurent*.

⁶En un grafo, el peso de un camino es el “producto” de los pesos de los arcos que lo componen (si la multiplicación \otimes no es conmutativa, entonces es importante el orden en que aparecen los arcos en el camino). La longitud de un camino es el número de arcos que lo componen.

Definición 2.40. Una *serie de potencias formal* en p variables (conmutativas) con coeficientes en S es una función f de \mathbb{N}^p o \mathbb{Z}^p en S . Para todo $k = (k_1, \dots, k_p) \in \mathbb{N}^p$ o \mathbb{Z}^p , $f(k)$ representa el coeficiente de $z_1^{k_1} \dots z_p^{k_p}$.

Definición 2.41. Un *polinomio* es una serie de potencias f con soporte finito, es decir, el conjunto

$$\text{sop}(f) = \{k \in \mathbb{Z}^p \mid f(k) \neq \varepsilon\}$$

es finito.

Observación 2.42. Contrariamente a lo que sucede en el álgebra lineal habitual, no debe confundirse a un polinomio o a una serie de potencias formal (con coeficientes en $\overline{\mathbb{R}}_{\text{máx}}$) con las funciones numéricas que ellas definen. A cada polinomio o serie de potencia f está asociada una única función \hat{f} definida por

$$\hat{f} : (\overline{\mathbb{R}}_{\text{máx}})^p \rightarrow \overline{\mathbb{R}}_{\text{máx}}, \quad z \mapsto \hat{f}(z) = \bigoplus_{k \in \mathbb{Z}^p} f(k_1, \dots, k_p) z_1^{k_1} \dots z_p^{k_p};$$

pero dos polinomios formales diferentes pueden tener asociada la misma función polinomial, con lo cual la correspondencia entre polinomios formales y funciones polinomiales no es uno a uno. Por ejemplo, los polinomios formales $f(z) = 2 \oplus z^2$ y $g(z) = 2 \oplus 1z \oplus z^2$ definen la misma función polinomial pues

$$2 \oplus z^2 = (1 \oplus z)^2 = 2 \oplus 1z \oplus z^2, \quad \forall z \in \overline{\mathbb{R}}_{\text{máx}}.$$

En el conjunto de las series de potencias formales se definen las siguientes operaciones:

$$\begin{aligned} f \oplus g : (f \oplus g)(k) &= f(k) \oplus g(k), \\ f \otimes g : (f \otimes g)(k) &= \bigoplus_{i+j=k} f(i) \otimes g(j). \end{aligned}$$

Estas son las definiciones habituales de suma y multiplicación de series de potencias. Notemos que el producto no es otra cosa que la convolución. Es fácil verificar que el conjunto de las series de potencias con estas operaciones es un dioide el cual representaremos con $S[[z_1, \dots, z_p]]$. Como es habitual utilizamos los mismos símbolos para representar a las operaciones de S y a las de $S[[z_1, \dots, z_p]]$. El elemento neutro de la suma de este nuevo dioide, que representaremos nuevamente con ε , está definido por $f(k) = \varepsilon, \forall k$, y el elemento neutro de la multiplicación, que representaremos con e , está dado por: $f(0, \dots, 0) = e$ y $f(k) = \varepsilon, \forall k \neq (0, \dots, 0)$. Observemos que cuando k varía en \mathbb{Z}^p , la definición de $f \otimes g$ involucra sumas infinitas, y entonces para que la misma tenga sentido, es necesario suponer que S es completo (esta condición no es necesaria para definir el dioide de los polinomios). El subconjunto de los polinomios es un subdioide de $S[[z_1, \dots, z_p]]$ que representaremos con $S[z_1, \dots, z_p]$.

La relación de orden inducida en el conjunto de las series está dada por

$$f \succeq g \Leftrightarrow f(k) \succeq g(k) \quad , \quad \forall k \quad ,$$

y por lo tanto, $S[[z_1, \dots, z_p]]$ no es totalmente ordenado aún cuando S si lo fuese. El dioide $S[[z_1, \dots, z_p]]$ es conmutativo si S es conmutativo (esto se debe a que consideramos variables conmutativas únicamente). Si S es completo, entonces $S[[z_1, \dots, z_p]]$ también es completo; pero $S[z_1, \dots, z_p]$ no lo es, como lo muestra el siguiente ejemplo.

Ejemplo 2.43. Consideremos el subconjunto infinito de polinomios $\left\{ z_1^k \dots z_p^k \right\}_{k \in \mathbb{N}}$. Como la suma de los elementos de este subconjunto no es un polinomio, resulta que $S[z_1, \dots, z_p]$ no es completo.

Si en S están definidas las cotas inferiores, por ejemplo cuando S es completo, entonces estas cotas inferiores se extienden a $S[z_1, \dots, z_p]$ y a $S[[z_1, \dots, z_p]]$ por

$$(f \wedge g)(k) = f(k) \wedge g(k) \quad , \quad \forall k \quad .$$

Si S es distributivo, entonces $S[[z_1, \dots, z_p]]$ también es distributivo. Aún cuando S sea arquimediano, $S[[z_1, \dots, z_p]]$ y $S[z_1, \dots, z_p]$ no lo son necesariamente cuando los exponentes varían en \mathbb{N}^p , como lo indica el siguiente ejemplo.

Ejemplo 2.44. Sea $p = 1$, $f = z$ y $g = e$. Notemos entonces que no existe h tal que $f \otimes h \succeq g$, pues $(f \otimes h)(0) = \varepsilon$ para toda serie h y además $g(0) = e$.

Cuando los exponentes varían en \mathbb{Z}^p tenemos el siguiente lema.

Lema 2.45. Si S es arquimediano, entonces $S[[z_1, \dots, z_p]]$ y $S[z_1, \dots, z_p]$ son también arquimediano si los exponentes varían en \mathbb{Z}^p .

Demostración. Sean $f \neq \varepsilon$ y g dos series de potencias. Debemos mostrar que existe una serie de potencias h tal que $f \otimes h \succeq g$. Como $f \neq \varepsilon$, existe $l \in \mathbb{Z}^p$ tal que $f(l) \neq \varepsilon$. Definamos la serie \hat{f} por: $\hat{f}(l) = f(l)$ y $\hat{f}(k) = \varepsilon$ cuando $k \neq l$. Claramente $f \succeq \hat{f}$, y entonces es suficiente mostrar que existe h tal que $\hat{f} \otimes h \succeq g$. Notemos que $(\hat{f} \otimes h)(k) = \hat{f}(l) \otimes h(k-l)$. Como S es arquimediano, para cada $k \in \mathbb{Z}^p$, existe un a_k tal que $\hat{f}(l) \otimes a_k \succeq g(k)$. Entonces basta con definir $h(k) = a_{k+l}$ para que se cumpla que $\hat{f} \otimes h \succeq g$. Notemos que si g es un polinomio, entonces h también lo es. \square

2.5. La ecuación implícita $x = ax \oplus b$

En esta sección estudiaremos la ecuación implícita $x = ax \oplus b$ la cual, como veremos más adelante, juega un rol importante. Consideraremos esta ecuación en el contexto de un dioide completo como lo son en particular $\overline{\mathbb{Z}}_{\text{máx}}$, $\overline{\mathbb{Z}}_{\text{mín}}$, $\overline{\mathbb{Z}}_{\text{máx}}^{n \times n}$, $\overline{\mathbb{Z}}_{\text{máx}}[[z]]$ y $\left(\overline{\mathbb{Z}}_{\text{máx}}[[z]] \right)^{n \times n}$.

Teorema 2.46. *La menor solución de $x = ax \oplus b$, y también de la desigualdad $x \succeq ax \oplus b$, es a^*b donde a^* (la estrella de Kleene de a) está definida por*

$$a^* \stackrel{\text{def}}{=} e \oplus a \oplus a^2 \oplus \dots . \quad (2.4)$$

Demostración. En primer lugar notemos que a^* existe porque el dioide se supone completo. Como

$$a(a^*b) \oplus b = a \left(\bigoplus_{k=0}^{\infty} a^k b \right) \oplus b = \bigoplus_{k=1}^{\infty} a^k b \oplus b = \bigoplus_{k=0}^{\infty} a^k b = a^*b ,$$

a^*b es solución de la ecuación $x = ax \oplus b$ y de la desigualdad $x \succeq ax \oplus b$. Si x es solución de $x = ax \oplus b$ (respectivamente de $x \succeq ax \oplus b$), entonces luego de k reemplazos de x por $ax \oplus b$ en el lado derecho de esta igualdad (respectivamente desigualdad) obtenemos que

$$x = a^{k+1}x \oplus (e \oplus a \oplus \dots \oplus a^k) b \succeq (e \oplus a \oplus \dots \oplus a^k) b$$

(respectivamente que $x \succeq a^{k+1}x \oplus (e \oplus a \oplus \dots \oplus a^k) b \succeq (e \oplus a \oplus \dots \oplus a^k) b$). Como esto es válido para todo k , obtenemos que $x \succeq a^*b$. \square

Capítulo 3

Grafos de eventos temporizados como sistemas lineales

Consideramos Grafos de Eventos Temporizados (GET) con pesos de arcos igual a 1 y temporización constante y solamente en lugares. Demostraremos que tales sistemas pueden ser modelados como sistemas “lineales”; pero con un significado diferente al visto anteriormente.

3.1. Punto de vista del “dater”

3.1.1. ¿Qué es necesario para comenzar una simulación ?

Es conveniente considerar que los GETs estan “delimitados por transiciones”, es decir, todos los lugares tienen transiciones anteriores y posteriores. Esto no implica ninguna restricción real porque:

- cualquier lugar de entrada tendría sus marcas provistas desde el exterior, las que pueden pensarse como provenientes de una transición anterior controlada desde el exterior.
- cualquier lugar de salida puede estar seguido por una transición que se activaría únicamente por las marcas que arriban a dicho lugar.

Llamaremos a las transiciones de entrada: u_j , siendo $j = 1, \dots, m$, etc. los daters asociados. De la misma forma, las transiciones de salida serán llamadas y_l , $l = 1, \dots, p$, y por último, las transiciones internas se denominarán x_i , $i = 1, \dots, n$.

Adoptando la regla de que las transiciones sean disparadas *tan pronto como sea posible*, y teniendo en cuenta la ausencia de conflictos, lo único que se necesita conocer para realizar una simulación es:

- cuando las transiciones de entrada son activadas (por decisiones externas) durante todo el transcurso de la simulación;
- cuando las marcas presentes en el mercado inicial estarán disponibles (puede considerarse que dichas marcas estaban presentes desde cierto tiempo antes de comenzada la simulación).

Conocido esto, será posible determinar cuándo ocurrirán todas las transiciones internas y de salida.

3.1.2. Ecuaciones de daters

Presentamos los *daters* asociados con cada transición. Para una transición llamada x_i , la variable correspondiente $x_i(k)$ se interpreta como el instante en el cual ocurrirá el k -ésimo disparo de la misma. Desde el comienzo de la simulación los disparos sucesivos de una transición son numerados de manera secuencial, a partir de un origen universal (generalmente cero, aunque podría ser negativo). Luego, la función $k \mapsto x_i(k)$ es *no decreciente* (debido a que varios disparos pueden ocurrir simultáneamente, *puede no ser estrictamente creciente*)

El tiempo puede medirse en una escala *real*, *racional* o *entera*. Por esto $x_i(k) \in \mathbb{R}$ o \mathbb{Q} o \mathbb{Z} .

Las ecuaciones de daters resultan de las siguientes consideraciones:

- si la transición x_i está ubicada posteriormente a la x_j y separada de esta por un lugar denominado P_{ij} , entonces el k -ésimo disparo de x_i consumirá la marca producida por el disparo número $k - M_{ij}$ de x_j , siendo M_{ij} el marcado inicial del lugar P_{ij} ;

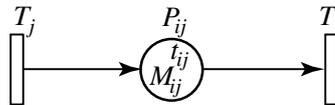


Figura 3.1:

- si el tiempo de espera en el lugar P_{ij} es t_{ij} , el disparo número k de la transición x_i no puede ocurrir hasta que hayan transcurrido al menos t_{ij} unidades de tiempo desde el disparo número $k - M_{ij}$ de la transición x_j ;
- teniendo en cuenta esta consideración para todas las transiciones x_j anteriores a x_i , el máximo de todos estos instantes determinará el instante del disparo número k de la transición x_i .

Finalmente, sean $\bullet i$ los índices de las transiciones anteriores a x_i . La ecuación fundamental del GET es:

$$x_i(k) = \max_{j \in \bullet i} (x_j(k - M_{ij}) + t_{ij}) . \quad (3.1)$$

3.1.3. Inicialización

Las ecuaciones vistas son válidas siempre y cuando las marcas consideradas hayan sido producidas por disparos de las transiciones durante la simulación. Si el numerado de los disparos comienza con $k = 0$, estas ecuaciones valdrán cuando

$k \geq M_{ij}$. Sin embargo, las ecuaciones son válidas sin restricciones cuando las marcas del marcado inicial no contribuyen al máximo. Esto es cierto con la siguiente condición suficiente:

las marcas del marcado inicial están *listas para usar* en tiempo $-\infty$. Se habla entonces de *condiciones iniciales canónicas*.

Más adelante, se verá también como forzar condiciones iniciales arbitrarias (no canónicas).

3.1.4. Forma matricial

De la ecuación genérica (3.1) (válida con la restricción hecha en §3.1.3), resulta evidente que la forma general de la ecuación de dater del GET completo es la siguiente:

$$x(k) = A_0x(k) \oplus A_1x(k-1) \oplus \dots \oplus B_0u(k) \oplus B_1u(k-1) \oplus \dots, \quad (3.2a)$$

$$y(k) = C_0x(k) \oplus C_1x(k-1) \oplus \dots \oplus D_0u(k) \oplus D_1u(k-1) \oplus \dots, \quad (3.2b)$$

- $x(\cdot), u(\cdot), y(\cdot)$ son vectores columna de dimensiones n, m y p respectivamente;
- A_i, B_i, C_i, D_i son matrices de dimensiones $n \times n, n \times m, p \times n$ y $p \times m$; El número máximo de matrices (no nulas) de cada tipo es igual al máximo marcado inicial de los lugares del GET, como se explicará más adelante;
- la regla para conformar el elemento (r, s) de la matriz A_i es el siguiente: si r es una transición *interna* inmediatamente posterior a la transición *interna* s y si hay i marcas en el marcado inicial del lugar P_{rs} , entonces $(A_i)_{rs}$ es no nula (i.e., distinta de ε) y es igual al *tiempo de espera* del lugar P_{rs} . En otras palabras, si se considera el grafo del GET con las transiciones como nodos y los lugares como arcos y si se mantienen solamente los nodos internos y los arcos con exactamente i marcas iniciales, entonces este es el *grafo de precedencia* de A_i (con pesos en los arcos igual a los tiempos de espera de los lugares correspondientes);
- de la misma forma, B_i se basa en el grafo que mantiene solamente los nodos correspondientes a las transiciones de entrada e internas y arcos con exactamente i marcas iniciales entre una transición de *entrada* y una transición *interna*; pero ahora este es el *grafo de transición* correspondiente;
- también de forma análoga, C_i se basa en el grafo que mantiene los nodos internos y de salida, y arcos con exactamente i marcas iniciales entre una transición *interna* y una transición de *salida*, siendo este también el *grafo de transición* de dicha matriz.

- la matriz D_i se define igual que las anteriores; sólo que ahora el grafo mantiene solamente los nodos de entrada y salida y arcos con exactamente i marcas iniciales;
- el álgebra utilizada es máx-plus;
- las *condiciones iniciales* son $x(k) = \varepsilon$ para todo k negativo, lo que refleja la suposición de §3.1.3 asumiendo también que el primer disparo de cada transición que modifica el marcado inicial del lugar anterior es el número 0.

3.1.5. Forma canónica

3.1.5.1. Eliminación de la parte implícita

Las ecuaciones (3.2) son aparentemente implícitas ya que $x(k)$ está presente en ambos lados de (3.2a). Desde un punto de vista matemático, se vio como resolver tales ecuaciones implícitas en §2.5. Utilizando el Teorema 2.46, sabemos que la menor solución de (3.2a) en este dioide completo está dada por:

$$x(k) = A_0^*(A_1x(k-1) \oplus \dots \oplus B_0u(k) \oplus B_1u(k-1) \oplus \dots).$$

Este “simple” argumento para librarse de la parte implícita debe ser examinado más cuidadosamente desde un punto de vista práctico.

1. Teniendo en cuenta que hemos estado confinados en $\mathbb{R}_{\text{máx}}$, algunos elementos de $x(k)$ podrían resultar iguales a $\top = +\infty$. ¿Qué implicará esto en la práctica ?
2. Hemos tomado la *menor* solución: si ésta no es única, ¿cuál es el sentido de esta elección ?

La respuesta a la última pregunta está en relación con las dos “reglas del juego” que repetimos a continuación:

1. las transiciones se disparan *tan rápido como sea posible*, lo que compatibiliza el menor posible “dater” con las ecuaciones;
2. las ecuaciones implícitas son válidas “fuera de la influencia del marcado inicial”; se seleccionaron las condiciones iniciales de forma tal que cualquier otra selección puede solamente retardar todos los eventos posteriores.

Respecto a la pregunta anterior, el significado de que algún $x_i(k)$ sea infinito implicaría que el k -ésimo disparo de la transición x_i ocurra en tiempo $+\infty$, o sea, no ocurra nunca. Esto podría sólo ocurrir si algunos elementos de A_0^* son iguales a $+\infty$. Teniendo en cuenta la definición (2.4) de la estrella de Kleene y a la interpretación de las potencias sucesivas de una matriz en términos de los pesos de los caminos de su grafo de precedencia, para que esto último ocurra, resulta necesario que el grafo de precedencia de A_0 contenga circuitos con pesos *positivos*.

Teniendo en cuenta que este grafo es el que se obtiene manteniendo solamente transiciones internas y lugares (arcos) entre las mismas sin marcas iniciales, la presencia de circuitos con peso positivo en dicho grafo implica que las transiciones presentes en tales circuitos no puedan ser disparadas nunca (es un problema de *vivacidad*).

Luego, es razonable asumir que tales situaciones no tendrán lugar. Si A_0 no contiene circuitos, podrá ser llevada a una forma *estrictamente triangular inferior* con una apropiada numeración de las transiciones internas, y luego, la ecuación (3.2a) no será *verdaderamente* implícita. Este es el caso obvio cuando la serie de potencias que define la estrella de Kleene tiene un número finito de términos no nulos.

En el análisis hecho, la presencia de circuitos sin marcas en el mercado inicial; pero con peso nulo (i.e. tiempo de espera de los lugares igual a 0), es algo controversial. Desde el punto de vista de los especialistas en redes de Petri esta situación no debe ser considerada como “viva”. Sin embargo, desde el punto de vista algebraico, si bien se pierde unicidad en la solución de (3.2a), la menor solución (que es la que interesa) sigue siendo dada por dicha fórmula.

Desde el punto de vista práctico, la “vivacidad” de un circuito sin marcas y con tiempo de espera nulo puede justificarse suponiendo que se permite “tomar marcas prestadas del exterior durante un tiempo cero”. Luego, cada transición involucrada en dichos ciclos puede dispararse y luego devolver la marca prestada instantáneamente.

3.1.5.2. Ecuaciones de estados

La etapa posterior a la obtención de la forma canónica es transformar las ecuaciones para limitar los retardos a retardos unitarios sólo en la parte interna del sistema. De esta forma, se intenta llegar a las siguientes ecuaciones *canónicas*:

$$X(k) = AX(k-1) \oplus Bu(k) , \quad (3.3a)$$

$$y(k) = CX(k) \oplus Du(k) , \quad (3.3b)$$

con nuevas definiciones de X , A , B , C , D . Esta es la manipulación estándar en la teoría de sistemas, que se lleva a cabo a través de la incorporación al nuevo vector de estados X de suficientes versiones retardadas de las anteriores variables x_i y u_j de forma tal que $X(k-1)$ contenga toda la información necesaria para calcular $X(k)$. Con esta nueva representación, el nuevo GET correspondiente tendrá exactamente una marca en el mercado inicial en todos los lugares ubicados entre dos transiciones internas y no tendrá marcas en los lugares ubicados entre transiciones de entrada e internas y entre transiciones internas y de salida. La manipulación necesaria para llegar a las ecuaciones canónicas será explicada durante el desarrollo del ejemplo que sigue.

3.1.6.3. Eliminación de las variables de estados “espúreas”

Analizando el GET correspondiente a las nuevas ecuaciones (Figura 3.3) podemos observar que ahora x_1 es una transición de salida en lugar de una transición interna. Matemáticamente, esto se traduce en el hecho que x_1 no aparece más en

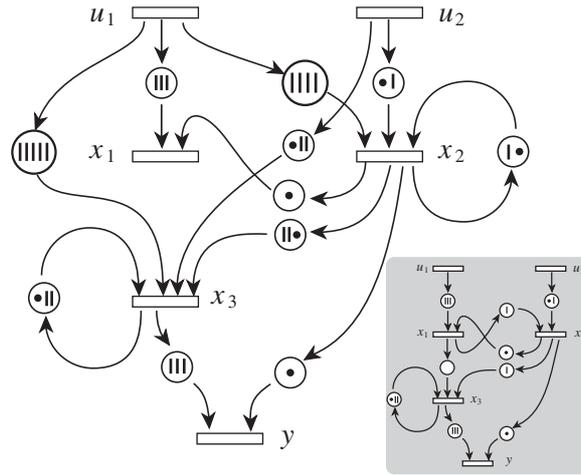


Figura 3.3: GET tras eliminar la parte implícita

el lado derecho de las ecuaciones (en todas las matrices A_i y C_i , la primer columna es cero). Luego, será posible eliminar x_1 del vector de estados. Con esto, las ecuaciones quedan:

$$\begin{pmatrix} x_2 \\ x_3 \end{pmatrix} (k) = \begin{pmatrix} 1 & \varepsilon \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} (k-1) \oplus \begin{pmatrix} 4 & \varepsilon \\ 5 & \varepsilon \end{pmatrix} u(k) \oplus \begin{pmatrix} \varepsilon & 1 \\ \varepsilon & 2 \end{pmatrix} u(k-1),$$

$$y(k) = (\varepsilon \quad 3) \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} (k) \oplus (e \quad \varepsilon) \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} (k-1).$$

El GET correspondiente se muestra en Figura 3.4.

3.1.6.4. Confinación de los retardos a la parte interna

Las últimas ecuaciones obtenidas no están aún en la forma canónica (3.3), porque $u_2(k-1)$ todavía aparece en la ecuación de $x(k)$, y $x_2(k-1)$ aparece en la ecuación de $y(k)$. Para llegar a la forma canónica será necesario agregar dos nuevas variables de estado: $x_1(k) \stackrel{\text{def}}{=} u_2(k)$ (esta nueva variable llamada x_1 no tiene relación con la anterior x_1 que fue eliminada del vector de estados viejo) y

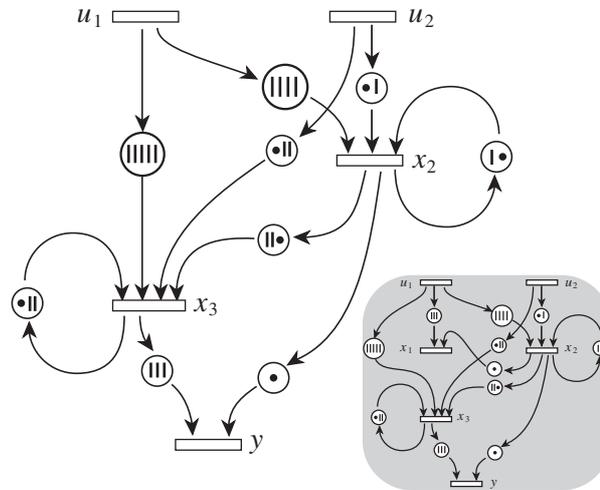


Figura 3.4: GET tras la eliminación de las variables de estado espúreas

$x_4(k) \stackrel{\text{def}}{=} x_2(k-1)$. Finalmente, tras este último paso, se obtiene la forma canónica:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} (k) = \begin{pmatrix} \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ 1 & 1 & \varepsilon & \varepsilon \\ 2 & 2 & 2 & \varepsilon \\ \varepsilon & e & \varepsilon & \varepsilon \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} (k-1) \oplus \begin{pmatrix} \varepsilon & e \\ 4 & \varepsilon \\ 5 & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix} u(k),$$

$$y(k) = (\varepsilon \ \varepsilon \ 3 \ e) (x_1 \ x_2 \ x_3 \ x_4)^\top (k).$$

El GET correspondiente se muestra en Figura 3.5.

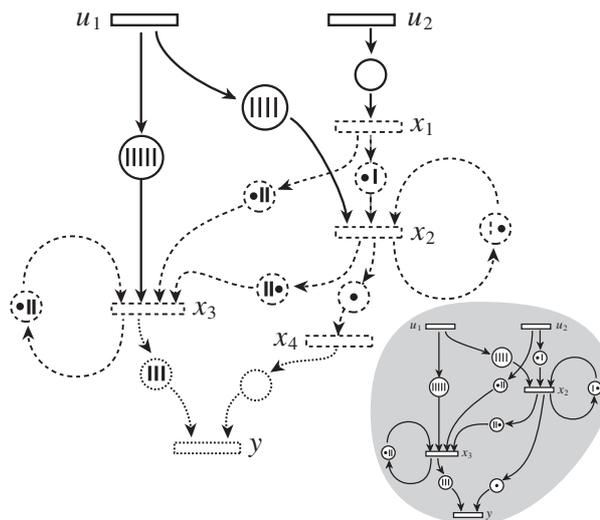


Figura 3.5: GET tras eliminar retardos en arcos de entrada y salida

Ejercicio 3.1. Considerar los esquemas (a) y (b) de la Figura 1.14, determinar las ecuaciones explícitas de dichos sistemas y mostrar que ambos conjuntos de ecuaciones son equivalentes al del esquema (c) en la misma figura.

3.2. Matrices transferencia via la transformada γ

3.2.1. La transformada γ

Por el uso repetido de (3.3a), se tiene (con $x(-1) = \varepsilon$ de acuerdo a las condiciones iniciales canónicas)

$$x(k) = \bigoplus_{i=0}^k A^i B u(k-i)$$

(con $A^0 = e$, matriz identidad) y luego, con (3.3b),

$$y(k) = \bigoplus_{i=0}^k C A^i B u(k-i) \oplus D u(k) . \quad (3.4)$$

Lo anterior al último término $Du(k)$, es la “convolución” de la secuencia de entradas $\{u(\cdot)\}$ con la secuencia de matrices

$$h(i) \stackrel{\text{def}}{=} C A^i B , \quad i \in \mathbb{N} , \quad (3.5)$$

secuencia que en la teoría clásica de sistemas es llamada *resupuesta al impulso* del sistema (los coeficientes $C A^i B$ son llamados *coeficientes de Markov*). Dado que todas las operaciones son en álgebra máx-plus, se trata más bien de una *sup-convolución*: suponiendo por simplicidad que u e y son vectores de dimensión 1, con la notación usual, la ecuación (3.4) se escribirá:

$$y(k) = \max \left(\max_{i=0, \dots, k} (h(i) + u(k-i)), D + u(k) \right) .$$

Una tarea básica en la teoría de sistemas es componer sistemas para obtener sistemas complejos a partir de bloques más simples de construir. Las principales operaciones son composiciones serie, paralelo y feedback como se muestra en la Figura 3.6. Realizar estas operaciones con la representación entrada-salida de (3.4) no es una tarea fácil. Este es el motivo por el cual la “transformada z ” ha sido introducida en la teoría clásica de sistemas. La transformada γ presentada ahora es el equivalente para sistemas máx-plus lineales (γ es más bien análoga a z^{-1}).

Para una “señal” $\{u(k)\}_{k \in \mathbb{Z}}$, la transformada γ , denominada $U(\gamma)$, es la *serie formal de potencias* (en realidad, una serie de Laurent con $k \in \mathbb{Z}$ en lugar de \mathbb{N}) en la variable formal γ con coeficientes $u(\cdot)$, definida como:

$$U(\gamma) = \bigoplus_{k \in \mathbb{Z}} u(k) \gamma^k . \quad (3.6)$$

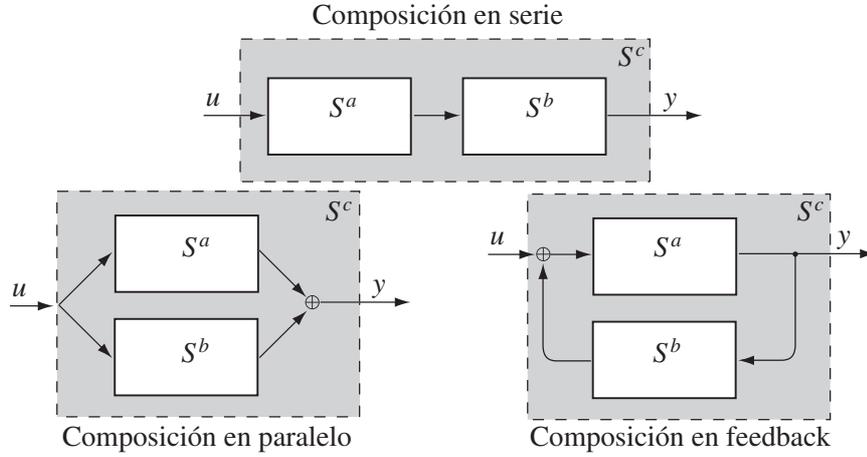


Figura 3.6: Composición de sistemas

Obsérvese que

$$\gamma U(\gamma) = \bigoplus_{k \in \mathbb{Z}} u(k) \gamma^{k+1} = \bigoplus_{k \in \mathbb{Z}} u(k-1) \gamma^k,$$

que es la transformada γ de la señal $\{v(k)\}_{k \in \mathbb{Z}}$ con $v(k) = u(k-1)$. Esta observación se traduce en la relación simbólica:

$$\gamma u(k) = u(k-1),$$

y γ es denominado operador “backward shift”.

3.2.2. Obtención de la matriz transferencia de un sistema

Volviendo a las ecuaciones canónicas (3.3) y considerando las transformadas γ de las señales $u(\cdot)$, $x(\cdot)$, $y(\cdot)$, de la ecuación (3.3a) se obtiene

$$X(\gamma) = \gamma A X(\gamma) \oplus B U(\gamma).$$

Esta es una ecuación implícita de $X(\gamma)$ en el dióide de matrices de series de potencia en γ con coeficientes en álgebra máx-plus. Por varias razones mencionadas anteriormente, es conveniente considerar álgebras máx-plus *completas* tales como $\overline{\mathbb{Z}}_{\text{máx}}$ o $\overline{\mathbb{R}}_{\text{máx}}$.

Debido a los argumentos discutidos en §3.1.5.1, tiene sentido tomar la *menor solución* de la ecuación anterior, lo que lleva a

$$X(\gamma) = (\gamma A)^* B U(\gamma).$$

Usando (3.3b), se obtiene

$$Y(\gamma) = (C(\gamma A)^* B \oplus D) U(\gamma). \quad (3.7)$$

Esta es la relación que expresa el comportamiento *entrada-salida* del sistema dando la transformada γ de la salida como función de la transformada γ de la entrada.

El “operador”

$$H(\gamma) = C(\gamma A)^* B \oplus D \quad (3.8)$$

se denomina *matriz transferencia* del sistema. Se trata de una matriz cuyo número de filas es igual a la dimensión del vector de salida y , el número de columnas es igual a la dimensión del vector de entrada u , y sus elementos son series formales de potencia en γ con coeficientes en $\overline{\mathbb{Z}}_{\text{máx}}$ o $\overline{\mathbb{R}}_{\text{máx}}$.

Obsérvese que $C(\gamma A)^* B$ en (3.8) es la transformada γ de la respuesta al impulso definida en (3.5). Obsérvese también que la (sup-)convolución (3.4) de la respuesta al impulso $h(\cdot)$ por la señal de entrada $u(\cdot)$ fue transformada en un *producto* de las series de potencias $C(\gamma A)^* B$ y $U(\gamma)$ (con coeficientes matriciales).¹

En un GET puede calcularse la matriz transferencia desde cualquier transición o subconjunto de transiciones \mathcal{J} (aunque no sean transiciones de entrada) hacia cualquier transición o subconjunto de transiciones \mathcal{O} (aunque no sean transiciones de salida) mediante el siguiente procedimiento:

- si una transición T_i en \mathcal{J} no es una transición de entrada, agregar un nuevo lugar anterior a dicha transición (sin marcas y con tiempo de espera nulo) y una nueva transición anterior a dicho lugar; considerar que esta nueva transición de entrada sustituyendo a T_i en \mathcal{J} ;
- si una transición T_j en \mathcal{O} no es una transición de salida, agregar un nuevo lugar (sin marcas y con tiempo de espera nulo) posterior a dicha transición y una nueva transición de salida posterior a este lugar; sustituir T_j en \mathcal{O} por esta nueva transición de salida;
- calcular la matriz transferencia del nuevo sistema completo con las entradas y salidas adicionales;
- extraer desde esta matriz transferencia “grande” la matriz transferencia con índices de columnas en \mathcal{J} e índices de filas en \mathcal{O} ; obsérvese que esto equivale a ignorar las salidas que no estén en \mathcal{O} y suponer todas las entradas que no estén en \mathcal{J} iguales a ε .

El valor ε es el *valor de entrada menos restrictivo* que pueda imaginarse, ya que corresponde a disparar la transición de entrada un número infinito de veces en tiempo $-\infty$, provocando una infinita cantidad de marcas listas para usar en los lugares posteriores al tiempo $-\infty$. Recordar que en general, una secuencia de disparos que ocurren en tiempos $d(0), d(1), \dots$, es representada por su transformada $\gamma \bigoplus_{k \in \mathbb{N}} d(k) \gamma^k$. Aquí, para todo k , $d(k) = -\infty = \varepsilon$.

¹Esto puede relacionarse con el hecho que la aplicación de la *transformada de Laplace* a una convolución la convierte en un producto, lo que constituye uno de los motivos por los cuales se utilizan matrices transferencia para los sistemas de tiempo continuo lineales y estacionarios en la teoría clásica de sistemas

3.2.3. Matriz transferencia y respuesta al impulso

Como ya se dijo, la matriz transferencia es simplemente la transformada γ de la respuesta al impulso. Para simplificar la notación, consideraremos un sistema SISO (Single Input Single Output). Esto implica considerar la función transferencia desde una entrada particular i hacia una salida particular j .

Supongamos que utilizamos una trayectoria de entrada representada por $U(\gamma) = e$. Luego, la salida $Y(\gamma) = H(\gamma)e = H(\gamma)$. Por consiguiente, $H(\gamma)$ es la transformada γ de la trayectoria de salida correspondiente a esa trayectoria de entrada particular: así, esta trayectoria de salida deberá coincidir con la respuesta al impulso.

Para deducir cual es la trayectoria de entrada representada por e , obsérvese que:

$$e = 0\gamma^0 \oplus \varepsilon\gamma \oplus \varepsilon\gamma^2 \oplus \dots$$

donde aparece un primer inconveniente: la secuencia de coeficientes de los monomios de potencias crecientes de γ no es *no decreciente*, como debiera ser para los “daters”. Supongamos, y luego volveremos sobre este tema, que se puede decir que $e \sim \gamma^*$ en donde todos los coeficientes son iguales a 0, esto es, un número infinito de disparos en tiempo 0 en la entrada. Esto es un “impulso” y entonces, se justifica llamar a la correspondiente trayectoria de salida como “respuesta al impulso”.

3.2.4. Resolución de un ejemplo (continuación)

Volvemos al ejemplo discutido en §3.1.6 correspondiente a las Figuras 3.2 a 3.5. Como se dijo, para calcular la matriz transferencia de este sistema se puede comenzar desde cualquier conjunto de ecuaciones de los que se obtuvieron en dicha sección. Comenzaremos desde la forma inicial correspondiente a la Figura 3.2.

Como también se dijo, no es práctico en general calcular expresiones tales como $(\gamma A)^*$ mediante la aplicación de la definición de la estrella de Kleene. Es más fácil resolver ecuaciones *escalares* implícitas de manera progresiva (como en la eliminación de Gauss) que resolver ecuaciones vectoriales implícitas. Esta idea sería explicada a través del ejemplo.

Para abreviar, la transformada γ será notada utilizando mayúsculas; pero sin mención explícita del argumento γ (por ejemplo, U en lugar de $U(\gamma)$). Luego, partiendo de las ecuaciones iniciales, obtenemos sucesivamente:

$$\begin{aligned} x_1(k) &= x_2(k-1) \oplus 3u_1(k) \Rightarrow X_1 = \gamma X_2 \oplus 3U_1, \\ x_2(k) &= 1x_1(k) \oplus 1u_2(k-1) \Rightarrow X_2 = 1X_1 \oplus 1\gamma U_2, \\ x_3(k) &= x_1(k) \oplus 1x_2(k) \oplus 2x_3(k-1) \Rightarrow X_3 = X_1 \oplus 1X_2 \oplus 2\gamma X_3, \\ y(k) &= x_2(k-1) \oplus 3x_3(k) \Rightarrow Y = \gamma X_2 \oplus 3X_3. \end{aligned}$$

Luego, substituyendo la expresión de X_1 en la ecuación de X_2 ,

$$\begin{aligned} X_2 &= 1\gamma X_2 \oplus 4U_1 \oplus 1\gamma U_2 \Rightarrow X_2 = (1\gamma)^*(4U_1 \oplus 1\gamma U_2) \\ &\Rightarrow X_1 = (3 \oplus 4\gamma(1\gamma)^*)U_1 \oplus 1\gamma^2(1\gamma)^*U_2. \end{aligned}$$

Estas expresiones de X_1 y X_2 utilizadas en la ecuación de X_3 llevan a:

$$\begin{aligned} X_3 &= (3 \oplus 4\gamma(1\gamma)^*)U_1 \oplus 1(1\gamma)^*(4U_1 \oplus 1\gamma U_2) \oplus 2\gamma X_3 \\ &= (3 \oplus (1\gamma)^*(4\gamma \oplus 5))U_1 \oplus 2\gamma(1\gamma)^*U_2 \oplus 2\gamma X_3 \\ &= (1\gamma)^*(5U_1 \oplus 2\gamma U_2) \oplus 2\gamma X_3, \end{aligned}$$

ya que se puede verificar directamente que $4\gamma(1\gamma)^* \leq 5(1\gamma)^*$. Más aún, más tarde veremos por qué se puede decir que $4\gamma \leq 5$, lo que en este momento parece extraño ya que en general no se pueden comparar monomios en γ de diferentes exponentes.

En consecuencia,

$$\begin{aligned} X_3 &= (2\gamma)^*(1\gamma)^*(5U_1 \oplus 2\gamma U_2) \\ &= (2\gamma)^*(5U_1 \oplus 2\gamma U_2). \end{aligned}$$

La última igualdad es una consecuencia del siguiente lema y del hecho que $1 \oplus 2 = 2$.

Lema 3.2. *En un dioide conmutativo (o sea, un dioide con multiplicación conmutativa), se tiene*

$$a^*b^* = (a \oplus b)^*.$$

La demostración se deja como ejercicio.

Siguiendo con el ejemplo, hasta acá hemos obtenido X_1, X_2, X_3 en forma explícita como expresiones de U_1, U_2 . Aún resta expresar Y en la misma manera explícita para obtener la matriz transferencia.

$$\begin{aligned} Y &= \gamma(1\gamma)^*(4U_1 \oplus 1\gamma U_2) \oplus 3(2\gamma)^*(5U_1 \oplus 2\gamma U_2) \\ &= (4\gamma(1\gamma)^* \oplus 8(2\gamma)^*)U_1 \oplus (1\gamma^2(1\gamma)^* \oplus 5\gamma(2\gamma)^*)U_2 \\ &= 8(2\gamma)^*U_1 \oplus 5\gamma(2\gamma)^*U_2. \end{aligned}$$

La última igualdad puede ser justificada por la expansión de las estrellas de Kleene involucradas y la utilización de las reglas de cálculo en álgebra máx-plus para los coeficientes de los monomios en γ de la misma potencia. Sin embargo, una forma mejor es observar que

- $(2\gamma)^* \succeq (1\gamma)^*$;
- $4\gamma \leq 8$ y $1\gamma^2 \leq 5\gamma$, que aún no parece claro debido a la comparación de monomios de diferente potencia.

Finalmente, factorizando $5(2\gamma)^*$ en la última expresión de Y (recordando que $8 = 3 \otimes 5$), se obtiene la matriz transferencia:

$$Y = 5(2\gamma)^*(3U_1 \oplus U_2) = (8(2\gamma)^* \quad 5(2\gamma)^*) \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}.$$

3.2.5. Simplificación del sistema por el cálculo de la matriz transferencia

Se puede ver ahora que, desde un punto de vista entrada-salida, el GET de la Figura 3.2 tiene el mismo comportamiento que el GET de la Figura 3.7 ya que la matriz transferencia es la misma. Una importante consecuencia de la manipulación algebraica en $\overline{\mathbb{Z}}_{\text{máx}}[\![\gamma]\!] o \overline{\mathbb{R}}_{\text{máx}}[\![\gamma]\!]$ es el poder de simplificación de sistemas en un modo automático (computacional).

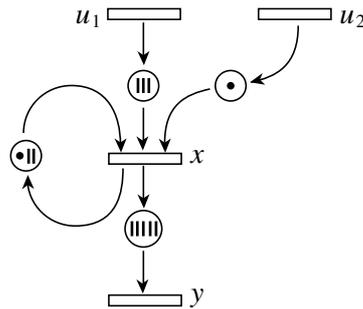


Figura 3.7: Un GET equivalente al de la Figura 3.2

Ejercicio 3.3. Demostrar que los dos GET de la Figura 3.8 tienen la misma función transferencia.

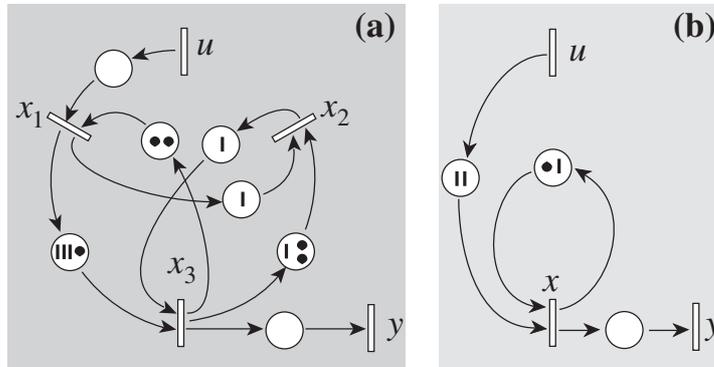


Figura 3.8: Dos GET equivalentes

3.2.6. Imposición de condiciones iniciales no canónicas

Volviendo al tema de las condiciones iniciales canónicas definidas en §3.1.3, por medio de un ejemplo simple (que puede ser fácilmente adaptado para considerar situaciones generales), se verá como tener en cuenta condiciones iniciales no canónicas. Supongamos que en la Figura 3.9-(a) queremos que las dos marcas iniciales del lugar P estén listas para ser utilizadas en los instantes de tiempo 1

y 3 respectivamente para el disparo de la transición x_2 (pensar cuáles serían sus tiempos de llegada a P). Una forma de conseguir esto manteniendo la regla de condiciones iniciales canónicas (esto es, todas las marcas del marcado inicial están listas para usar al comienzo de la simulación) es la siguiente:

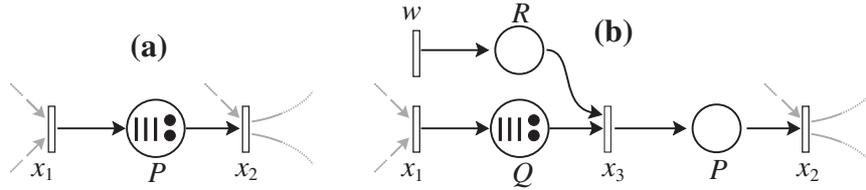


Figura 3.9: Control de las condiciones iniciales

- agregar una nueva transición x_3 anterior a P y un nuevo lugar Q anterior a x_3 como se ve en la Figura 3.9-(b); además, x_3 recibe una nueva transición de entrada w a través de un nuevo lugar R ;
- marcar Q con el marcado anterior de P y transferir el tiempo de espera de P a Q (ahora P no tiene más marcas iniciales y tiene tiempo de espera nulo);
- considerar la siguiente trayectoria para la nueva transición de entrada w : 1, 3, 3, \dots , o sea, la primer activación (número 0) se hace en tiempo 1, la segunda en tiempo 3, y todas las demás también en tiempo 3 (de forma tal que haya un número infinito de marcas disponibles en tiempo 3 en el lugar R).

Puede verse fácilmente que las dos primeras marcas que lleguen al lugar P estarán listas para usar en tiempos 1 y 3 como se quería. Las siguientes marcas que llegarán al lugar P (producidas por los disparos de x_1 tras el comienzo de la simulación) estarán listas para usar en P al mismo tiempo que estén listas para usar en Q , y al mismo tiempo que están en P en la parte (a) de la figura.

La transformada γ de esta trayectoria de w es

$$W(\gamma) = 1\gamma^0 \oplus 3\gamma \oplus 3\gamma^2 \oplus \dots = 1 \oplus 3\gamma^+$$

con la notación tradicional

$$a^+ = a \otimes a^* = a \oplus a^2 \oplus \dots \quad \text{donde} \quad a^* = e \oplus a^+ . \quad (3.9)$$

Sea $H(\gamma)$ la matriz transferencia global (correspondiente a las condiciones iniciales canónicas) del sistema del que la Figura 3.9 muestra sólo una parte, y sea $H_w(\gamma)$ la matriz transferencia desde la transición de entrada w hasta el vector de salida completo. Luego, la salida correspondiente a las condiciones iniciales deseadas está dada por:

$$Y(\gamma) = H(\gamma)U(\gamma) \oplus H_w(\gamma)(1 \oplus 3\gamma^+) .$$

Nótese que ahora la relación entrada salida no es más lineal, sino *afín*, debido a la presencia de condiciones iniciales no nulas (equivalente a una entrada adicional w con los valores dados).²

3.2.7. Una controversia sobre las condiciones iniciales

Hemos considerado que las marcas iniciales (con condiciones iniciales canónicas) están disponibles desde el tiempo $-\infty$, lo cual puede interpretarse como una aparente paradoja.

Considérese por ejemplo la situación de la forma canónica de las ecuaciones obtenidas en §3.1.5: es evidente que corresponden a un GET con exactamente una marca en los lugares ubicados entre transiciones internas, y de acuerdo a las convenciones hechas, esas marcas están allí desde $-\infty$. Si consideramos ahora una transición interna con todas las transiciones anteriores siendo también internas, esta transición puede ser disparada desde $-\infty$ (ya que hay una marca en cada lugar anterior lista para ser utilizada). Entonces, ¿por qué esta transición no se disparó antes consumiendo las marcas que luego no deberían estar allí en el marcado inicial en el tiempo 0 ?

La primer observación la haremos sobre una situación más simple. El siguiente GET es quizás el más simple que pueda imaginarse. Esta sería la representación

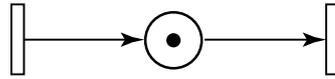


Figura 3.10: Shift γ

en redes de Petri del shift γ (la función transferencia entrada-salida es γ). Nuevamente, la marca inicial debería haberse ido por el disparo de la transición de salida. Dada una trayectoria de entrada cuya transformada es $U(\gamma)$, la correspondiente trayectoria de salida es $\gamma U(\gamma)$, o sea, la k -ésima marca que entra a través de la transición de entrada (en tiempo $u(k)$), lo hace al mismo tiempo que la marca número $k + 1$ sale por la transición de salida. De aquí que el rol de este GET es simplemente desplazar la numeración en 1 entre la entrada y la salida debido a la presencia de una marca en el marcado inicial.

Respecto a la primer marca (número 0) puede decirse que salió en tiempo $-\infty = \varepsilon$ dado que en la transformada γ de la trayectoria de salida, esto es $\gamma U(\gamma)$, el coeficiente que acompaña a γ^0 es “cero” (no existe el monomio), esto es, ε . Tenemos entonces la confirmación de que la marca del marcado inicial se fue en $-\infty$ y luego, debido a este evento, los siguientes eventos (disparos de la transición de salida debidos a marcas que realmente entran al sistema por la transición de salida) son numerados en una escala desplazada en uno con respecto a los disparos de la entrada.

²En la teoría de sistemas clásica, se suele utilizar un procedimiento análogo para introducir condiciones iniciales no nulas a través de la incorporación de entradas adicionales. En este caso las funciones tales como la H_w se suelen denominar pseudo-transferencias.

Como vemos, *desde un punto de vista algebraico*, todo parece ser consistente. La paradoja permanece si se insiste en pensar que el mercado inicial es “una foto del GET tomada en tiempo 0” (el tiempo inicial de la simulación).

Capítulo 4

Teoría de residuación

La mayor parte del material de este capítulo pertenece a la teoría de lattices. Como hemos visto en el capítulo 2, un dioide completo es, en particular, un lattice completo debido a la estructura de orden inducida por la suma. La estructura adicional que provee la multiplicación prácticamente no será utilizada, exceptuando por supuesto cuando consideremos funciones en las que interviene la multiplicación (por ejemplo, cuando consideremos la residual de la función $x \mapsto a \otimes x$).

4.1. Funciones monótonas y continuidad

4.1.1. Funciones monótonas

En esta sección caracterizaremos a las funciones monótonas (ver Definición 2.12) en términos de conjuntos inferiores y superiores.

Definición 4.1. Sea D un conjunto ordenado. Un *conjunto inferior* es un subconjunto no vacío I de D que verifica la siguiente propiedad:

$$\{b \leq a \text{ y } a \in I\} \Rightarrow b \in I .$$

Un *conjunto inferior cerrado* (generado por a) es un conjunto inferior de la forma $\{b \mid b \leq a\}$ el cual será representado por $[\leftarrow, a]$. Un *conjunto superior* es un subconjunto no vacío S de D que verifica la siguiente propiedad:

$$\{a \leq b \text{ y } a \in S\} \Rightarrow b \in S .$$

Un *conjunto superior cerrado* (generado por a) es un conjunto superior de la forma $\{b : a \leq b\}$ el cual será representado por $[a, \rightarrow]$.

Notemos que un conjunto inferior cerrado es un conjunto inferior que contiene a su (menor) cota superior. Análogamente, un conjunto superior cerrado es un conjunto superior que contiene a su (mayor) cota inferior. En $\mathbb{R}_{\text{máx}}$ los conjuntos inferiores cerrados son intervalos de la forma $(-\infty, a]$, mientras que los conjuntos superiores cerrados son intervalos de la forma $[a, +\infty)$ o bien $(-\infty, a)$. Análogamente, en $\mathbb{R}_{\text{máx}}$, los conjuntos superiores cerrados son intervalos de la forma $[a, +\infty)$. En la Figura 4.1 se representan estos conjuntos en el caso de un lattice parcialmente ordenado. En dicha figura se representa a un lattice utilizando un *diagrama de Hasse*,

en donde los nodos representan elementos, los arcos entre elementos indican si los mismos son comparables por la relación de orden (se utiliza la propiedad transitiva para representar un número mínimo de arcos) y el elemento que se encuentra arriba de otro (conectado por un camino) es el mayor.

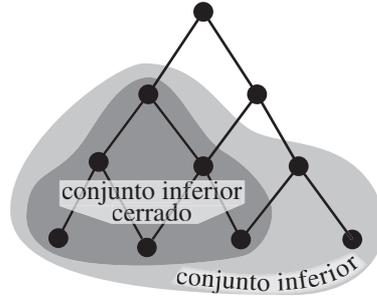


Figura 4.1: Conjunto inferior y conjunto inferior cerrado

Si Π es un \vee -morfismo (\oplus -morfismo en el caso de dioides) o un \wedge -morfismo,¹ entonces Π es una función monótona (ver §2.2.1). En particular, si D es un dioide, como para cada $a \in D$ la función $x \mapsto a \otimes x$ es un \oplus -morfismo, resulta que la misma es una función monótona. Sin embargo, como fue comentado en §2.2.1, una función monótona no es necesariamente un \oplus -morfismo ni un \wedge -morfismo (ver Ejemplo 2.24).

Lema 4.2. *Sea Π una función del lattice D en el lattice S . Entonces las siguientes afirmaciones son equivalentes:*

1. Π es monótona;
2. la pre-imagen $\Pi^{-1}([\leftarrow, a])$ de todo conjunto inferior cerrado es un conjunto inferior o es el conjunto vacío;
3. la pre-imagen $\Pi^{-1}([a, \rightarrow])$ de todo conjunto superior cerrado es un conjunto superior o es el conjunto vacío;
4. Π es un \vee -supermorfismo, es decir,

$$\Pi(x \vee y) \succeq \Pi(x) \vee \Pi(y), \quad \forall x, y \in D;$$

5. Π es un \wedge -submorfismo, es decir,

$$\Pi(x \wedge y) \preceq \Pi(x) \wedge \Pi(y), \quad \forall x, y \in D.$$

¹Una función Π de D en S (lattices) es \wedge -morfismo si $\Pi(a \wedge b) = \Pi(a) \wedge \Pi(b)$, $\forall a, b \in D$. Un \vee -morfismo se define en forma análoga.

Demostración.

1 \Rightarrow 2 Supongamos que Π es monótona. Sea $x \in \Pi^{-1}([\leftarrow, a])$ si este conjunto no es vacío. Entonces, por definición de pre-imagen, $\Pi(x) \preceq a$. Sea $y \preceq x$. Entonces $\Pi(y) \preceq \Pi(x) \preceq a$, y por lo tanto $y \in \Pi^{-1}([\leftarrow, a])$. Con esto queda demostrado que $\Pi^{-1}([\leftarrow, a])$ es un conjunto inferior.

2 \Rightarrow 1 Supongamos que $y \preceq x$. Como $x \in \Pi^{-1}([\leftarrow, \Pi(x)])$ y además como por hipótesis $\Pi^{-1}([\leftarrow, \Pi(x)])$ es un conjunto inferior, resulta que $y \in \Pi^{-1}([\leftarrow, \Pi(x)])$. Entonces $\Pi(y) \preceq \Pi(x)$ y por lo tanto Π es una función monótona.

1 \Leftrightarrow 3 Es análoga a $1 \Leftrightarrow 2$.

1 \Rightarrow 4 Supongamos que Π es monótona. Entonces

$$\left. \begin{array}{l} x \vee y \geq x \Rightarrow \Pi(x \vee y) \geq \Pi(x) \\ x \vee y \geq y \Rightarrow \Pi(x \vee y) \geq \Pi(y) \end{array} \right\} \Rightarrow \Pi(x \vee y) \geq \Pi(x) \vee \Pi(y)$$

donde la última implicación se debe a que $\Pi(x) \vee \Pi(y)$ es la menor de las cotas superiores de $\Pi(x)$ y $\Pi(y)$. Por lo tanto, Π es un \vee -supermorfismo.

4 \Rightarrow 1 Supongamos que $y \preceq x$. Entonces $x = x \vee y$. Por hipótesis tenemos que $\Pi(x) = \Pi(x \vee y) \geq \Pi(x) \vee \Pi(y)$ y entonces $\Pi(y) \preceq \Pi(x)$. Por lo tanto, Π es una función monótona.

1 \Leftrightarrow 5 Es análoga a $1 \Leftrightarrow 4$. □

Cuando D es una cadena las afirmaciones 4 y 5 del lema anterior se verifican con igualdad. Por ejemplo, si D es una cadena, entonces $x \vee y$ es igual a y o x , y por lo tanto $\Pi(x \vee y)$ es igual a $\Pi(y)$ o $\Pi(x)$, con lo cual $\Pi(x \vee y)$ es igual a $\Pi(y) \vee \Pi(x)$. Entonces en este caso Π es un \vee -morfismo. Lo mismo se verifica para la afirmación 5 y en ese caso Π es un \wedge -morfismo.

Si D y S son lattices completos, entonces es fácil verificar que las afirmaciones 4 y 5 también se verifican cuando \vee y \wedge operan sobre subconjuntos infinitos.

4.1.2. Continuidad

Definición 4.3. Una función Π de un lattice completo D en un lattice completo S se dice *semicontinua inferior* (s.c.i.) si para cada subconjunto (finito o infinito) R de D se verifica:

$$\Pi\left(\bigvee_{x \in R} x\right) = \bigvee_{x \in R} \Pi(x).$$

Análogamente, Π se dice *semicontinua superior* (s.c.s.) si se verifica:

$$\Pi\left(\bigwedge_{x \in R} x\right) = \bigwedge_{x \in R} \Pi(x).$$

Una función se dice *continua* cuando es semicontinua inferior y superior.

Una función semicontinua inferior (respectivamente semicontinua superior) es en particular un \vee -morfismo (respectivamente un \wedge -morfismo).

Observación 4.4. Notemos que obtenemos la noción habitual de semicontinuidad inferior (respectivamente superior) cuando aplicamos la definición anterior de semicontinuidad inferior (respectivamente superior) a una función monótona de \mathbb{R} en \mathbb{R} . De aquí la terminología utilizada.

Lema 4.5. *Las siguientes afirmaciones son equivalentes:*

- Π es semicontinua inferior;
- la pre-imagen $\Pi^{-1}([\leftarrow, a])$ de todo conjunto inferior cerrado es un conjunto inferior cerrado o es el conjunto vacío.

Análogamente, las siguientes dos afirmaciones son equivalentes:

- Π es semicontinua superior;
- la pre-imagen $\Pi^{-1}([a, \rightarrow])$ de todo conjunto superior cerrado es un conjunto superior cerrado o es el conjunto vacío.

Demostración. Probaremos únicamente la primera equivalencia. Supongamos que Π es semicontinua inferior. Entonces, por el Lema 4.2, Π es monótona y la pre-imagen $R = \Pi^{-1}([\leftarrow, a])$ de todo conjunto inferior cerrado es un conjunto inferior (si es distinto del vacío). Si $x \in R$ entonces $\Pi(x) \leq a$ y por lo tanto

$$\Pi\left(\bigvee_{x \in R} x\right) = \bigvee_{x \in R} \Pi(x) \leq a.$$

Entonces la cota superior de $R = \Pi^{-1}([\leftarrow, a])$ pertenece a dicho conjunto con lo cual $\Pi^{-1}([\leftarrow, a])$ es un conjunto inferior cerrado.

Supongamos ahora que la pre-imagen de todo conjunto inferior cerrado es un conjunto inferior cerrado. Entonces, por el Lema 4.2, Π es monótona y para todo subconjunto R de D se verifica que:

$$\Pi\left(\bigvee_{x \in R} x\right) \geq \bigvee_{x \in R} \Pi(x). \quad (4.1)$$

Como

$$R \subset \Pi^{-1}\left([\leftarrow, \bigvee_{x \in R} \Pi(x)]\right)$$

y como por hipótesis $\Pi^{-1}\left(\left[\leftarrow, \bigvee_{x \in R} \Pi(x)\right]\right)$ es un conjunto inferior cerrado, resulta que la cota superior $\bigvee_{x \in R} x$ de los elementos de R pertenece a

$$\Pi^{-1}\left(\left[\leftarrow, \bigvee_{x \in R} \Pi(x)\right]\right),$$

y entonces

$$\Pi\left(\bigvee_{x \in R} x\right) \leq \bigvee_{x \in R} \Pi(x). \quad (4.2)$$

De (4.1) y (4.2) obtenemos entonces que

$$\Pi\left(\bigvee_{x \in R} x\right) = \bigvee_{x \in R} \Pi(x),$$

y por lo tanto Π es semicontinua inferior.

La demostración de la segunda equivalencia es análoga. \square

Ejemplo 4.6. Consideremos la función $\Pi : \overline{\mathbb{R}}_{\text{máx}} \rightarrow \overline{\mathbb{Z}}_{\text{máx}}$ definida por:

$$\Pi(x) = \bigoplus_{\{y \in \overline{\mathbb{Z}}_{\text{máx}} \mid y \leq x\}} y,$$

donde el símbolo \leq tiene el significado habitual. Notemos que $\Pi(x)$ es simplemente la parte entera del número real x . Esta función es un \oplus -morfismo y un \wedge -morfismo. También es semicontinua superior (esto es consecuencia, como veremos más adelante, de que Π es la residual de la función $x \mapsto x$ de $\overline{\mathbb{Z}}_{\text{máx}}$ en $\overline{\mathbb{R}}_{\text{máx}}$). Sin embargo Π no es semicontinua inferior, pues tomando por ejemplo $R = \{5 - (\frac{1}{n+1}) \mid n \in \mathbb{N}\}$, resulta:

$$\Pi\left(\bigvee_{x \in R} x\right) = \Pi(5) = 5 \neq 4 = \bigvee_{x \in R} 4 = \bigvee_{x \in R} \Pi(x).$$

Lema 4.7. El conjunto de las funciones semicontinuas inferiores de un dioide completo D en sí mismo es un dioide completo con las operaciones suma $\widehat{\oplus}$ y multiplicación $\widehat{\otimes}$ definidas por:

$$\begin{aligned} \Pi \widehat{\oplus} \Phi &: x \mapsto \Pi(x) \oplus \Phi(x); \\ \Pi \widehat{\otimes} \Phi &: x \mapsto \Pi(\Phi(x)). \end{aligned}$$

Análogamente, el conjunto de las funciones semicontinuas superiores de un dioide completo D en sí mismo es un dioide completo con las operaciones suma $\widetilde{\oplus}$ y multiplicación $\widetilde{\otimes}$ definidas por:

$$\begin{aligned} \Pi \widetilde{\oplus} \Phi &: x \mapsto \Pi(x) \wedge \Phi(x); \\ \Pi \widetilde{\otimes} \Phi &: x \mapsto \Pi(\Phi(x)). \end{aligned}$$

Demostración. Probaremos solamente la primera afirmación pues la demostración de la segunda es análoga. Es sencilla la verificación de que la multiplicación y la suma de dos funciones semicontinuas inferiores es una función semicontinua inferior. El elemento neutro de la suma $\widehat{\varepsilon}$ es claramente la función que es idénticamente igual a ε , mientras que el elemento neutro de la multiplicación \widehat{e} es la función identidad en D . La propiedad distributiva a derecha de la multiplicación con respecto a la suma es consecuencia directa de las definiciones de la suma $\widehat{\oplus}$ y la multiplicación $\widehat{\otimes}$, mientras que la propiedad distributiva a izquierda es consecuencia de estas definiciones y del hecho de que las funciones son semicontinuas inferiores. El hecho de que este nuevo dioide sea completo es consecuencia directa de la definición de la suma $\widehat{\oplus}$ y de que el dioide D es completo. Los restantes axiomas de un dioide pueden comprobarse fácilmente. \square

Observación 4.8. Como por el lema anterior el conjunto de las funciones semicontinuas inferiores de un dioide completo D en sí mismo es un dioide completo, entonces sabemos que está definida la operación cota inferior sobre este conjunto, la cual representaremos con $\widehat{\wedge}$. Sin embargo, en general no se verifica que

$$(\Pi \widehat{\wedge} \Phi)(x) = \Pi(x) \wedge \Phi(x) ,$$

pues la función de la derecha no es necesariamente semicontinua inferior.

Ejemplo 4.9. Sea $D = (\overline{\mathbb{R}}_{\text{máx}})^2$, en donde las operaciones están definidas componente a componente a partir de las operaciones de $\overline{\mathbb{R}}_{\text{máx}}$, y sea $S = (\overline{\mathbb{R}}^2, \widehat{\oplus}, \widehat{\otimes})$, en donde la multiplicación es la misma que en D ; pero la suma está definida por:

$$x \widehat{\oplus} y = \begin{cases} x & \text{si } (x_1 > y_1) \text{ o } (x_1 = y_1 \text{ y } x_2 \geq y_2) ; \\ y & \text{si } (x_1 < y_1) \text{ o } (x_1 = y_1 \text{ y } x_2 \leq y_2) ; \end{cases}$$

es decir, $\widehat{\oplus}$ es la suma inducida por el orden lexicográfico de \mathbb{R}^2 . La relación de orden en D es la relación de orden parcial habitual. Consideremos la función $\Pi : D \rightarrow S$ dada por la biyección canónica $x \mapsto x$. Esta función es monótona pues $x \leq y$ en D implica que $x \preceq y$ en S . Sin embargo, esta función no es semicontinua inferior ni superior. Para ver que Π no es semicontinua inferior, consideremos por ejemplo el conjunto inferior cerrado generado por $(2, 3)$. Este conjunto está representado en la Figura 4.2. Como la pre-imagen por Π de este conjunto coincide con dicho conjunto, y como además los conjuntos inferiores cerrados en D son conos sud-oeste cerrados, resulta por el Lema 4.5 que Π no es semicontinua inferior. Análogamente se puede mostrar que Π no es semicontinua superior.

Este ejemplo es interesante pues muestra que una función biyectiva y monótona no tiene necesariamente una inversa monótona (que $x \preceq y$ en S no implica que $x \leq y$ en D). Sin embargo, si Π es una función monótona de un dioide D con una relación de orden total en un dioide S , entonces Π^{-1} también debe ser monótona. Para probar esta propiedad sean $a, b \in S$ tales que $a \prec b$ ($a \preceq b$, $a \neq b$). Entonces, como D es totalmente ordenado (y Π biyectiva), se debe verificar que $\Pi^{-1}(a) <$

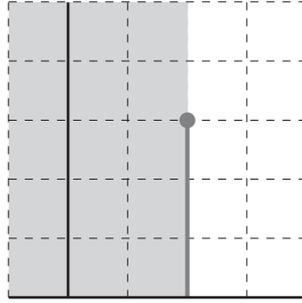


Figura 4.2: Conjunto inferior cerrado para el orden lexicográfico

$\Pi^{-1}(b)$ o bien $\Pi^{-1}(b) < \Pi^{-1}(a)$. Supongamos que $\Pi^{-1}(b) < \Pi^{-1}(a)$. Entonces como Π es monótona se tendría que $\Pi(\Pi^{-1}(b)) < \Pi(\Pi^{-1}(a))$, es decir, $b < a$, lo cual es un absurdo. Por lo tanto $\Pi^{-1}(a) < \Pi^{-1}(b)$, y entonces Π^{-1} es monótona. En este último caso, si D y S son también dioides completos, entonces se puede probar la continuidad tanto de Π como de Π^{-1} .

Lema 4.10. Sean D y S dos dioides completos y Π un homomorfismo de D en S . Consideremos la congruencia definida por

$$x \mathcal{R}_\Pi y \Leftrightarrow \Pi(x) = \Pi(y), \quad \forall x, y \in D.$$

Si Π es semicontinua inferior (respectivamente superior), entonces cada clase de equivalencia posee un máximo elemento (respectivamente mínimo) el cual puede entonces considerarse como un representante canónico de dicha clase.

Demostración. Supongamos que Π es semicontinua inferior. Sea $[x]$ la clase de equivalencia de x y sea $\bar{x} \stackrel{\text{def}}{=} \bigoplus_{y \in [x]} y$ y la (menor) cota superior de los elementos de $[x]$. Debemos mostrar que $\bar{x} \in [x]$. Como Π es semicontinua inferior tenemos que:

$$\Pi(\bar{x}) = \Pi\left(\bigoplus_{y \in [x]} y\right) = \bigoplus_{y \in [x]} \Pi(y) = \bigoplus_{y \in [x]} \Pi(x) = \Pi(x),$$

y por lo tanto $\bar{x} \in [x]$. El caso en el cual Π es semicontinua superior se demuestra en forma similar. \square

4.2. Elementos de la teoría de residuación

4.2.1. Resultados básicos

Los principales objetivos de la teoría de residuación son los de invertir funciones monótonas y el de dar una respuesta al problema de resolución de una ecuación de la forma $\Pi(x) = b$, donde Π es una función monótona entre dos lattices completos. Cuando Π no es sobreyectiva, la ecuación anterior no tiene solución para

ciertos valores de b ; mientras que cuando Π no es inyectiva, la solución puede no ser única. La idea para resolver este problema es la de debilitar la noción de solución a la de *subsolución*, es decir un elemento x que verifica que $\Pi(x) \preceq b$, o bien a la de *supersolución*, es decir un elemento x que verifica que $\Pi(x) \succeq b$. Luego se toma a la menor cota superior (respectivamente la mayor cota inferior) del conjunto de las subsoluciones (respectivamente supersoluciones) como la “solución” de la ecuación. Esto es posible siempre y cuando el conjunto de las subsoluciones (respectivamente supersoluciones) sea no vacío. La elección del enfoque que se utilizará depende de las propiedades de continuidad que verifica Π . El primer enfoque se toma cuando Π es semicontinua inferior ya que en ese caso la menor cota superior del conjunto de las subsoluciones también es una subsolución (es la mayor subsolución). Si en este caso representamos con $\Pi^\sharp(b)$ a la mayor subsolución, entonces se verifica que:

$$\Pi^\sharp(b) = \bigvee_{\{x|\Pi(x)\preceq b\}} x \quad \text{y} \quad \Pi(\Pi^\sharp(b)) \preceq b .$$

Dualmente, se considera el segundo enfoque cuando Π es semicontinua superior ya que en dicho caso la mayor cota inferior del conjunto de las supersoluciones es una supersolución (es la menor supersolución). Si representamos con $\Pi^b(b)$ a esta menor supersolución, entonces se verifica:

$$\Pi^b(b) = \bigwedge_{\{x|\Pi(x)\succeq b\}} x \quad \text{y} \quad \Pi(\Pi^b(b)) \succeq b .$$

Observación 4.11. Notemos que si existe una solución “real” de la ecuación $\Pi(x) = b$ (es decir, que la verifica con igualdad), entonces cualquiera de los dos enfoques (siempre y cuando sean aplicables) dará como respuesta una solución “real”.

Teorema 4.12. *Sea Π una función monótona del lattice completo D en el lattice completo S . Entonces las siguientes afirmaciones son equivalentes:*

1. *Para todo $b \in S$ existe la mayor subsolución de $\Pi(x) = b$ (esto es, el conjunto de las subsoluciones no es vacío y posee un elemento máximo).*
2. *$\Pi(\varepsilon) = \varepsilon$ y Π es semicontinua inferior² (o equivalentemente, la pre-imagen de todo conjunto inferior cerrado es distinto del vacío y es un conjunto inferior cerrado).*
3. *Existe una función Π^\sharp de S en D que es monótona y semicontinua superior tal que:*

$$\Pi \circ \Pi^\sharp \preceq I_S \quad (\text{identidad en } S); \quad (4.3)$$

$$\Pi^\sharp \circ \Pi \succeq I_D \quad (\text{identidad en } D). \quad (4.4)$$

²Representaremos con ε al menor elemento (o elemento base) de cualquier lattice completo.

Cuando Π verifica estas propiedades se dice que es residuable, y entonces la función Π^\sharp (univocamente determinada por (4.3) y (4.4)) recibe el nombre de residual de Π .

Demostración.

1 \Rightarrow 3 Para cada $b \in S$, definamos $\Pi^\sharp(b)$ como la mayor subsolución de $\Pi(x) = b$ (la cual existe por hipótesis). Es fácil verificar que la función Π^\sharp definida de esta manera es monótona. La desigualdad (4.3) es directa a partir de la definición de subsolución. Para probar la desigualdad (4.4), tomemos $b = \Pi(x)$ para cada $x \in D$. Como en particular x es una subsolución correspondiente a dicho b , por la definición de $\Pi^\sharp(b)$ resulta que $x \leq \Pi^\sharp(b) = \Pi^\sharp \circ \Pi(x)$, y por lo tanto se cumple (4.4).

Probaremos ahora que Π^\sharp es semicontinua superior. Como Π^\sharp es monótona sabemos que para cada subconjunto $B \subset S$ se verifica:

$$\Pi^\sharp\left(\bigwedge_{b \in B} b\right) \leq \bigwedge_{b \in B} \Pi^\sharp(b). \quad (4.5)$$

Utilizando el hecho de que Π es monótona y la desigualdad (4.3) obtenemos:

$$\Pi\left(\bigwedge_{b \in B} \Pi^\sharp(b)\right) \leq \bigwedge_{b \in B} \Pi \circ \Pi^\sharp(b) \leq \bigwedge_{b \in B} b.$$

Por la definición de Π^\sharp resulta entonces que:

$$\bigwedge_{b \in B} \Pi^\sharp(b) \leq \Pi^\sharp\left(\bigwedge_{b \in B} b\right). \quad (4.6)$$

De (4.5) y (4.6) obtenemos entonces que Π^\sharp es semicontinua superior.

3 \Rightarrow 2 Por (4.3) tenemos: $\Pi \circ \Pi^\sharp(\varepsilon) \leq \varepsilon$, y entonces $\Pi \circ \Pi^\sharp(\varepsilon) = \varepsilon$. Como $\varepsilon \leq \Pi^\sharp(\varepsilon)$, resulta que $\Pi(\varepsilon) \leq \Pi \circ \Pi^\sharp(\varepsilon) = \varepsilon$, y por lo tanto $\Pi(\varepsilon) = \varepsilon$.

Veamos que Π es semicontinua inferior. Sea $X \subset D$. Como Π es monótona sabemos que se verifica:

$$\bigvee_{x \in X} \Pi(x) \leq \Pi\left(\bigvee_{x \in X} x\right). \quad (4.7)$$

Para cada $x \in X$ definamos $b_x = \Pi(x)$. Por la desigualdad (4.4), $\Pi^\sharp(b_x) \geq x$, y por lo tanto

$$\begin{aligned} \Pi\left(\bigvee_{x \in X} x\right) &\leq \Pi\left(\bigvee_{x \in X} \Pi^\sharp(b_x)\right) \leq \Pi \circ \Pi^\sharp\left(\bigvee_{x \in X} b_x\right) \\ &\leq \bigvee_{x \in X} b_x = \bigvee_{x \in X} \Pi(x), \end{aligned} \quad (4.8)$$

donde utilizamos el hecho de que Π^\sharp es monótona y que se cumple la desigualdad (4.3). De las desigualdades (4.7) y (4.8) resulta entonces que Π es semicontinua inferior.

2 \Rightarrow 1 Debido a que $\Pi(\varepsilon) = \varepsilon$, el subconjunto X_b de las subsoluciones de $\Pi(x) = b$ es no vacío para cada $b \in S$. Como $\Pi(x) \leq b$ para cada $x \in X_b$ y como además Π es semicontinua inferior, resulta:

$$\Pi\left(\bigvee_{x \in X_b} x\right) = \bigvee_{x \in X_b} \Pi(x) \leq b.$$

Por lo tanto $\bigvee_{x \in X_b} x$ también es una subsolución de $\Pi(x) = b$ (es la mayor subsolución).

Notemos finalmente que como por definición la mayor subsolución es única, las equivalencias anteriores definen una única función Π^\sharp . \square

Observación 4.13. Si Π es una función residuable, entonces se verifica:

$$\Pi^{-1}([\leftarrow, x]) = [\leftarrow, \Pi^\sharp(x)].$$

Además como $\Pi(\top) \leq \top$, resulta que $\Pi^\sharp(\top) \geq \top$, y por lo tanto $\Pi^\sharp(\top) = \top$.

Tenemos el siguiente teorema para el caso en el cual nos interesamos en buscar la menor supersolución en lugar de la mayor subsolución.

Teorema 4.14. *Sea Π una función monótona del lattice completo D en el lattice completo S . Entonces las siguientes afirmaciones son equivalentes:*

1. *para todo $b \in S$ existe la menor supersolución de $\Pi(x) = b$ (esto es, el conjunto de las supersoluciones no es vacío y posee un elemento mínimo);*
2. *$\Pi(\top) = \top$ y Π es semicontinua superior (o equivalentemente, la preimagen de todo conjunto superior cerrado es distinto del vacío y es un conjunto superior cerrado);*
3. *existe una función Π^\flat de S en D que es monótona y semicontinua inferior tal que:*

$$\Pi \circ \Pi^\flat \geq I_S \quad (\text{identidad en } S); \quad (4.9)$$

$$\Pi^\flat \circ \Pi \leq I_D \quad (\text{identidad en } D). \quad (4.10)$$

Cuando Π verifica estas propiedades se dice que es dualmente residuable, y entonces la función Π^\flat (unívocamente determinada por (4.9) y (4.10)) recibe el nombre de residual dual de Π .

Notemos que si Π es una función residuable, entonces su residual Π^\sharp es dualmente residuable pues, como demostramos en el Teorema 4.12, Π^\sharp es semicontinua superior y además verifica que $\Pi^\sharp(\top) = \top$ (ver Observación 4.13). Además, de la afirmación 3 del teorema anterior y de las desigualdades (4.3) y (4.4), obtenemos que $(\Pi^\sharp)^\flat = \Pi$.

Observación 4.15. Si Π es una función dualmente residuable, entonces tenemos que

$$\Pi^{-1}([x, \rightarrow]) = [\Pi^\flat(x), \rightarrow] .$$

Además como $\varepsilon \preceq \Pi(\varepsilon)$, resulta que $\Pi^\flat(\varepsilon) \preceq \varepsilon$, y por lo tanto $\Pi^\flat(\varepsilon) = \varepsilon$.

Ejemplo 4.16. Consideremos la función $\Pi : \overline{\mathbb{Z}}_{\text{máx}} \rightarrow \overline{\mathbb{R}}_{\text{máx}}$ donde $\Pi(x) = x$. Entonces, como claramente la pre-imagen de cada conjunto inferior cerrado es un conjunto inferior cerrado no vacío, resulta que Π es residuable. La residual de Π es la función que fue estudiada en el Ejemplo 4.6, es decir, es la parte entera “desde abajo” de un número real. Un razonamiento análogo muestra que Π también es dualmente residuable y que su residual dual es la parte entera “desde arriba” de un número real.

Ejemplo 4.17. Sea S un lattice completo. Consideremos la función Π de S en S^2 definida por: $\Pi(x) = (x, x)$. Notemos que una subsolución de $\Pi(z) = b = (x, y)$ es un elemento z de D que verifica: $z \preceq x$ y $z \preceq y$. Entonces Π es una función residuable y su residual está dada por:

$$\Pi^\sharp(x, y) = x \wedge y .$$

Análogamente, Π es dualmente residuable y su residual dual es $\Pi^\flat(x, y) = x \vee y$.

4.2.2. Resultados adicionales

En el siguiente teorema se resumen algunas propiedades de las funciones residuales y dualmente residuales, así como de las residuales y de las residuales duales.

Teorema 4.18.

1. Si Π es una función residuable de D en S , entonces

$$\begin{aligned} \Pi \circ \Pi^\sharp \circ \Pi &= \Pi ; \\ \Pi^\sharp \circ \Pi \circ \Pi^\sharp &= \Pi^\sharp . \end{aligned}$$

Además se verifican las siguientes equivalencias:

$$\begin{aligned} \Pi^\sharp \circ \Pi &= I_D \Leftrightarrow \Pi \text{ es inyectiva} \Leftrightarrow \Pi^\sharp \text{ es sobreyectiva;} \\ \Pi \circ \Pi^\sharp &= I_S \Leftrightarrow \Pi^\sharp \text{ es inyectiva} \Leftrightarrow \Pi \text{ es sobreyectiva.} \end{aligned}$$

Las mismas afirmaciones se verifican para las funciones dualmente residuales cambiando \sharp por \flat .

2. Si $\Pi : D \rightarrow S$ y $\Phi : S \rightarrow C$ son funciones residuables, entonces $\Phi \circ \Pi$ también es residuable y

$$(\Phi \circ \Pi)^\sharp = \Pi^\sharp \circ \Phi^\sharp .$$

Nuevamente, las mismas afirmaciones se verifican con \flat en lugar de \sharp .

3. Si Π, Φ, Ψ y Θ son funciones de D en D , y si Π y Θ son residuables, entonces

$$\Pi \circ \Phi \preceq \Psi \circ \Theta \Leftrightarrow \Phi \circ \Theta^\sharp \preceq \Pi^\sharp \circ \Psi .$$

Como corolario tenemos:

$$\Pi \preceq \Theta \Leftrightarrow \Theta^\sharp \preceq \Pi^\sharp ,$$

y

$$\Pi \preceq I_D \Leftrightarrow \Pi^\sharp \succeq I_D ,$$

$$\Pi \succeq I_D \Leftrightarrow \Pi^\sharp \preceq I_D .$$

Afirmaciones similares se verifican para residuales duales bajo apropiadas hipótesis. En particular se verifica:

$$\Phi \circ \Pi \preceq \Theta \circ \Psi \Leftrightarrow \Theta^\flat \circ \Phi \preceq \Psi \circ \Pi^\flat .$$

4. Si Π y Φ son dos funciones residuables del dioide D (en el cual \wedge está definida) en sí mismo, entonces $\Pi \oplus \Phi$ es residuable y

$$(\Pi \oplus \Phi)^\sharp = \Pi^\sharp \wedge \Phi^\sharp .$$

Si Π y Φ son dualmente residuables, entonces $\Pi \wedge \Phi$ es dualmente residuable y

$$(\Pi \wedge \Phi)^\flat = \Pi^\flat \oplus \Phi^\flat .$$

5. Si Π y Φ son dos funciones residuables del dioide D (en el cual \wedge está definida) en sí mismo y si $\Pi \wedge \Phi$ es residuable, entonces

$$(\Pi \wedge \Phi)^\sharp \succeq \Pi^\sharp \oplus \Phi^\sharp .$$

Si Π, Φ y $\Pi \oplus \Phi$ son dualmente residuables, entonces

$$(\Pi \oplus \Phi)^\flat \preceq \Pi^\flat \wedge \Phi^\flat .$$

Observación 4.19. Consideremos nuevamente el Lema 4.10. Notemos que si Π es residuable, entonces el máximo elemento de la clase de equivalencia de x está dado por $\Pi^\sharp \circ \Pi(x)$.

La función $\Pi^\sharp \circ \Pi$ es un *proyector*,³ pues por la primera afirmación del teorema anterior tenemos:

$$\Pi^\sharp \circ \Pi \circ \Pi^\sharp \circ \Pi = \Pi^\sharp \circ \Pi .$$

Más adelante consideraremos nuevamente a estas funciones.

³Una función $\Pi : D \rightarrow D$ se dice que es un proyector cuando $\Pi = \Pi \circ \Pi$.

4.2.3. Residuación restringida

La residuación tiene como objetivo el de dar la mejor (en el sentido de que es la mayor o la menor según el caso) subsolución o supersolución de la ecuación $\Pi(x) = b$, cuando $\Pi : D \rightarrow S$ es una función monótona que verifica convenientes propiedades de semicontinuidad. A veces puede ser necesario buscar la mejor “solución aproximada” de dicha ecuación sobre un subconjunto $D^{\text{ad}} \subset D$ y no sobre todo el conjunto D (supondremos que D^{ad} es un sup-semilattice, es decir, es cerrado con respecto a \vee). Este problema recibe el nombre de *residuación restringida*. Representemos con I a la inclusión canónica⁴ de D^{ad} en D . Notemos entonces que “resolver” la ecuación $\Pi(x) = b$ sobre D^{ad} conduce a estudiar la ecuación:

$$\Pi \circ I(x) = b . \quad (4.11)$$

Si D^{ad} es un sup-semilattice completo, entonces I es una función residuable, con lo cual (por el Teorema 4.18) resulta que la mayor subsolución de la ecuación (4.11) está dada por:

$$x = (\Pi \circ I)^{\sharp}(b) = I^{\sharp} \circ \Pi^{\sharp}(b) .$$

Este elemento x puede considerarse nuevamente como elemento de D si le aplicamos I . Finalmente obtenemos:

$$I(x) = I \circ I^{\sharp}(y) \quad \text{donde} \quad y = \Pi^{\sharp}(b) .$$

Notemos que este procedimiento puede dividirse en dos pasos:

1. primero se considera el problema sin restricciones (es decir, se busca la solución de $\Pi(x) = b$ en D); esto conduce a la solución y ;
2. como segundo paso se “proyecta” la solución y sobre el subconjunto de las soluciones admisibles D^{ad} ; notemos que $I \circ I^{\sharp}$ es un proyector con imagen D^{ad} .

Se puede realizar un razonamiento análogo en el caso en el cual Π sea dualmente residuable; pero ahora para que I sea dualmente residuable es necesario suponer que D^{ad} sea cerrado con respecto a \wedge (de forma tal de que D^{ad} sea un inf-semilattice completo).

Observación 4.20. A veces puede suceder que una función $\Pi : D \rightarrow S$ no sea residuable; pero que esta propiedad sí se verifique cuando se restringe el lado derecho de la ecuación $\Pi(x) = b$ a cierto subconjunto $S^{\text{ad}} \subset S$. Por ejemplo, en un dioide completo D , la función $\Pi : x \mapsto x^*$ (ver §2.5) de D en D no es residuable ya que ni siquiera es un \oplus -morfismo. Para mostrar que Π no es un \oplus -morfismo, tomemos por ejemplo $D = (\overline{\mathbb{R}}_{\text{máx}})^{2 \times 2}$. Entonces si

$$a = \begin{pmatrix} \varepsilon & 1 \\ -1 & \varepsilon \end{pmatrix} \quad \text{y} \quad b = \begin{pmatrix} \varepsilon & -1 \\ 1 & \varepsilon \end{pmatrix} ,$$

⁴ I es la función de D^{ad} en D definida por: $I(x) = x$.

resulta que

$$\begin{aligned} (a \oplus b)^* &= \begin{pmatrix} \varepsilon & 1 \\ 1 & \varepsilon \end{pmatrix}^* = \begin{pmatrix} \top & \top \\ \top & \top \end{pmatrix} \neq a^* \oplus b^* \\ &= \begin{pmatrix} e & 1 \\ -1 & e \end{pmatrix} \oplus \begin{pmatrix} e & -1 \\ 1 & e \end{pmatrix} = \begin{pmatrix} e & 1 \\ 1 & e \end{pmatrix}. \end{aligned}$$

Sin embargo, como veremos en el Teorema 4.27, Π es residuable cuando se toma a S^{ad} como la imagen de Π .

4.2.4. ¿ Cuándo la ecuación $\Pi(x) = b$ admite solución ?

Antes de finalizar esta sección es bueno insistir en el principal objetivo de la teoría de residuación que es el de resolver ecuaciones de la forma $\Pi(x) = b$ cuando Π es una función monótona. La mejor técnica para resolver este problema es la siguiente:

1. buscar la mayor subsolución o la menor supersolución dependiendo de si Π es semicontinua inferior (y $\Pi(\varepsilon) \leq b$) o semicontinua superior (y $\Pi(\top) \geq b$) respectivamente;⁵
2. ver si

$$\Pi \circ \Pi^{\sharp}(b) = b \quad \text{o bien} \quad \Pi \circ \Pi^{\flat}(b) = b. \quad (4.12)$$

Observemos que existe una solución de $\Pi(x) = b$ si y solamente si se verifica una de las igualdades de (4.12) (siempre y cuando se cumpla la propiedad de semicontinuidad correspondiente). Por lo tanto, la residuación nos provee de un método efectivo y constructivo para determinar si la ecuación $\Pi(x) = b$ posee solución.

4.3. Funciones clausura y clausura de funciones

4.3.1. Funciones clausura y clausura dual

Estudiaremos ahora una clase especial de funciones de un lattice en sí mismo que serán de interés más adelante.

Definición 4.21. Sea D un lattice y $\Pi : D \rightarrow D$ una función monótona tal que $\Pi \circ \Pi = \Pi \geq I_D$, entonces Π recibe el nombre de *función clausura*. Si $\Pi \circ \Pi = \Pi \leq I_D$, entonces Π recibe el nombre de *función clausura dual*.

Ejemplo 4.22. Como primer ejemplo de una función clausura consideremos sobre cualquier lattice D la función $\Pi : x \mapsto a \vee x$. Entonces como

$$\Pi \circ \Pi(x) = a \vee (a \vee x) = a \vee x = \Pi(x) \geq x,$$

obtenemos que Π es una función clausura.

⁵Si $\Pi(\varepsilon) \not\leq b$ o $\Pi(\top) \not\geq b$, entonces claramente la ecuación no tiene solución.

Ejemplo 4.23. Sea D un dioide completo. Consideremos la función $\Pi : x \mapsto x^*$ (ver §2.5). Entonces como claramente $x^* \succeq x$ y $(x^*)^* = x^*$, obtenemos que Π es una función clausura.

Teorema 4.24. Si $\Pi : D \rightarrow D$ es una función residuable, entonces las siguientes cuatro afirmaciones son equivalentes:

1. $\Pi \circ \Pi = \Pi \succeq I_D$ (es decir, Π es una función clausura);
2. $\Pi^\sharp \circ \Pi^\sharp = \Pi^\sharp \preceq I_D$ (es decir, Π^\sharp es una función clausura dual);
3. $\Pi^\sharp = \Pi \circ \Pi^\sharp$;
4. $\Pi = \Pi^\sharp \circ \Pi$.

Demostración.

1 \Rightarrow 2 Esta implicación es consecuencia de los ítems 2 y 3 del Teorema 4.18.

2 \Rightarrow 3 Como $\Pi^\sharp \circ \Pi^\sharp = \Pi^\sharp$ resulta que $\Pi \circ \Pi^\sharp \circ \Pi^\sharp = \Pi \circ \Pi^\sharp$. Por (4.3) tenemos: $\Pi^\sharp \succeq \Pi \circ \Pi^\sharp \circ \Pi^\sharp$. Además $\Pi \circ \Pi^\sharp \succeq \Pi^\sharp$, pues $\Pi^\sharp \preceq I_D \Rightarrow \Pi \succeq I_D$ (ítem 3 del Teorema 4.18). Por lo tanto, tenemos que $\Pi^\sharp \succeq \Pi \circ \Pi^\sharp \circ \Pi^\sharp = \Pi \circ \Pi^\sharp \succeq \Pi^\sharp$, y entonces $\Pi^\sharp = \Pi \circ \Pi^\sharp$.

3 \Rightarrow 4 Como $\Pi^\sharp = \Pi \circ \Pi^\sharp$, resulta que $\Pi \circ \Pi^\sharp \circ \Pi = \Pi^\sharp \circ \Pi$. Pero por el ítem 1 del Teorema 4.18 sabemos que: $\Pi \circ \Pi^\sharp \circ \Pi = \Pi$. Por lo tanto, $\Pi = \Pi^\sharp \circ \Pi$.

4 \Rightarrow 1 Como $\Pi = \Pi^\sharp \circ \Pi$, tenemos que $\Pi \succeq I_D$ por (4.4). Además, $\Pi = \Pi^\sharp \circ \Pi \Rightarrow \Pi \circ \Pi = \Pi \circ \Pi^\sharp \circ \Pi$, y entonces (por el ítem 1 del Teorema 4.18) resulta: $\Pi \circ \Pi = \Pi$. \square

El teorema anterior nos dice que toda función clausura residuable Π puede factorizarse como $\Pi^\sharp \circ \Pi$. En realidad, se puede probar que toda función clausura Π definida sobre D puede factorizarse como $\Psi^\sharp \circ \Psi$ para cierta función $\Psi : D \rightarrow S$, donde S es un conjunto ordenado. Más adelante daremos otra caracterización de las funciones clausura semicontinuas inferiores. El siguiente teorema es el análogo del teorema anterior para el caso de funciones dualmente residuales.

Teorema 4.25. Si $\Pi : D \rightarrow D$ es una función dualmente residuable, entonces las siguientes cuatro afirmaciones son equivalentes:

1. $\Pi \circ \Pi = \Pi \succeq I_D$ (es decir, Π es una función clausura);
2. $\Pi^b \circ \Pi^b = \Pi^b \preceq I_D$ (es decir, Π^b es una función clausura dual);
3. $\Pi = \Pi \circ \Pi^b$;
4. $\Pi^b = \Pi^b \circ \Pi$.

Lema 4.26. Si Π y Φ son funciones clausura sobre D que también son \wedge -morfismos, entonces $\Pi \wedge \Phi$ es una función clausura. Análogamente, si Π y Φ son funciones clausura dual sobre D que también son \vee -morfismos, entonces $\Pi \vee \Phi$ es una función clausura dual. Las afirmaciones anteriores se extienden a un número infinito de funciones si las mismas son semicontinuas superiores e inferiores respectivamente.

Demostración. Probaremos la primera afirmación solamente. Como $\Pi \succeq I_D$ y $\Phi \succeq I_D$, claramente resulta que $\Pi \wedge \Phi \succeq I_D$. Además

$$\begin{aligned} (\Pi \wedge \Phi) \circ (\Pi \wedge \Phi) &= \Pi \circ \Pi \wedge \Pi \circ \Phi \wedge \Phi \circ \Pi \wedge \Phi \circ \Phi \\ &= \Pi \wedge \Pi \circ \Phi \wedge \Phi \circ \Pi \wedge \Phi = \Pi \wedge \Phi, \end{aligned}$$

pues $\Pi \succeq I_D \Rightarrow \Pi \circ \Phi \succeq \Phi$ y $\Phi \succeq I_D \Rightarrow \Phi \circ \Pi \succeq \Pi$. Por lo tanto, $\Pi \wedge \Phi$ es una función clausura. \square

Una función clausura sobre un lattice D no es necesariamente un \wedge -morfismo ni tampoco un \vee -morfismo. Por ejemplo, ya mostramos (ver Observación 4.20) que la función clausura $\Pi(x) = x^*$ sobre un dioide completo no es necesariamente un \vee -morfismo. Sin embargo, probaremos a continuación que la imagen de una función clausura Π es cerrada con respecto a \wedge , y que también es un sup-semilattice; pero no con \vee (la cota superior de D) como la operación (menor) cota superior en la imagen de Π . Probaremos además que si D es un lattice completo, entonces Π puede considerarse como una función residuable cuando se restringe su codominio a la imagen de Π .

Teorema 4.27. Sea Π es una función clausura sobre un lattice D . Entonces

1. $\text{im } \Pi$ es un \wedge -semilattice. Más precisamente se verifica:

$$\Pi(x) \wedge \Pi(y) = \Pi(\Pi(x) \wedge \Pi(y)) \in \text{im } \Pi, \quad \forall x, y \in D;$$

2. $\text{im } \Pi$ es un $\widehat{\vee}$ -semilattice con la siguiente cota superior definida sobre $\text{im } \Pi$ (con la relación de orden inducida por la relación de orden de D):

$$x \widehat{\vee} y \stackrel{\text{def}}{=} \Pi(x \vee y), \quad \forall x, y \in \text{im } \Pi;$$

3. si D es un lattice completo, entonces estas operaciones también están definidas para subconjuntos infinitos, y además $\text{im } \Pi$ con el orden inducido por el orden de D es un lattice completo;
4. representemos con $\widehat{\Pi}$ a la función del lattice (D, \vee, \wedge) en el lattice $(\text{im } \Pi, \widehat{\vee}, \wedge)$ definida por $\widehat{\Pi}(x) = \Pi(x)$, $\forall x \in D$; entonces, $\widehat{\Pi}$ es un sup-morfismo,⁶

⁶Esto quiere decir que $\widehat{\Pi}$ verifica: $\widehat{\Pi}(x \vee y) = \widehat{\Pi}(x) \widehat{\vee} \widehat{\Pi}(y)$, $\forall x, y \in D$.

5. si D es un lattice completo, entonces $\widehat{\Pi}$ es residuable y $\widehat{\Pi}^\sharp$ es la inclusión canónica de $\text{im } \Pi$ en D .

Demostración.

1. Como

$$\begin{aligned} \Pi(x) \wedge \Pi(y) &= \Pi^2(x) \wedge \Pi^2(y) && \text{(pues } \Pi^2 \stackrel{\text{def}}{=} \Pi \circ \Pi = \Pi) \\ &\geq \Pi(\Pi(x) \wedge \Pi(y)) && \text{(por el Lema 4.2)} \\ &\geq \Pi(x) \wedge \Pi(y) && \text{(pues } \Pi \geq I_D) \end{aligned}$$

obtenemos que $\Pi(x) \wedge \Pi(y) = \Pi(\Pi(x) \wedge \Pi(y))$. Notemos que esto significa que la cota inferior de dos elementos de la imagen de Π pertenece a la imagen de Π . Por lo tanto, $\text{im } \Pi$ es un \wedge -semilattice.

2. Sea $z \stackrel{\text{def}}{=} \Pi(x) \widehat{\vee} \Pi(y) = \Pi(\Pi(x) \vee \Pi(y))$. Entonces, $z \geq \Pi(x) \vee \Pi(y)$ ya que $\Pi \geq I_D$. Por lo tanto, z , que pertenece a $\text{im } \Pi$ por definición, es mayor o igual que $\Pi(x)$ y $\Pi(y)$. Sea u otro elemento de la imagen de Π mayor o igual que $\Pi(x)$ y $\Pi(y)$. Notemos que como $u \in \text{im } \Pi$, existe $v \in D$ tal que $u = \Pi(v)$, y entonces $u = \Pi(v) = \Pi^2(v) = \Pi(u)$. Por lo tanto,

$$u \geq \Pi(x) \vee \Pi(y) \Rightarrow u = \Pi(u) \geq \Pi(\Pi(x) \vee \Pi(y)) = z .$$

Hemos probado entonces que z es la menor cota superior de $\Pi(x)$ y $\Pi(y)$ que pertenece a $\text{im } \Pi$. En consecuencia, $\text{im } \Pi$ es un sup-semilattice con $\widehat{\vee}$ como cota superior (con la relación de orden inducida por la relación de orden de D).

3. Si D es un lattice completo, entonces los mismos argumentos que fueron utilizados en los ítems anteriores pueden repetirse sin modificaciones cambiando al subconjunto $\{x, y\}$ por cualquier subconjunto infinito. Por lo tanto, $\text{im } \Pi$ es un lattice completo. Notemos que el mínimo elemento de $\text{im } \Pi$ es $\Pi(\varepsilon)$ y el máximo elemento es $\Pi(\top)$.
4. Para cualquier par $x, y \in D$, como $\Pi \geq I_D$, resulta que $\Pi(x) \geq x$ y $\Pi(y) \geq y$, con lo cual $\Pi(x) \vee \Pi(y) \geq x \vee y$, y entonces

$$\Pi(x) \widehat{\vee} \Pi(y) \stackrel{\text{def}}{=} \Pi(\Pi(x) \vee \Pi(y)) \geq \Pi(x \vee y) .$$

Por otro lado, $\Pi(x \vee y) \geq \Pi(x) \vee \Pi(y)$ ya que Π es monótona (ver Lema 4.2). Aplicando la función Π a ambos lados de esta desigualdad y teniendo en cuenta que $\Pi^2 = \Pi$, obtenemos:

$$\Pi(x \vee y) = \Pi^2(x \vee y) \geq \Pi(\Pi(x) \vee \Pi(y)) \stackrel{\text{def}}{=} \Pi(x) \widehat{\vee} \Pi(y) .$$

Por lo tanto, $\Pi(x \vee y) = \Pi(x) \widehat{\vee} \Pi(y)$, lo cual implica que $\widehat{\Pi}$ es un sup-morfismo.

5. Nuevamente, si D es un lattice completo, entonces la demostración anterior puede repetirse para subconjuntos infinitos en lugar de $\{x, y\}$, lo cual implica que $\widehat{\Pi}$ es semicontinua inferior. Como además la imagen del mínimo elemento de D es el mínimo elemento de $\text{im } \Pi$, obtenemos que $\widehat{\Pi}$ es residuable (ver Teorema 4.12). Consideremos ahora la inclusión canónica $I_{\text{im } \Pi}$ de $\text{im } \Pi$ en D y veamos que es igual a $\widehat{\Pi}^\sharp$. Esta función es semicontinua superior por los ítems 1 y 3. Además como $I_{\text{im } \Pi} \circ \widehat{\Pi}(x) = \Pi(x) \succeq x$ (pues $\Pi \succeq I_D$), resulta que se verifica (4.4) (ver Teorema 4.12). Finalmente, notemos que $\widehat{\Pi} \circ I_{\text{im } \Pi}(x) = \Pi(x) = x$ (recordemos que demostramos que si $x \in \text{im } \Pi$, entonces $\Pi(x) = x$), con lo cual se verifica (4.3). Por el Teorema 4.12 resulta entonces que $\widehat{\Pi}^\sharp$ es la inclusión canónica $I_{\text{im } \Pi}$ de $\text{im } \Pi$ en D . \square

El siguiente teorema es el dual del teorema anterior.

Teorema 4.28. 1. Si Π es una función clausura dual sobre un lattice D , entonces $\text{im } \Pi$ es un \vee -semilattice. Más precisamente se verifica:

$$\Pi(x) \vee \Pi(y) = \Pi(\Pi(x) \vee \Pi(y)) \in \text{im } \Pi, \quad \forall x, y \in D.$$

2. $\text{im } \Pi$ es un $\widehat{\wedge}$ -semilattice con la siguiente cota inferior definida sobre $\text{im } \Pi$ (con la relación de orden inducida por la relación de orden de D):

$$x \widehat{\wedge} y \stackrel{\text{def}}{=} \Pi(x \wedge y), \quad \forall x, y \in \text{im } \Pi.$$

3. Si D es un lattice completo, entonces estas operaciones también están definidas para subconjuntos infinitos, y además $\text{im } \Pi$ con el orden inducido por el orden de D es un lattice completo.
4. Representemos con $\widehat{\Pi}$ a la función del lattice (D, \vee, \wedge) en el lattice $(\text{im } \Pi, \vee, \widehat{\wedge})$ definida por $\widehat{\Pi}(x) = \Pi(x)$, $\forall x \in D$. Entonces, $\widehat{\Pi}$ es un inf-morfismo.
5. Si D es un lattice completo, entonces $\widehat{\Pi}$ es dualmente residuable y $\widehat{\Pi}^\flat$ es la inclusión canónica de $\text{im } \Pi$ en D .

4.3.2. Mejor función clausura y clausura dual

Consideremos el conjunto de las funciones monótonas de un lattice completo D en sí mismo. Este conjunto es un lattice completo con la relación de orden inducida por la relación de orden de D , es decir

$$\Pi \succeq \Phi \Leftrightarrow \Pi(x) \succeq \Phi(x), \quad \forall x \in D.$$

Notemos entonces que

$$\begin{aligned}\Pi \vee \Phi &: x \mapsto \Pi(x) \vee \Phi(x) , \\ \Pi \wedge \Phi &: x \mapsto \Pi(x) \wedge \Phi(x) .\end{aligned}$$

Si nos restringimos únicamente al conjunto de las funciones semicontinuas inferiores, entonces este conjunto es un dioide completo con la operación suma \oplus igual a la cota superior \vee y la operación multiplicación \otimes igual a la composición \circ (la demostración de esta afirmación es similar a la del Lema 4.7). Análogamente, el conjunto de las funciones semicontinuas superiores es dioide completo con las operaciones $\oplus = \wedge$ y $\otimes = \circ$.

Representaremos con Π^k a la función:

$$\underbrace{\Pi \circ \cdots \circ \Pi}_{k \text{ veces}}$$

y con Π^0 a I_D . También utilizaremos la siguiente notación:

$$\Pi^* = \bigvee_{k=0}^{+\infty} \Pi^k , \quad \Pi^+ = \bigvee_{k=1}^{+\infty} \Pi^k , \quad \Pi_* = \bigwedge_{k=0}^{+\infty} \Pi^k , \quad \Pi_+ = \bigwedge_{k=1}^{+\infty} \Pi^k .$$

Notemos que no hay problemas para definir Π^* o Π_* aún en el caso en el cual la función Π no sea semicontinua inferior ni semicontinua superior pues suponemos que el lattice D es completo.

De la definición resulta claramente que $\Pi^* = I_D \vee \Pi^+$, y por lo tanto $\Pi^* \succeq \Pi^+$; pero observemos que se verifica la igualdad si $\Pi \succeq I_D$. Además, por el Lema 4.2, tenemos:

$$\Pi \circ \Pi^* \succeq \Pi^* \circ \Pi = \Pi^+ ;$$

pero la igualdad se verifica si Π es semicontinua inferior. Análogamente, $\Pi_* = I_D \wedge \Pi_+$, y por lo tanto $\Pi_* \preceq \Pi_+$; pero la igualdad se verifica si $\Pi \preceq I_D$. Además nuevamente por el Lema 4.2 tenemos:

$$\Pi \circ \Pi_* \preceq \Pi_* \circ \Pi = \Pi_+ ;$$

pero la igualdad se verifica si Π es semicontinua superior.

Si Π es una función clausura, entonces $\Pi = \Pi^*$; y si Π es una función clausura dual, entonces $\Pi = \Pi_*$. Cuando Π no es una función clausura (respectivamente clausura dual), entonces como lo indica el siguiente lema, la función Π^* (respectivamente Π_*) puede considerarse como su “clausura” (respectivamente “clausura dual”) siempre y cuando se cumplan apropiadas condiciones de semicontinuidad.

Lema 4.29. *Sea Π una función de un lattice completo D en sí mismo. Si Π es semicontinua inferior, entonces Π^* es la menor función clausura que es mayor o igual que Π . Análogamente, si Π es semicontinua superior, entonces Π_* es la mayor función clausura dual que es menor o igual que Π .*

Demostración. Probaremos la primera afirmación únicamente. Como Π es semicontinua inferior, es fácil verificar que $(\Pi^*)^2 = \Pi^*$. Como además $\Pi^* \succeq I_D$, resulta que Π^* es una función clausura. Por la definición de Π^* tenemos que $\Pi^* \succeq \Pi$.

Sea ahora Ψ otra función clausura mayor o igual que Π . Entonces de $\Psi \succeq I_D$ y $\Psi \succeq \Pi$ obtenemos:

$$\Psi = \Psi^2 \succeq \Psi \circ \Pi \succeq \Pi^2.$$

Análogamente se demuestra que $\Psi = \Psi^k \succeq \Pi^k, \forall k \in \mathbb{N}$. Por lo tanto,

$$\Psi \succeq \bigvee_{k=0}^{+\infty} \Pi^k = \Pi^*,$$

con lo cual queda demostrado que Π^* es la menor función clausura que es mayor o igual que Π . \square

Como consecuencia directa del lema anterior tenemos el siguiente corolario. El mismo nos proporciona una afirmación equivalente adicional para los Teoremas 4.24 y 4.25.

Corolario 4.30. *Sea Π una función de un lattice completo D en sí mismo. Si Π es semicontinua inferior, entonces Π es una función clausura si y sólo si $\Pi = \Pi^*$. Si Π es semicontinua superior, entonces Π es una función clausura dual si y sólo si $\Pi = \Pi_*$.*

El siguiente corolario completa los resultados de los Teoremas 4.27 y 4.28.

Corolario 4.31. *Sea Π una función clausura de un lattice completo D en sí mismo. Si Π es semicontinua inferior, entonces $\text{im } \Pi$ es un lattice con las operaciones cota superior \vee y cota inferior \wedge de D . La misma afirmación se verifica si Π es una función clausura dual semicontinua superior.*

Demostración. En el Teorema 4.27 probamos que si Π es una función clausura, entonces $\text{im } \Pi$ es un \wedge -semilattice. Como Π es semicontinua inferior, claramente $\text{im } \Pi$ es también un \vee -semilattice. Por lo tanto $\text{im } \Pi$ es un lattice. Una demostración similar puede darse para la segunda afirmación. \square

4.4. Residuación de la suma y la multiplicación

4.4.1. Fórmulas principales

En esta sección estudiaremos las siguientes funciones de un dioide D en sí mismo:

$$\begin{aligned} T_a : x &\mapsto a \oplus x && \text{(traslación por } a\text{);} \\ L_a : x &\mapsto a \otimes x && \text{(multiplicación a izquierda por } a\text{);} \\ R_a : x &\mapsto x \otimes a && \text{(multiplicación a derecha por } a\text{).} \end{aligned}$$

Observemos que tenemos como consecuencia directa de la propiedades asociativa y conmutativa de la suma las siguientes igualdades:

$$T_a \circ T_b = T_b \circ T_a = T_{a \oplus b} = T_a \oplus T_b .$$

Notemos además que si D es un dioide distributivo, entonces se verifica:

$$T_a \wedge T_b = T_{a \wedge b} .$$

Para la multiplicación, la propiedad asociativa implica la igualdad:

$$L_a \circ L_b = L_{ab} ,$$

y también que

$$L_a \circ R_b = R_b \circ L_a .$$

Una consecuencia directa de la propiedad distributiva de la multiplicación con respecto a la suma es:

$$L_a \oplus L_b = L_{a \oplus b}$$

y además que

$$L_a \circ T_b = T_{ab} \circ L_a$$

(por supuesto que propiedades análogas se verifican para la multiplicación a derecha por a).

Observemos que la función L_a es semicontina inferior si y sólo si la propiedad distributiva de la multiplicación (a izquierda) con respecto a la suma se extiende a la suma de conjuntos con un número infinito de elementos. Si suponemos que esta propiedad se verifica, como además $L_a(\varepsilon) = \varepsilon$, resulta por el Teorema 4.12 que L_a es una función residuable. Las mismas consideraciones pueden realizarse para la multiplicación a derecha R_a . En particular, notemos que las funciones L_a y R_a son semicontinuas inferiores (y por lo tanto residuales) si D es un dioide completo.

Notación. Utilizaremos dos tipos de notaciones para representar a las residuales de las funciones L_a y R_a . La *división a izquierda por a* será representada por $L_a^\sharp(x) = a \backslash x$, mientras que la *división a derecha por a* será representada por $R_a^\sharp(x) = x \not\! / a$. También utilizaremos la notación:

$$L_a^\sharp(x) = \frac{x}{a} , \quad R_a^\sharp(x) = \frac{x}{a} .$$

En el caso de T_a , como $T_a(\varepsilon) \neq \varepsilon$ salvo que $a = \varepsilon$, tenemos que esta función no es residuable (observemos que la ecuación $x \oplus a = b$ no posee subsoluciones cuando $a > b$; sin embargo dicha ecuación siempre posee supersoluciones). Pero como T_a es claramente una función clausura, resulta por el Teorema 4.27 que esta

función es residuable si la consideramos como una función de D en $a \oplus D$. Sin embargo, esto es poco interesante ya que su residual es prácticamente la identidad. Por tal motivo, supondremos que D es un dioide distributivo, pues en ese caso T_a es semicontinua superior, y como además se verifica que $T_a(\top) = \top$, resulta que T_a es dualmente residuable.

Notación. Utilizaremos la siguiente notación: $T_a^b(x) = x \ominus a$.

A partir de la definición de $x \ominus a$ claramente tenemos:

$$x \ominus a = \varepsilon \Leftrightarrow a \succeq x .$$

Observación 4.32. Cada vez que consideremos a las funciones L_a^\sharp , R_a^\sharp y T_a^b supondremos implícitamente que D es un dioide distributivo y completo de manera tal que las mismas estén bien definidas.

En las siguientes tablas resumimos las principales propiedades que verifican las nuevas operaciones “división” y “substracción”. Las mismas son consecuencia directa de las propiedades generales que fueron presentadas en §4.2.

Cuadro 4.1: Fórmulas que involucran la división (primera parte)

$\frac{x \wedge y}{a} = \frac{x}{a} \wedge \frac{y}{a}$	$\frac{x \wedge y}{a} = \frac{x}{a} \wedge \frac{y}{a}$	(f.1)
$\frac{x \oplus y}{a} \succeq \frac{x}{a} \oplus \frac{y}{a}$	$\frac{x \oplus y}{a} \succeq \frac{x}{a} \oplus \frac{y}{a}$	(f.2)
$\frac{x}{a \oplus b} = \frac{x}{a} \wedge \frac{x}{b}$	$\frac{x}{a \oplus b} = \frac{x}{a} \wedge \frac{x}{b}$	(f.3)
$\frac{x}{a \wedge b} \succeq \frac{x}{a} \oplus \frac{x}{b}$	$\frac{x}{a \wedge b} \succeq \frac{x}{a} \oplus \frac{x}{b}$	(f.4)
$a \frac{x}{a} \preceq x$	$\frac{x}{a} a \preceq x$	(f.5)
$\frac{ax}{a} \succeq x$	$\frac{xa}{a} \succeq x$	(f.6)
$a \frac{ax}{a} = ax$	$\frac{xa}{a} a = xa$	(f.7)

4.4.2. Aplicaciones y ejemplos

Una consecuencia interesante de las propiedades anteriores es la descomposición de cualquier x con respecto a cualquier y dada por la fórmula (f.21) de la

Cuadro 4.2: Fórmulas que involucran la división (segunda parte)

$\frac{a(a \setminus x)}{a} = \frac{x}{a}$	$\frac{(x \not\phi a)a}{a} = \frac{x}{a}$	(f.8)
$\frac{x}{ab} = \frac{a \setminus x}{b}$	$\frac{x}{ba} = \frac{x \not\phi a}{b}$	(f.9)
$\frac{a \setminus x}{b} = \frac{x \not\phi b}{a}$	$\frac{x \not\phi a}{b} = \frac{b \setminus x}{a}$	(f.10)
$b \frac{x}{a} \preceq \frac{x}{a \not\phi b}$	$\frac{x}{a} b \preceq \frac{x}{b \setminus a}$	(f.11)
$\frac{x}{a} b \preceq \frac{xb}{a}$	$b \frac{x}{a} \preceq \frac{bx}{a}$	(f.12)
$\frac{x}{a} \oplus b \preceq \frac{x \oplus ab}{a}$	$\frac{x}{a} \oplus b \preceq \frac{x \oplus ba}{a}$	(f.13)

Cuadro 4.3: Fórmulas que involucran la sustracción

$(x \oplus y) \ominus a = (x \ominus a) \oplus (y \ominus a)$	(f.14)
$(x \wedge y) \ominus a \preceq (x \ominus a) \wedge (y \ominus a)$	(f.15)
$(x \ominus a) \oplus a = x \oplus a$	(f.16)
$(x \oplus a) \ominus a = x \ominus a$	(f.17)
$x \ominus (a \oplus b) = (x \ominus a) \ominus b = (x \ominus b) \ominus a$	(f.18)
$x \ominus (a \wedge b) = (x \ominus a) \oplus (x \ominus b)$	(f.19)
$ax \ominus ab \preceq a(x \ominus b)$	(f.20)
$x = (x \wedge y) \oplus (x \ominus y)$	(f.21)

Tabla 4.3. Su justificación es la siguiente:

$$\begin{aligned} (x \wedge y) \oplus (x \ominus y) &= (x \oplus (x \ominus y)) \wedge (y \oplus (x \ominus y)) \\ &= x \wedge (x \oplus y) \\ &= x , \end{aligned}$$

donde la primera igualdad se debe a que suponemos que D es un dioide distributivo, la segunda está basada en que $x \ominus y \preceq x$ (pues x siempre es una supersolución de $a \oplus y = x$) y en la fórmula (f.16), mientras que la tercera es obvia. Como corolario de esta descomposición tenemos:

$$\begin{aligned} x \oplus y &= (x \wedge y) \oplus (x \ominus y) \oplus (y \wedge x) \oplus (y \ominus x) \\ &= (x \ominus y) \oplus (x \wedge y) \oplus (y \ominus x) . \end{aligned}$$

Observación 4.33. Observemos que la fórmula (f.3) es equivalente a:

$$L_{a \oplus b}^{\sharp}(x) = L_a^{\sharp}(x) \wedge L_b^{\sharp}(x) ,$$

mientras que la fórmula (f.9) es equivalente a:

$$L_{ab}^{\sharp}(x) = L_b^{\sharp} \circ L_a^{\sharp}(x) .$$

Por lo tanto, la función $a \mapsto L_a^{\sharp}$ es un homomorfismo de D en el dioide de las funciones semicontinuas superiores de D en D (ver Lema 4.7).

Análogamente, la fórmula (f.19) puede escribirse como:

$$T_{a \wedge b}^b(x) = T_a^b(x) \oplus T_b^b(x) ,$$

mientras que la fórmula (f.18) puede escribirse como:

$$T_{a \oplus b}^b(x) = T_a^b \circ T_b^b(x) .$$

Recordemos que D posee una estructura de dioide con la suma $\tilde{\oplus} \stackrel{\text{def}}{=} \wedge$ y la multiplicación $\tilde{\otimes} \stackrel{\text{def}}{=} \oplus$, pues suponemos que D es un dioide distributivo. Entonces las igualdades anteriores nos dicen que la función $a \mapsto T_a^b$ es un homomorfismo del dioide $(D, \tilde{\oplus}, \tilde{\otimes})$ en el dioide de las funciones semicontinuas inferiores de D en D (ver Lema 4.7). Notemos que la multiplicación de funciones semicontinuas inferiores es idempotente y conmutativa cuando nos restringimos a las funciones de la forma T_a^b .

Ejemplo 4.34. Consideremos el dioide completo $\overline{\mathbb{R}}_{\text{máx}}$. Como consecuencia directa de la definición tenemos:

$$a \ominus b = \begin{cases} a & \text{si } b < a ; \\ \varepsilon & \text{en caso contrario.} \end{cases}$$

En cuanto a $a \not\leq b$, que en este caso es igual a $b \setminus a$ pues la multiplicación es conmutativa, notemos que por la definición tenemos: $a \not\leq b = a - b$ (la resta habitual)

si $a, b \in \mathbb{R}$. También, como consecuencia directa de la definición, obtenemos que $a \not\leq b = \top$ si $b = \varepsilon$ o bien $a = \top$. En particular resulta:

$$\varepsilon \not\leq \varepsilon = \top \not\leq \top = \top,$$

lo cual conduce a la regla $\infty - \infty = +\infty$ en la notación habitual.

Observemos que la notación habitual no debería utilizarse ya que conduce a confusión. En realidad, ya mostramos que la igualdad

$$\top \otimes \varepsilon = \varepsilon$$

conduce a la regla $\infty - \infty = -\infty$ con la notación habitual, la cual contradice la regla que acabamos de encontrar.

Ejemplo 4.35. Como segundo ejemplo consideremos el dioide $(2^{\mathbb{R}^2}, \cup, +)$. Supongamos en primer lugar que A y B son los discos representados en la Figura 4.3. Entonces, por definición, $C = B \ominus A$ es el menor subconjunto $C \subset \mathbb{R}^2$ tal que $B \subseteq C \oplus A = C \cup A$. Este subconjunto está representado en Figura 4.3, en donde también se ilustra la fórmula (f.16) en este caso particular.

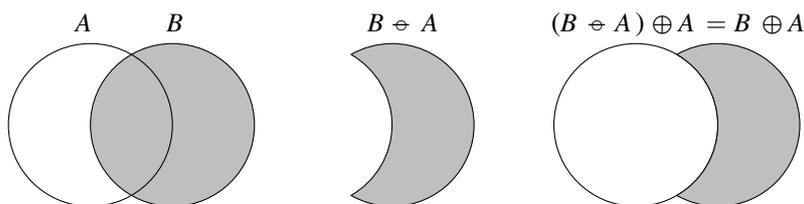


Figura 4.3: La operación \ominus en $(2^{\mathbb{R}^2}, \cup, +)$

Supongamos ahora que A es el disco centrado en el origen de la Figura 4.4, mientras que B es el cuadrado de dicha figura. Entonces, por definición, $C = B \not\leq A$ es el mayor subconjunto $C \subset \mathbb{R}^2$ que verifica: $C \otimes A = C + A \subseteq B$. Este subconjunto es el cuadrado que está representado en la parte media de la Figura 4.4, mientras que en el lado derecho de dicha figura se ilustra la propiedad (f.5).

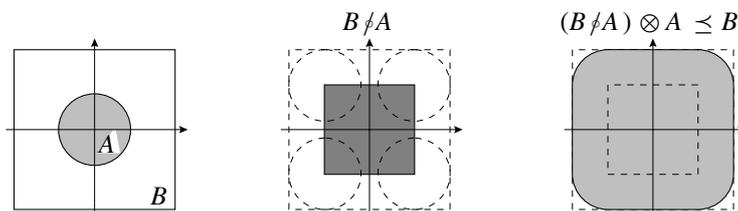


Figura 4.4: La operación $\not\leq$ en $(2^{\mathbb{R}^2}, \cup, +)$

Finalizaremos esta sección estudiando el problema de la resolución de ecuaciones de la forma

$$ax \oplus b = c \tag{4.13}$$

en el sentido de la mayor subsolución. Esto conduce a calcular la residual de la función $H = \widehat{A}_b \circ L_a$, en donde \widehat{A}_b es igual a la función T_b salvo por el hecho de que su codominio es $b \oplus D$. Recordemos que es necesario restringir el codominio de T_b a $b \oplus D$ para obtener la función residuable \widehat{A}_b . Más precisamente, el subconjunto de las subsoluciones de la ecuación (4.13) no es vacío si y sólo si $b \preceq c$. Entonces $H^\sharp = L_a^\sharp \circ \widehat{A}_b^\sharp = L_a^\sharp$, pues $\widehat{A}_b^\sharp = I_{b \oplus D}$ (ver Teorema 4.27). Estas observaciones se resumen en el siguiente lema.

Lema 4.36. *Existe la mayor subsolución \widehat{x} de $ax \oplus b = c$ si y sólo si $b \preceq c$. En dicho caso $\widehat{x} = L_a^\sharp(c) = a \backslash c$.*

Ejemplo 4.37. Consideremos la función $H : x \mapsto x \backslash a$. Esta función verifica la propiedad:

$$x \preceq y \Rightarrow H(x) \succeq H(y) ,$$

pues $y \backslash a$ es una subsolución de $x \otimes z = a$ ($x \otimes (y \backslash a) \preceq y \otimes (y \backslash a) \preceq a$). Además, por la fórmula (f.3), sabemos que se verifica:

$$H(x \oplus y) = H(x) \wedge H(y) ,$$

lo cual también se cumple para sumas con infinitos términos. Por último, sabemos que $H(\varepsilon) = \top$. Por lo tanto, si la relación de orden es invertida en el codominio de H (es decir, $y \succeq z$ si $y \preceq z$, con lo cual la suma \oplus “se transforma” en \wedge , ε en \top , etc.), entonces H resulta ser una función residuable. Por lo tanto tiene sentido buscar la mayor (en el sentido original) solución de

$$x \backslash a \succeq b$$

(esto corresponde a buscar la mayor subsolución de la ecuación $H(x) = b$ cuando se modifica el codominio de H en la forma indicada). Veamos que la mayor solución está dada por $a \not\backslash b$. En primer lugar veamos que es solución. Por la propiedad (f.11) (primera columna) tenemos:

$$(a \not\backslash b) \backslash a \succeq b(a \backslash a) \succeq b ,$$

en donde la última desigualdad se debe a que $a \backslash a \succeq e$ (e siempre es subsolución de $a \otimes z = a$). Veamos ahora que $a \not\backslash b$ es la mayor solución. Sea x otra solución. Entonces, por la propiedad (f.11) (segunda columna), resulta:

$$x \backslash a \succeq b \Rightarrow a \not\backslash (x \backslash a) \preceq a \not\backslash b \Rightarrow (a \not\backslash a)x \preceq a \not\backslash b \Rightarrow x \preceq a \not\backslash b ,$$

en donde utilizamos nuevamente el hecho de que $a \not\backslash a \succeq e$ y además que $y \preceq z \Rightarrow a \not\backslash y \succeq a \not\backslash z$.

4.4.3. Algunas ecuaciones de punto fijo

En esta sección consideraremos una generalización de la ecuación y de la desigualdad de punto fijo que fueron estudiadas en §2.5 debido a que este tipo de ecuaciones y desigualdades aparecen frecuentemente en la práctica. En primer lugar consideraremos ecuaciones y desigualdades de la forma:

$$x = \Pi(x) \oplus b \quad \text{o bien} \quad x \succeq \Pi(x) \oplus b, \quad (4.14)$$

donde Π es una función semicontinua inferior de un dioide completo D en sí mismo. Para la resolución de este problema tenemos el siguiente teorema.

Teorema 4.38. *Sea Π una función semicontinua inferior de un dioide completo D en sí mismo. Entonces la menor solución de (4.14) está dada por $\Pi^*(b)$ (esta solución verifica con igualdad el caso del problema con desigualdad).*

Demostración. La demostración es una adaptación directa de la demostración del teorema presentado en §2.5 que corresponde al caso particular $\Pi(x) = ax$. \square

Para el caso de las ecuaciones y de las desigualdades de la forma

$$x = \Pi(x) \wedge b \quad \text{o bien} \quad x \preceq \Pi(x) \wedge b, \quad (4.15)$$

en donde Π es una función semicontinua superior de un dioide completo D en sí mismo, tenemos el siguiente teorema.

Teorema 4.39. *Sea Π una función semicontinua superior de un dioide completo D en sí mismo. Entonces la mayor solución de (4.15) está dada por $\Pi_*(b)$ (esta solución verifica con igualdad el caso del problema con desigualdad).*

Ejemplo 4.40. Consideremos la ecuación

$$x = a \backslash x \wedge b.$$

Esta ecuación es un caso particular de (4.15) en el cual $\Pi(x) = a \backslash x$. Una consecuencia directa de la fórmula (f.9) es que:

$$\Pi^k(x) = a^k \backslash x, \quad \forall k \in \mathbb{N}.$$

Por lo tanto, utilizando la fórmula (f.3) obtenemos:

$$\Pi_*(x) = a^* \backslash x.$$

Entonces la mayor solución de la ecuación anterior es $a^* \backslash b$.

Para finalizar esta sección consideraremos algunas identidades que involucran a a^* . Recordemos que la función L_a verifica las siguientes propiedades:

$$L_a \circ L_b = L_{ab} \quad \text{y} \quad L_a \oplus L_b = L_{a \oplus b}.$$

Una consecuencia directa de las mismas es que $(L_a)^* = L_{a^*}$. Como $(L_a)^*$ es una función clausura (ver Lema 4.29) sabemos que $((L_a)^*)^2 = (L_a)^*$, y por lo tanto $(L_{a^*})^2 = L_{a^*}$. Como además $(L_{a^*})^2 = L_{(a^*)^2}$, obtenemos que $L_{(a^*)^2} = L_{a^*}$, y en particular $(a^*)^2 = a^*$. Por lo tanto, $(a^*)^k = a^*$, $\forall k \in \mathbb{N}$. Otra consecuencia del hecho de que $(L_a)^* = L_{a^*}$ sea una función clausura es que $(L_{a^*})^* = L_{a^*}$ (ver Corolario 4.30). Pero como ya sabemos $(L_{a^*})^* = L_{(a^*)^*}$, y entonces $L_{a^*} = L_{(a^*)^*}$ de donde resulta en particular que $a^* = (a^*)^*$. Finalmente por el Teorema 4.24 sabemos que $L_{a^*} = L_{a^*}^{\sharp} \circ L_{a^*}$. Entonces, si en particular aplicamos estas funciones a e , obtenemos: $a^* = a^* \searrow a^*$.

4.5. Residuación de matrices e inversibilidad

4.5.1. Residuación de matrices

Sea D un dioide en el que existe la cota inferior \wedge y sea A una matriz de $D^{m \times n}$. Consideremos la función $L_A : D^n \rightarrow D^m$ definida por: $L_A(x) = Ax$. El propósito de esta sección es el de encontrar una fórmula para L_A^{\sharp} y además el de establecer las condiciones bajo las cuales la matriz A admite una inversa a izquierda.

En realidad, es posible tratar el caso más general de una matriz de operadores como indicamos a continuación. Con el propósito de mantener una notación simple, nos restringiremos al caso $n = 3$ y $m = 2$, ya que su generalización es directa. Entonces consideremos cinco dioides $\{D_i\}_{i=1,2,3}$ y $\{C_i\}_{i=1,2}$ y seis funciones residuales Π_{ji} de D_i en C_j . Definamos la función Π de $D_1 \times D_2 \times D_3$ en $C_1 \times C_2$ por:

$$\Pi : x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \Pi_{11}(x_1) \oplus \Pi_{12}(x_2) \oplus \Pi_{13}(x_3) \\ \Pi_{21}(x_1) \oplus \Pi_{22}(x_2) \oplus \Pi_{23}(x_3) \end{pmatrix}.$$

Para calcular Π^{\sharp} resulta conveniente considerar a Π como la suma de las siguientes tres funciones:

$$\Pi_1 = \begin{pmatrix} \Pi_{11}(x_1) \\ \Pi_{22}(x_2) \end{pmatrix}, \quad \Pi_2 = \begin{pmatrix} \Pi_{12}(x_2) \\ \Pi_{23}(x_3) \end{pmatrix}, \quad \Pi_3 = \begin{pmatrix} \Pi_{13}(x_3) \\ \Pi_{21}(x_1) \end{pmatrix},$$

ya que cada componente de cada una de estas funciones depende de una única variable x_i (que además es diferente de las variables de las cuales dependen las restantes componentes de la función), con lo cual calcular la residual de las mismas resulta una tarea sencilla. Por ejemplo,

$$x = \Pi_3^{\sharp}(y) = \begin{pmatrix} \Pi_{21}^{\sharp}(y_2) \\ \top \\ \Pi_{13}^{\sharp}(y_1) \end{pmatrix}.$$

Entonces, como $\Pi = \Pi_1 \oplus \Pi_2 \oplus \Pi_3$, por el ítem 4 del Teorema 4.18 obtenemos:

$$\Pi^\sharp(y) = \begin{pmatrix} \Pi_{11}^\sharp(y_1) \wedge \Pi_{21}^\sharp(y_2) \\ \Pi_{12}^\sharp(y_1) \wedge \Pi_{22}^\sharp(y_2) \\ \Pi_{13}^\sharp(y_1) \wedge \Pi_{23}^\sharp(y_2) \end{pmatrix}.$$

Regresando a la función L_A , el siguiente lema es un corolario de las consideraciones que acabamos de hacer si utilizamos la notación habitual $A \bowtie y$ para representar a $L_A^\sharp(y)$.

Lema 4.41. Si $A = (A_{ij}) \in D^{m \times n}$, en donde D es un dioide en el cual la cota inferior \wedge está definida, y además $y \in D^m$, entonces

$$(A \bowtie y)_i = \bigwedge_{j=1}^m (A_{ji} \bowtie y_j).$$

Observación 4.42. Notemos que L_A^\sharp no es en general un operador lineal, es decir, en general no puede expresarse como la multiplicación a izquierda por cierta matriz. Por el lema anterior, el cálculo de $A \bowtie y$ conduce a realizar, de cierta manera, una multiplicación matricial (a izquierda), en donde la multiplicación se reemplaza por la división (a izquierda) y la suma por la cota inferior. Observemos además que en el caso particular en el cual $D = \overline{\mathbb{R}}_{\text{máx}}$ obtenemos:

$$(A \bowtie y)_i = \bigwedge_{j=1}^m (A_{ji} \bowtie y_j) = \underset{j=1, \dots, m}{\text{mín}} (-A_{ji} + y_j),$$

y entonces $A \bowtie y = -A^t \otimes y$, en donde las operaciones que se utilizan para calcular $-A^t \otimes y$ son las operaciones del dioide $\overline{\mathbb{R}}_{\text{mín}}$.

Si $A, F \in D^{m \times n}$, $B \in D^{m \times p}$ y $C \in D^{n \times p}$, entonces como consecuencia directa del lema anterior obtenemos las siguientes fórmulas para $C = A \bowtie B$ y $F = B \not\leftarrow C$:

$$C_{ij} = \bigwedge_{k=1}^m (A_{ki} \bowtie B_{kj}), \quad F_{ij} = \bigwedge_{k=1}^p (B_{ik} \not\leftarrow C_{jk}). \quad (4.16)$$

Observación 4.43. Como resultado de las observaciones realizadas en §4.2.4, tenemos que la ecuación $AX = B$ tiene solución si y sólo si $A(A \bowtie B) = B$.

4.5.2. Una observación importante

Sean $A \in D^{m \times n}$, $B \in D^{m \times p}$ y $x \in D^p$. Entonces es importante tener cuidado con expresiones de la forma $A \bowtie Bx$, pues la mismas escritas de esa manera son ambiguas. Por un lado, $A \bowtie (Bx)$ es igual a $L_A^\sharp \circ L_B(x)$. La función $L_A^\sharp \circ L_B$ en general no es un \oplus -morfismo y la misma verifica solamente la desigualdad:

$$L_A^\sharp \circ L_B(x \oplus y) \geq L_A^\sharp \circ L_B(x) \oplus L_A^\sharp \circ L_B(y)$$

pues es monótona. Por otro lado, $x \mapsto (A \backslash B)x$ es un operador lineal de D^p en D^n , el cual no es otra cosa que la multiplicación a izquierda por la matriz $C = A \backslash B$ (ver (4.16)). Los operadores $L_A^\sharp \circ L_B$ y $A \backslash B$ verifican la desigualdad $L_A^\sharp \circ L_B \succeq A \backslash B$ pues por la propiedad (f.12) tenemos que: $A \backslash (Bx) \succeq (A \backslash B)x$, $\forall x \in D$.

4.5.3. Inversibilidad

Daremos ahora las condiciones bajo las cuales existe la inversa a izquierda de una matriz $A \in D^{n \times n}$, es decir, un operador B de D^n en D^n tal que $B \circ A = I$ ($I = I_{D^n}$). Esta inversa a izquierda B no debe necesariamente poder expresarse como la multiplicación a izquierda por cierta matriz. Si D es un dioide conmutativo y si $B \in D^{n \times n}$, entonces se puede probar que $BA = I \Leftrightarrow AB = I$. El siguiente lema generaliza un resultado similar para las matrices Booleanas (observemos que el álgebra de Boole es en particular un dioide arquimediano completo).

Lema 4.44. *Sea D un dioide arquimediano completo y sea A una matriz de dimensión $n \times n$ con componentes en D . Entonces una condición necesaria y suficiente para la existencia del operador inversa a izquierda de A es que exista una y solamente una componente en cada fila y en cada columna de A que sea diferente de ε y que dichas componentes tengan una inversa a izquierda.*

Observación 4.45. La función $y \mapsto A \backslash y$ (la cual representaremos con $A \backslash \cdot$) es un \wedge -morfismo pues, como ya sabemos, es semicontinua superior. Si A posee un operador inversa a izquierda B , entonces $A \backslash \cdot$ también es un \oplus -morfismo cuando nos restringimos a la imagen de A . En realidad, como $B \circ A = I$, se puede probar fácilmente que A es una función inyectiva y entonces por el ítem 1 del Teorema 4.18 tenemos que: $A \backslash (Ax) = x$, $\forall x \in D$. Por lo tanto

$$x \oplus y = A \backslash (A(x \oplus y)) = A \backslash (Ax \oplus Ay) \succeq A \backslash (Ax) \oplus A \backslash (Ay) ,$$

en donde la desigualdad es consecuencia del hecho de que $A \backslash \cdot$ es una función monótona. Pero como el último término es igual a $x \oplus y$ obtenemos:

$$A \backslash (Ax \oplus Ay) = A \backslash (Ax) \oplus A \backslash (Ay) ,$$

que es lo que queríamos mostrar.

Capítulo 5

Otras representaciones de grafos de eventos temporizados

Hemos visto que los GET pueden ser modelados como sistemas lineales de dos formas:

1. en el *dominio de los eventos* donde el índice k cuenta eventos, obtuvimos ecuaciones lineales recurrentes en álgebra máx-plus en términos de *daters*;
2. usando la transformada γ de los *daters*, donde γ es el operador *backward shift* en el conteo, obtuvimos matrices transferencias que sintetizan el comportamiento entrada-salida del sistema. Éstas son matrices de polinomios en γ con coeficientes en álgebra máx-plus.

Ahora veremos, en primer lugar, que debido a que los *daters* tienen trayectorias no decrecientes es posible considerar funciones recíprocas que asocian número de eventos al tiempo, que son llamadas “*counters*” y están asociadas con el punto de vista del *dominio temporal* ya que ahora el índice estará representando el tiempo (y por eso será denominado t en lugar de k). Este punto de vista estará más acorde con los paradigmas clásicos de la teoría de sistemas en los cuales todas las variables son expresadas en función de la variable temporal.

Sin embargo, utilizar las recíprocas o inversas de las ecuaciones de *daters* implica más bien un cambio de variables *no lineal*, por lo que parecería extraño poder encontrar ecuaciones *lineales* para los *counters*. Como veremos, la linealidad se preservará ya que el álgebra es adaptada a un nuevo punto de vista, donde en lugar de usar álgebra máx-plus, utilizaremos álgebra mín-plus (que es isomórfica)

También veremos que con los *counters* podemos asociar otra transformada, la *transformada* δ , donde δ es ahora el operador *backward shift* en el tiempo, y podremos obtener matrices transferencias cuyos elementos sean series de potencia en δ con coeficientes en álgebra mín-plus.

Podemos entonces preguntarnos cuál de los puntos de vista (el de los *daters* o de los *counters*) es el mejor. Como se verá, mostraremos que ambos tienen ventajas y la mejor forma de utilizarlos será una síntesis de ambos que se plasmará en una tercer representación “bidimensional”, lo que nos llevará a una nueva álgebra denominada álgebra $\mathcal{M}_{\text{in}}^{\text{ás}}[[\gamma, \delta]]$.

5.1. Punto de vista del “counter”

5.1.1. Introducción a los counters

En lugar de considerar la función $k \mapsto d(k)$ donde k es el índice que numera eventos y $d(k)$ es el tiempo en el cual tiene lugar el evento número k de una transición, puede considerarse la función inversa $t \mapsto c(t) = k$ tal que $d(k) = t$. Esta posibilidad de invertir los daters radica en el hecho que $d(\cdot)$ es *monótona*.

Debido a que esta función no es *estrictamente monótona* (dado que dos o más eventos sucesivos pueden ocurrir al mismo tiempo) y más aún, en cierto sentido es *discontinua*, esto es, $d(k+1)$ puede saltar a $d(k) + \theta$ con $\theta > 1$, puede ocurrir que no haya ninguno o bien que haya varios k tales que $d(k) = t$ para un determinado t . Sin embargo el propósito de la teoría de residuación es justamente solucionar este tipo de dificultades.

Para ser más precisos y preservar cierta simetría entre el dominio de los eventos (donde la variable es el índice k) y el dominio temporal (donde la variable es el índice t), asumiremos que los daters toman valores en \mathbb{Z} . Por otro lado, para aplicar la teoría de residuación, agregaremos el punto $-\infty$ y $+\infty$ a ambos dominios. Más aún, para un dater d , definimos:

$$d(-\infty) = -\infty \text{ and } d(+\infty) = +\infty, \quad (5.1)$$

que constituyen condiciones de borde que nunca contradicen el hecho que los daters son no decrecientes.

Resta definir con qué álgebras se dotarán los respectivos dominios. Para el dominio temporal, en el cual los daters adoptan valores, es natural considerar el álgebra máx-plus dado que se vió que los daters satisfacen ecuaciones lineales en la misma. Debido a que el álgebra para los counters es en cambio el álgebra mín-plus, sería lógico equipar al dominio de los eventos con la misma. Sin embargo, si se hace esto, debido a que el orden asociado con el álgebra mín-plus es inverso con respecto al orden usual, esto nos forzaría a hablar de daters *antítonos* (o sea, no crecientes) en lugar de *isotonos*. Luego, por el momento supondremos que el dominio de los eventos también estará equipado con el álgebra máx-plus, que de todos modos es isomórfica a la mín-plus.

Por último, para aplicar los resultados de la teoría de residuación (sobre todo, los Teoremas 4.12 y 4.14), debemos verificar propiedades de *semicontinuidad* de los daters. Si queremos definir counters como residuales de daters, estos daters deberán ser semicontinuos inferiores (s.c.i.); o bien s.c.s. si los queremos definir como residuales duales. De hecho, dado que el conjunto involucrado es $\overline{\mathbb{Z}}$, la única dificultad podría aparecer con expresiones de la forma

$$d\left(\bigoplus_{i \in I} k_i\right),$$

respectivamente

$$d\left(\bigwedge_{i \in I} k_i\right),$$

tal que I es infinito y el subconjunto $\{k_i\}_{i \in I}$ satisface

$$\bigoplus_{i \in I} k_i = \top = +\infty , \quad (5.2a)$$

$$\text{respectivamente } \bigwedge_{i \in I} k_i = \varepsilon = -\infty . \quad (5.2b)$$

Debido a las condiciones de borde (5.1), el problema de semicontinuidad se responde positivamente si

$$\bigoplus_{i \in I} d(k_i) = d\left(\bigoplus_{i \in I} k_i\right) = d(\top) = \top ,$$

$$\text{respectivamente } \bigwedge_{i \in I} d(k_i) = d\left(\bigwedge_{i \in I} k_i\right) = d(\varepsilon) = \varepsilon$$

para todas las familias $\{k_i\}$ que satisfacen (5.2): esto conduce a decir que

$$\lim_{k \rightarrow +\infty} d(k) = +\infty , \quad (5.3a)$$

$$\text{respectivamente } \lim_{k \rightarrow -\infty} d(k) = -\infty . \quad (5.3b)$$

En palabras,

- la condición (5.3a) significa que un número infinito de eventos requiere una cantidad infinita de tiempo (en otras palabras, no puede ocurrir un número infinito de eventos en un intervalo finito de tiempo);
- la condición (5.3b), es siempre satisfecha ya que si k_0 es el número del primer evento tenido en cuenta (generalmente, $k_0 = 0$), luego para todo $k < k_0$ por convención será $d(k) = -\infty$.

Por último, si los daters son residuales, el counter $\tilde{c} \stackrel{\text{def}}{=} d^\#$ queda explícitamente definido por

$$\tilde{c}(t) = \sup\{k \mid d(k) \leq t\} , \quad (5.4a)$$

mientras que si son dualmente residuales, $c \stackrel{\text{def}}{=} d^b$ queda explícitamente definido por

$$c(t) = \inf\{k \mid d(k) \geq t\} . \quad (5.4b)$$

En palabras, esto significa que

- $\tilde{c}(t)$ es el número del último evento ocurrido hasta el instante t incluido,
- $c(t)$ es el número del primer evento que ocurrirá desde el instante t incluido.

Se puede demostrar (ver [1, Lemma 5.30]) que

$$c(t) = \tilde{c}(t - 1) + 1 . \tag{5.5}$$

Observación 5.1. Si se satisface la condición de residuación correspondiente en (5.3) (que vimos que no son condiciones muy restrictivas), de acuerdo con (4.3)–(4.4), respectivamente (4.9)–(4.10), se tiene

$$d(\tilde{c}(t)) \leq t \text{ and } \tilde{c}(d(t)) \geq t ,$$

respectivamente $d(c(t)) \geq t \text{ and } c(d(t)) \leq t .$

De todos modos, las definiciones (5.4) pueden adoptarse siempre; pero las desigualdades anteriores podrían no cumplirse.

Veamos, a través de un ejemplos, las dos posibles definiciones del counter.

Ejemplo 5.2. Sea la siguiente trayectoria del dater

k	...	-2	-1	0	1	2	3	4	5	6	7	8	...
$d(k)$...	ε	ε	0	1	1	1	3	3	5	6	7	...

que se muestra en la Figura 5.1 (gráfica de la izquierda). Utilizando las definiciones (5.4), se obtiene

t	...	-2	-1	0	1	2	3	4	5	6	7	...
$c(t)$...	0	0	0	1	4	4	6	6	7	8	...
$\tilde{c}(t)$...	-1	-1	0	3	3	5	5	6	7	8	...

que se muestran respectivamente en las gráficas del centro y de la derecha de la Figura 5.1.

Nótese que t (el dominio temporal) está aún representado en el eje y mientras k (el dominio de los eventos) se representa sobre el eje x para poder comparar con la gráfica del dater. Obsérvese también que la zona sombreada (que será más precisamente definida luego) es la misma para las gráficas de la izquierda y del centro, mientras que está desplazada en una unidad hacia abajo y a la izquierda en la gráfica de la derecha. Este “desplazamiento en uno” horizontal y vertical es simplemente una ilustración de (5.5). El hecho que la zona sombreada sea la misma para el dater y su residual dual será el motivo por el cual preferiremos utilizar el counter c en lugar del counter \tilde{c} .

5.1.2. Ecuación de counters

Las ecuaciones recursivas satisfechas por los counters pueden establecerse mediante un razonamiento directo sobre el funcionamiento de los mismos de forma análoga al razonamiento hecho para obtener las ecuaciones de los daters. Otra forma de hacer esto es comenzar desde las ecuaciones de daters y utilizar la teoría

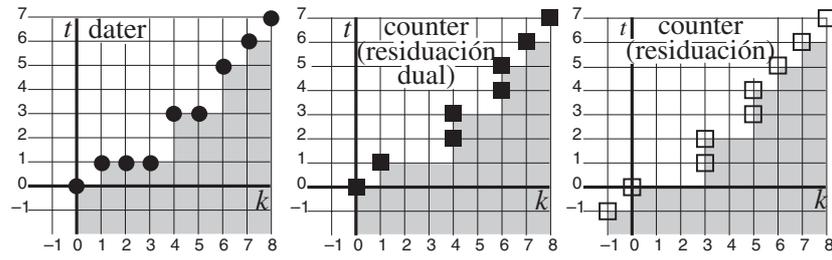


Figura 5.1: Un dater y sus dos counters asociados

de residuación para obtener dichas ecuaciones. Esto último es lo que haremos a continuación.

Observemos primero que debido a la (5.5), cualquier definición de counter que adoptemos, la ecuación recursiva debe ser la misma ya que cada definición de counter puede deducirse de la otra mediante un desplazamiento de uno en ambos ejes, y las ecuaciones “estacionarias” son invariantes frente a tales desplazamientos.¹ A continuación, deduciremos las ecuaciones para los counters $\tilde{c} = d^\sharp$.

Por definición, $x_i^\sharp(t)$ es el mayor valor k tal que $x_i(k) \leq t$, de acá que cualquier otro valor de k con esta propiedad es menor o igual que $x_i^\sharp(t)$. Dado que $x_i(k)$ está dado por la ecuación (3.1), cualquier k tal que $x_i(k) \leq t$ verifica que:

$$\begin{aligned} \forall j \in \bullet i, \quad x_j(k - M_{ij}) + t_{ij} \leq t &\Rightarrow x_j(k - M_{ij}) \leq t - t_{ij} \\ &\Rightarrow k - M_{ij} \leq x_j^\sharp(t - t_{ij}) \\ &\Rightarrow k \leq \text{mín}_{j \in \bullet i} (x_j^\sharp(t - t_{ij}) + M_{ij}) . \end{aligned}$$

Luego,

$$x_i^\sharp(t) = \text{mín}_{j \in \bullet i} (x_j^\sharp(t - t_{ij}) + M_{ij}) ,$$

o sea, los counters satisfacen ecuaciones lineales en álgebra mín-plus. Obsérvese, además, que

- los retardos en las ecuaciones están relacionados con los tiempos de espera;
- los coeficientes están relacionados con el mercado inicial.

De aquí en más no continuaremos utilizando la notación x^b ni x^\sharp para los counters. Para una transición x_i , se asocian un dater $x_i(k)$ y un counter $x_i^\sharp(t)$, que se distinguirán a través del argumento.

¹La única diferencia puede surgir respecto a las condiciones iniciales. Respecto a este tema, mencionaremos simplemente que es algo más complicado con los counters que con los daters.

Ejemplo 5.3. Volviendo al GET de §3.1.6. Con la interpretación mín-plus de la notación, las ecuaciones de counters de este GET son:

$$\begin{aligned}x_1(t) &= 1x_2(t) \oplus u_1(t - 3) , \\x_2(t) &= x_1(t - 1) \oplus 1u_2(t - 1) , \\x_3(t) &= x_1(t) \oplus x_2(t - 1) \oplus 1x_3(t - 2) , \\y(t) &= 1x_2(t) \oplus x_3(t - 3) .\end{aligned}$$

Ejercicio 5.4. Verificar las ecuaciones anteriores. Colocarlas en forma matricial como en §3.1.6.1. ¿ Cuántas matrices A_i , B_i , C_i son necesarias ? Comparar con el número correspondiente para las ecuaciones de daters. Intentar realizar una transformación similar a la de §3.1.6.2–3.1.6.3–3.1.6.4 para llegar a una forma canónica similar a (3.3). ¿ Cuántas variables de estado se necesitan (cuál es la dimensión del vector x en dichas ecuaciones canónicas) ? Comparar con lo obtenido mediante la utilización de daters.

5.1.3. La transformada δ y las matrices transferencia de counters

Un procedimiento análogo al del §3.2 puede seguirse para obtener matrices transferencia desde las ecuaciones de counters. Para esto, el nuevo operador *backward shift* δ debe introducirse. El mismo actúa simbólicamente sobre los counters $c(\cdot)$ de la siguiente forma:

$$\delta c(t) = c(t - 1) ,$$

o sea, δ es un desplazamiento en el *dominio temporal* mientras que γ lo es en el *dominio de los eventos* (desplaza la numeración de los eventos). Luego, podemos definir la transformada δ como

$$C(\delta) = \bigoplus_{t \in \mathbb{Z}} c(t) \delta^t .$$

Con esta herramienta, las ecuaciones de counters en $\overline{\mathbb{Z}}_{\min}$ se transforman mediante la transformada δ . El resultado final del proceso de eliminación de todas las variables internas es una matriz de la misma dimensión que antes (el número de filas es igual al número de salidas y el número de columnas es igual al número de entradas); pero ahora los elementos están en $\overline{\mathbb{Z}}_{\min}[[\delta]]$. En el caso de counters, las convoluciones serán en realidad *inf-convoluciones*.

Ejercicio 5.5. Volver al ejemplo del cual se obtuvieron las ecuaciones de counters en el Ejemplo 5.3. Obtener la matriz transferencia (en δ) y dibujar un GET lo más simple posible que admita dicha matriz. Compararlo con el GET original de la Figura 3.2 y con el GET obtenido en la Figura 3.7.

5.2. Más sobre los enfoques de daters y counters

Como ya se mencionó, surge naturalmente la pregunta sobre qué enfoque es conveniente. Como veremos a continuación, ambos enfoques cuentan con ventajas y desventajas comunes, y en definitiva, ninguno resulta totalmente satisfactorio. La comparación será realizada a través del análisis de ciertas propiedades y limitaciones de ambas representaciones. De este análisis se desprenderá la justificación para utilizar una nueva representación que combine las propiedades de ambos enfoques, como se verá más adelante.

5.2.1. Orden del sistema

Comenzaremos comparando el orden de los sistemas de ecuaciones obtenidos por ambos enfoques (y el costo computacional asociado) a través del análisis del ejemplo simple de la Figura 5.2.

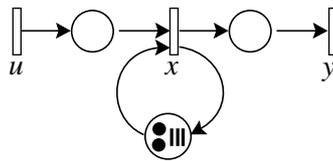


Figura 5.2: Un ejemplo simple

Las ecuaciones obtenidas para daters y counters son respectivamente:

$$\begin{array}{ll} \text{ecuaciones de daters:} & \text{ecuaciones de counters:} \\ y(k) = 3y(k-2) \oplus u(k) ; & y(t) = 2y(t-3) \oplus u(t) . \end{array} \quad (5.6)$$

Este sistema es de orden 2 en la representación por daters (requiere dos variables de estado) y de orden 3 en la representación por counters. Naturalmente, esta conclusión hubiera sido al revés intercambiando el número de barras y marcas en los lugares.

Sin embargo, suponiendo que la entrada no es restrictiva (o sea, considerando el sistema “autónomo” con $u = \varepsilon$), $y(k)$ se incrementa en 3 a cada paso de la simulación y lo mismo hacen las dos variables de estado necesarias (que podrían ser por ejemplo $y(k)$ e $y(k-1)$ para cierta realización). En la representación por counters, las tres variables de estados (que podrían ser definidas como $y(t)$, $y(t-1)$ e $y(t-2)$) se incrementan sólo en 2 a cada paso de simulación. Luego, se necesitará exactamente el mismo número de *bits* de memoria de la computadora para realizar una simulación con ambos modelos.

5.2.2. Monotonía de las trayectorias

Una dificultad que ilustraremos a través de un ejemplo, la constituye el hecho que la monotonía no es una propiedad intrínseca de las soluciones de las ecuaciones

tales como (5.6). Por ejemplo, para la entrada

$$u(k) = \begin{cases} \varepsilon & \text{si } k < 0, \\ e & \text{si } k \geq 0, \end{cases}$$

se puede verificar fácilmente que la salida

$$y(k) = \begin{cases} 3k/2 + 1 & \text{si } k \text{ es par,} \\ 3(k - 1)/2 & \text{si } k \text{ es impar,} \end{cases}$$

es una solución de las ecuaciones de daters anteriores. Explicitando los valores de dicha trayectoria se tiene:

k	...	-2	-1	0	1	2	...
$y(k)$...	-2	-3	1	4	3	...

Puede verse claramente que ésta no es una trayectoria monótona. Por supuesto puede hacerse una consideración análoga para las ecuaciones de counters.

5.2.3. Simplificaciones

A continuación deduciremos, para ambos enfoques, nuevas reglas de simplificación a partir de los ejemplos de la Figura 5.3.

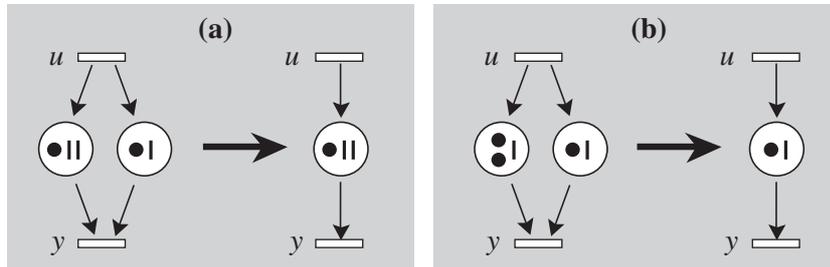


Figura 5.3: Reglas de simplificación

Consideremos en primer lugar el GET de la Figura 5.3-(a). Con daters se tiene:

$$y(k) = 2u(k - 1) \oplus 1u(k - 1) = (2 \oplus 1)u(k - 1) = 2u(k - 1),$$

lo que puede deducirse fácilmente tanto del significado de \oplus en $\overline{\mathbb{Z}}_{\text{máx}}$ como de la observación intuitiva del funcionamiento del GET.

Si ahora se utilizan counters, se tiene:

$$y(t) = 1u(t - 2) \oplus 1u(t - 1) = 1(u(t - 2) \oplus u(t - 1)) = 1u(t - 2),$$

que resulta claro teniendo en cuenta que $u(t - 2) \leq u(t - 1)$ ya que $t \mapsto u(t)$ es monótona y ahora \oplus es la operación mín.² Para la aplicación de las transformadas

²En el enfoque de counters, dado que la relación de orden en $\overline{\mathbb{Z}}_{\text{mín}}$ es la opuesta con respecto a la convencional, debe tenerse en cuenta que “monótona” aquí significa “no decreciente” respecto al orden convencional y “no creciente” respecto al orden del dioide.

γ y δ , podemos derivar entonces las siguientes reglas:

$$t\gamma^k \oplus \tau\gamma^k = \text{máx}(t, \tau)\gamma^k, \quad k\delta^t \oplus k\delta^\tau = k\delta^{\text{máx}(t, \tau)}. \quad (5.7a)$$

Consideremos ahora la Figura 5.3-(b). Para counters se tiene

$$y(t) = 2u(t-1) \oplus 1u(t-1) = (2 \oplus 1)u(t-1) = 1u(t-1),$$

lo que se puede deducir de la misma forma que el caso anterior para daters. Para daters, se tiene

$$y(k) = 1u(k-2) \oplus 1u(k-1) = 1(u(k-2) \oplus u(k-1)) = 1u(k-1),$$

que también puede deducirse de manera análoga al caso anterior para counters. Siguiendo el procedimiento de aplicar las transformadas δ y γ se obtienen las nuevas reglas:

$$k\delta^t \oplus \kappa\delta^t = \text{mín}(k, \kappa)\delta^t, \quad t\gamma^k \oplus t\gamma^\kappa = t\gamma^{\text{mín}(k, \kappa)}. \quad (5.7b)$$

Si ahora escribimos los monomios $t\gamma^k$ y $k\delta^t$ como $\gamma^k\delta^t$, las reglas (5.7a) y (5.7b) pueden sintetizarse como:

$$\gamma^k \oplus \gamma^\kappa = \gamma^{\text{mín}(k, \kappa)}, \quad \delta^t \oplus \delta^\tau = \delta^{\text{máx}(t, \tau)}. \quad (5.7c)$$

5.2.4. Cambios de base

La Figura 5.4 representa un GET antes y después del disparo de la transición x_1 o ξ_1 . Las siguientes ecuaciones son las obtenidas para daters antes y después del

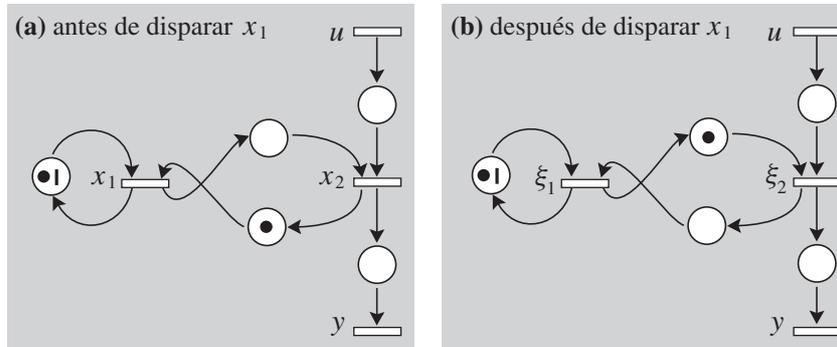


Figura 5.4: Disparo de una transición

disparo.

Antes del disparo

$$\begin{aligned} x_1(k) &= 1x_1(k-1) \oplus x_2(k-1), \\ x_2(k) &= x_1(k) \oplus u(k), \\ y(k) &= x_2(k), \end{aligned}$$

Después del disparo

$$\begin{aligned}\xi_1(k) &= 1\xi_1(k-1) \oplus \xi_2(k) , \\ \xi_2(k) &= \xi_1(k-1) \oplus u(k) , \\ y(k) &= \xi_2(k) .\end{aligned}$$

Mediante algunas sustituciones, se llega a las siguientes formas canónicas:

Antes del disparo

$$\begin{aligned}\begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} &= \begin{pmatrix} 1 & e \\ 1 & e \end{pmatrix} \begin{pmatrix} x_1(k-1) \\ x_2(k-1) \end{pmatrix} \oplus \begin{pmatrix} \varepsilon \\ e \end{pmatrix} u(k) , \\ y(k) &= (\varepsilon \quad e) \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} ,\end{aligned}$$

Después del disparo

$$\begin{aligned}\begin{pmatrix} \xi_1(k) \\ \xi_2(k) \end{pmatrix} &= \begin{pmatrix} 1 & \varepsilon \\ e & \varepsilon \end{pmatrix} \begin{pmatrix} \xi_1(k-1) \\ \xi_2(k-1) \end{pmatrix} \oplus \begin{pmatrix} e \\ e \end{pmatrix} u(k) , \\ y(k) &= (\varepsilon \quad e) \begin{pmatrix} \xi_1(k) \\ \xi_2(k) \end{pmatrix} .\end{aligned}$$

Estas ecuaciones en el espacio de estados son dos realizaciones de la misma función transferencia γ . Puede demostrarse que esta función transferencia es, tras aplicar todas las simplificaciones posibles:

$$e \oplus \gamma(1\gamma)^* .$$

En notación matricial se tiene:

$$\begin{aligned}x(k) &= Ax(k-1) \oplus Bu(k) , & y(k) &= Cx(k) , \\ \xi(k) &= \bar{A}\xi(k-1) \oplus \bar{B}u(k) , & y(k) &= \bar{C}\xi(k) .\end{aligned}$$

Sin embargo, no es posible encontrar una transformación matricial para pasar de una realización a la otra. Este paso requeriría la existencia de una matriz invertible T de 2×2 tal que $x = T\xi$, lo que implicaría por ejemplo que

$$B = T\bar{B} , \quad \text{i.e.} \quad \begin{pmatrix} \varepsilon \\ e \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} e \\ e \end{pmatrix} .$$

La primer fila de esta relación matricial implica que $T_{11} \oplus T_{12} = \varepsilon$, de donde $T_{11} = T_{12} = \varepsilon$, lo que evidentemente no es compatible con la condición de que T sea invertible.

Partiendo de la interpretación sobre el funcionamiento del GET (recordando que una transición interna se disparó una vez), o directamente desde las ecuaciones, puede verse que la relación entre ξ y x es simplemente

$$\xi_2(k) = x_2(k) ; \quad \xi_1(k) = x_1(k-1) .$$

Naturalmente, esta transformación no puede realizarse mediante un cambio de base estático y lineal en el espacio de estados.

Debido a que en la representación de counters los coeficientes y retardos resultan “intercambiados”, en este ejemplo, mediante la utilización de counters en lugar de dater sí será posible encontrar un cambio de base lineal entre los sistemas de ecuaciones obtenidos antes y después del disparo. En esta representación, los elementos de las matrices se corresponden con el número de marcas en el marcado inicial. El disparo de una transición interna elimina una marca de cada lugar anterior a la misma, lo que resta 1 a cada elemento de la fila de la matriz A correspondiente a dicha transición³ y a cada elemento de la misma fila de la matriz B . De la misma forma, este disparo agrega una marca en cada lugar posterior a la transición. Desde el punto de vista algebraico, esto implica sumar 1 a cada elemento de la columna correspondiente de A y de C .

Todas estas operaciones pueden ser realizadas en $\mathbb{Z}_{\text{mín}}$ pre-multiplicando y post-multiplicando por cierta matriz apropiada y su inversa respectivamente. Para el ejemplo anterior, la matriz de pre-multiplicación será:

$$\begin{pmatrix} -1 & \varepsilon \\ \varepsilon & e \end{pmatrix}.$$

Observación 5.6. Con el ejemplo tratado, podría concluirse en que el enfoque del counter tiene alguna superioridad sobre el del dater. Sin embargo, se puede analizar una situación dual. Supongamos que se saca una barra de todos los lugares anteriores a una dada transición interna y se agrega una barra a todos los lugares posteriores a la misma. Es simple de observar que esto no producirá ningún cambio en la relación entrada-salida. Este caso, visto desde el enfoque del dater corresponderá a un cambio de bases, mientras que desde el enfoque del counter no ocurrirá lo mismo.

Tras analizar las propiedades de orden, de monotonía, de simplificación y las posibilidades de los cambios de bases, queda evidenciado que los enfoques vistos hasta aquí cuentan esencialmente con las mismas ventajas y limitaciones. Según el caso considerado, uno u otro puede resultar el más conveniente.

En la siguiente sección, como ya fue anticipado, se presentará una representación bidimensional en la cual, informalmente hablando, los monomios tales como $t\gamma^k$ en la transformada γ y $k\delta^l$ en la transformada δ serán representados por monomios de la forma $\gamma^k\delta^l$; los objetos básicos serán entonces series de potencias en (γ, δ) con coeficientes Booleanos y además de la suma y producto convencionales de series se podrá utilizar la regla (5.7c).

5.3. Construcción del álgebra $\mathcal{M}_{\text{in}}^{\text{áx}}[[\gamma, \delta]]$

Hay muchas formas de construir el álgebra $\mathcal{M}_{\text{in}}^{\text{áx}}[[\gamma, \delta]]$. Aquí esbozaremos alguno de estos caminos de una manera relativamente informal.

³Obviamente la resta es en álgebra convencional.

5.3.1. Observaciones previas**5.3.1.1. Un isomorfismo entre subconjuntos y series de potencias**

En el ejemplo 2.8, se consideraron subconjuntos de \mathbb{R}^2 con \cup como \oplus y $+$ (suma vectorial) como \otimes . La misma idea se puede extender a \mathbb{Z} e incluso a $\overline{\mathbb{Z}}$. Para un conjunto S de \mathbb{Z}^2 , o sea, una colección de “pixels” con coordenadas (k_i, t_i) , con i tomando valores en un conjunto numerable I_S (I_S puede ser infinito), se asocia la siguiente serie formal (Laurent) de potencias

$$p(S) \stackrel{\text{def}}{=} \bigoplus_{i \in I_S} \gamma^{k_i} \delta^{t_i}$$

en $\mathbb{B}[\gamma, \delta]$ (o sea, los coeficientes son números Booleanos y hay dos variables formales, γ y δ , con exponentes en \mathbb{Z} , consideradas *conmutativas*: $\gamma\delta = \delta\gamma$).

Como siempre, solamente los monomios con coeficientes no nulos son explicitados. El dioide $\mathbb{B}[\gamma, \delta]$ es por supuesto *completo* y el \wedge de dos elementos corresponde a la \cap de sus correspondientes subconjuntos.

Obsérvese que esta función $S \mapsto p(S)$ además de ser una biyección es también un isomorfismo de $(\mathbb{Z}^2, \cup, +)$ en $\mathbb{B}[\gamma, \delta]$ provisto de la suma y producto usual de las series de potencias formales. De hecho, es simple verificar que:

$$p(S_1 \cup S_2) = p(S_1) \oplus p(S_2), \quad p(S_1 + S_2) = p(S_1) \otimes p(S_2).$$

Las mismas consideraciones pueden realizarse en dimensión 1 tomando series de potencia en una única variable formal γ , respectivamente δ : las series de potencia en γ , respectivamente δ , representan subconjuntos del eje x , respectivamente y . El producto de una serie de potencias en γ por una en δ representa un subconjunto *rectangular* de \mathbb{Z}^2 .

5.3.1.2. Dioide de semirectas infinitas

Multiplicar una serie de potencias por γ (respectivamente δ) conduce a trasladar el correspondiente subconjunto en 1 en la dirección positiva de x (respectivamente y). En el ejemplo 2.9, se vió que $\mathbb{R}_{\text{máx}}$ es isomorfa al subconjunto de $(2^{\mathbb{R}}, \cup, +)$ formado por todas las semirectas que se extienden de $-\infty$ a cualquier número. Consideraciones similares pueden hacerse para $\mathbb{Z}_{\text{máx}}$ o para el dioide completo $\overline{\mathbb{Z}_{\text{máx}}}$. Nótese que una semirecta $S = (-\infty, t]$ del eje y es caracterizada por

$$p(S) = p(S) \oplus \delta^{-1} p(S) \text{ o equivalentemente } p(S) = (\delta^{-1})^* p(S).$$

Luego, $\overline{\mathbb{Z}_{\text{máx}}}$ es isomorfa al subconjunto de $\mathbb{B}[\delta]$ de elementos que son múltiplos de $(\delta^{-1})^*$. Este subconjunto, que será llamado $(\delta^{-1})^* \mathbb{B}[\delta]$ no es un subdioide de $\mathbb{B}[\delta]$ ya que e no pertenece al mismo; pero es un dioide con $(\delta^{-1})^*$ como elemento neutro para la multiplicación. Los elementos de $(\delta^{-1})^* \mathbb{B}[\delta]$ son siempre *absorbentes para la multiplicación* ya que el producto de dos elementos de $\mathbb{B}[\delta]$ pertenecerá a $(\delta^{-1})^* \mathbb{B}[\delta]$ cuando uno de dichos elementos pertenezca a dicho conjunto.

De esta sección, se debe tener principalmente en cuenta el hecho que $\overline{\mathbb{Z}}_{\text{máx}}$ es isomorfa a $(\delta^{-1})^* \mathbb{B}[[\delta]]$.

5.3.2. Filtrado de los daters decrecientes

Consideremos ahora el conjunto de las transformadas γ de daters: se trata de elementos de $\overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$ con *secuencias de coeficientes no decrecientes*. La propiedad “ $k \mapsto d(k)$ no decreciente” se traduce en

$$\forall k \in \mathbb{Z}, \{d(k) \geq d(k-1)\} \iff \{d(k) = d(k) \oplus d(k-1)\} \quad (5.8a)$$

en álgebra máx-plus. Esto a su vez equivale a

$$\{D(\gamma) = D(\gamma) \oplus \gamma D(\gamma)\} \iff \{D(\gamma) = \gamma^* D(\gamma)\}. \quad (5.8b)$$

Luego, las transformadas γ de trayectorias no decrecientes pertenecen a $\gamma^* \overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$.

Como antes se dijo para $(\delta^{-1})^* \mathbb{B}[[\delta]]$, $\gamma^* \overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$ tampoco es un subdioide de $\overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$ ya que no contiene a e ; pero es un dioide con γ^* como elemento unitario y nuevamente es un “multiplicativo ideal” (los elementos son absorbentes multiplicativos).

Hay de todos modos maneras más formales para realizar esto. Para esto pueden considerarse uno de los dos siguientes caminos.

1. Considerar la inyección canónica de $\gamma^* \overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$ en $\overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$, o sea, en términos de daters, la inyección canónica de trayectorias no decrecientes en trayectorias no necesariamente monótonas. Esta función es residuable y dualmente residuable. Si se considera residuación dual, para un dado $G(\gamma) \in \overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$, se busca el menor $D(\gamma) \in \gamma^* \overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$ que sea mayor que $G(\gamma)$. Es más simple pensar al mismo como la menor solución de

$$D(\gamma) = D(\gamma) \oplus \gamma D(\gamma) \oplus G(\gamma).$$

Ejercicio 5.7. Demostrar que la ecuación anterior expresa las dos condiciones: $D(\gamma)$ corresponde a una trayectoria no decreciente y es mayor que $G(\gamma)$. Demostrar que la menor solución es $\gamma^* G(\gamma)$. Mostrar que, en términos de trayectorias

$$d(k) = \bigoplus_{l \leq k} g(l) = \sup_{l \leq k} g(l),$$

lo que corresponde a una especie de “integración” de g en álgebra máx-plus. La Figura 5.5 muestra esta operación.

Por último, considérese lo mismo utilizando residuación en lugar de residuación dual.

2. Se puede también considerar la función $H : G(\gamma) \mapsto \gamma^* G(\gamma)$ que tiene numerosas propiedades:

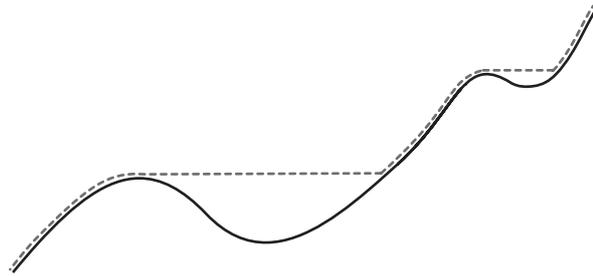


Figura 5.5: Mejor aproximación superior de una trayectoria no monótona por una no decreciente

- es un homomorfismo;
- es s.c.i. (de hecho, es la residual dual de la inyección canónica anterior, por lo que debe ser s.c.i.);
- es una función “cerrada” (ver el Corolario 4.30);
- es residuable si se restringe su rango a su imagen (Teorema 4.27 y Corolario 4.31).

Ambos elementos tienen la misma imagen si tienen la misma “aproximación superior” por una trayectoria no decreciente. Esto define una congruencia (Lema 2.17) y cada clase de equivalencia tiene un máximo representativo (Lema 4.10) que es simplemente $H(G)$ (esto resulta de la Observación 4.19 y del hecho que H^\sharp es la inyección canónica de acuerdo al Teorema 4.27). Todos estos resultados pueden demostrarse directamente sin necesidad de utilizar teoremas más generales.

En síntesis, multiplicar un elemento $G \in \overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$ por γ^* implica seleccionar el *único* elemento no decreciente D de la misma clase de equivalencia, que es además el máximo elemento de dicha clase y la “mejor aproximación no decreciente superior” de G . Desde un punto de vista geométrico, esta multiplicación por γ^* implica trazar una semirrecta hacia $+\infty$ desde cada pixel de G como se muestra en la Figura 5.6.

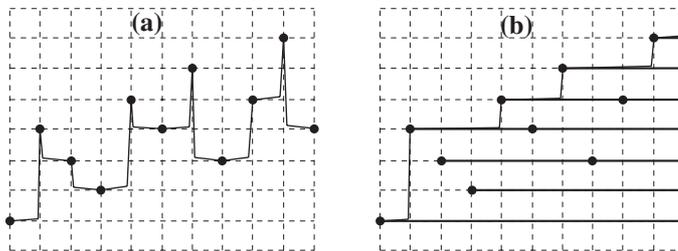


Figura 5.6: Filtrado de una trayectoria no monótona para producir una trayectoria no decreciente

5.3.3. El álgebra $\mathcal{M}_{\text{in}}^{\text{áx}}[[\gamma, \delta]]$

5.3.3.1. Diferentes caminos hacia $\mathcal{M}_{\text{in}}^{\text{áx}}[[\gamma, \delta]]$

Estamos interesados en los elementos de $\gamma^* \overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$; pero vimos que $\overline{\mathbb{Z}}_{\text{máx}}$ es isomorfa a $(\delta^{-1})^* \mathbb{B}[[\delta]]$. Luego, podríamos también considerar $\gamma^*((\delta^{-1})^* \mathbb{B}[[\delta]])[[\gamma]]$, que fácilmente puede demostrarse que es isomorfa a

$$\mathcal{M}_{\text{in}}^{\text{áx}}[[\gamma, \delta]] \stackrel{\text{def}}{=} \gamma^*(\delta^{-1})^* \mathbb{B}[[\gamma, \delta]] .$$

Esta estructura algebraica puede también obtenerse por los siguientes caminos.

1. Se puede comenzar con $\mathbb{B}[[\gamma, \delta]]$, cuyos elementos están asociados con subconjuntos de pixels en \mathbb{Z}^2 (con las operaciones discutidas antes); luego, considerar la imagen de este dioide en la función

$$G(\gamma, \delta) \mapsto \gamma^*(\delta^{-1})^* G(\gamma, \delta) .$$

Esta función es nuevamente un homomorfismo s.c.i. y una función cerrada ya que puede ser escrita (ver Lema 3.2)

$$G(\gamma, \delta) \mapsto (\gamma \oplus \delta^{-1})^* G(\gamma, \delta) . \quad (5.9)$$

Todo lo que se dijo en la sección anterior sobre la imagen de $\overline{\mathbb{Z}}_{\text{máx}}[[\gamma]]$ para $G(\gamma) \mapsto \gamma^* G(\gamma)$ puede repetirse aquí para la imagen de $\mathbb{B}[[\gamma, \delta]]$ para la función (5.9).

Así como multiplicar por γ^* implica trazar una semirrecta horizontal hacia $+\infty$ desde cada pixel, multiplicar por $(\delta^{-1})^*$ implica trazar una semirrecta vertical que va hacia $-\infty$ desde cada pixel (este era el modo de obtener una estructura isomorfa a $\overline{\mathbb{Z}}_{\text{máx}}$). Multiplicar por $(\gamma \oplus \delta^{-1})^* = \gamma^*(\delta^{-1})^*$ implica trazar, desde cada pixel, *un cono sudeste* que es la suma vectorial de las anteriores semirrectas representadas por γ^* y $(\delta^{-1})^*$ (recordar que la suma vectorial se traduce en el producto de series de potencias). Esto se muestra en la Figura 5.7.

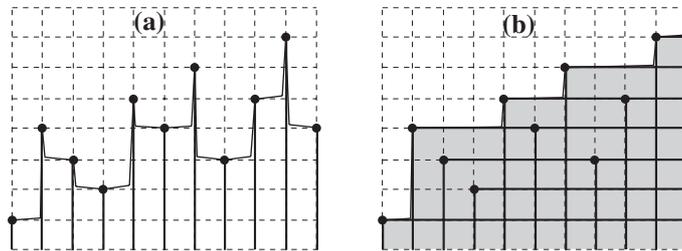


Figura 5.7: Trazado de semirrectas verticales y conos sudeste

2. La tercer manera de obtener $\mathcal{M}_{\text{in}}^{\text{áx}}[[\gamma, \delta]]$ es empezar con counters que tienen trayectorias monótonas representadas en series de potencia de $\overline{\mathbb{Z}}_{\text{mín}}[[\delta]]$.

- Filtrar las trayectorias monótonas implica multiplicar los elementos de este dioide por $(\delta^{-1})^*$, ya que el orden es invertido en $\overline{\mathbb{Z}}_{\min}$ con respecto al orden habitual y luego, en lugar de (5.8), se tiene que

$$\begin{aligned}
 c(t) \geq c(t-1) &\Leftrightarrow c(t-1) = c(t) \oplus c(t-1) \\
 &\Leftrightarrow C(\delta) = \delta^{-1}C(\delta) \oplus C(\delta) \Leftrightarrow C(\delta) = (\delta^{-1})^*C(\delta) . \quad (5.10)
 \end{aligned}$$

- Más aún, $\overline{\mathbb{Z}}_{\min}$ es isomorfa a $\gamma^*\mathbb{B}[\gamma]$. Finalmente, $(\delta^{-1})^*(\gamma^*\mathbb{B}[\gamma])[\delta]$ es isomorfa a $\mathcal{M}_{\min}^{\text{dx}}[\gamma, \delta]$ nuevamente.

Todos estos caminos desde $\mathbb{B}[\gamma, \delta]$ hacia $\mathcal{M}_{\min}^{\text{dx}}[\gamma, \delta]$ se sintetizan en el diagrama de la Figura 5.8.

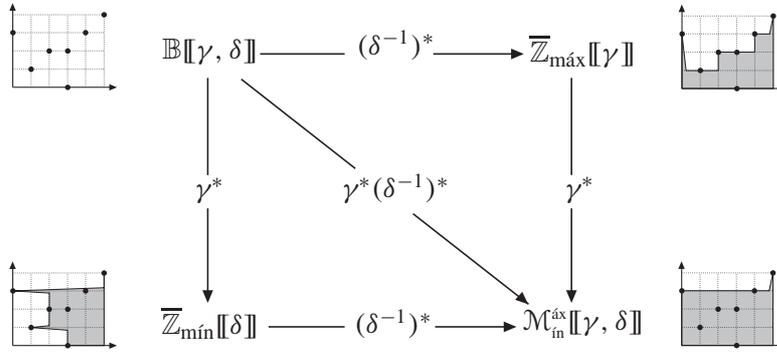


Figura 5.8: Caminos hacia $\mathcal{M}_{\min}^{\text{dx}}[\gamma, \delta]$

5.3.3.2. Punto de vista práctico de $\mathcal{M}_{\min}^{\text{dx}}[\gamma, \delta]$

Desde el punto de vista práctico, se debe tener en cuenta que manipular elementos de $\mathcal{M}_{\min}^{\text{dx}}[\gamma, \delta]$ implica manipular *clases de equivalencia* de $\mathbb{B}[\gamma, \delta]$. Dos elementos están en la misma clase si su producto por $\gamma^*(\delta^{-1})^*$ es formalmente el mismo. Esto implica la existencia de un representativo canónico de cada clase que es también el *máximo* y por lo tanto es el que tiene el máximo número de coeficientes no nulos. Por esta razón, *no es el más económico para manipular en cálculos prácticos*.

Cada clase puede ser representada por cualquiera de sus miembros. El siguiente lema dice que es lo que los distintos miembros de una misma clase tienen en común (esto es por supuesto una condición necesaria; pero no suficiente para pertenecer a la misma clase). Este lema utiliza nociones de *grado* y de *valuación* de elementos de $\mathcal{M}_{\min}^{\text{dx}}[\gamma, \delta]$. Debido a que cualquier elemento puede ser multiplicado por γ^* y/o por $(\delta^{-1})^*$ sin cambiar de clase, es claro que carece de sentido hablar de grado en γ de tales elementos (definido generalmente como el máximo grado en γ de los monomios no nulos en una serie de potencia) o de valuación en δ (definido en

general como el menor grado en δ de los monomios no nulos). Por el contrario, el siguiente lema nos muestra que la valuación en γ y el grado en δ son nociones bien definidas en $\mathcal{M}_m^{\text{áx}}[[\gamma, \delta]]$. Por esto, serán llamadas de aquí en más simplemente valuación y grado de un elemento en $\mathcal{M}_m^{\text{áx}}[[\gamma, \delta]]$.

Lema 5.8. *Para cualquier elemento de $\mathcal{M}_m^{\text{áx}}[[\gamma, \delta]]$, todos los representativos en $\mathbb{B}[[\gamma, \delta]]$ tienen la misma valuación en γ y el mismo grado en δ .*

Es simple realizar una argumentación geométrica. Para una prueba formal, ver [1, Lemma 5.30].

Desde el punto de vista práctico, sería bueno trabajar siempre con el *mínimo* representativo de cada clase, esto es, con el que tiene el número mínimo de monomios con coeficientes no nulos. Sin embargo, aunque muchas de ellas sí, no todas las clases admiten tal representativo mínimo. En las que éste existe, el mismo es obtenido conservando solamente las esquinas Noroeste de cualquier representativo de la clase. Las clases que no admiten representativos mínimos son aquellas que, hablando informalmente,

- contienen elementos con líneas horizontales hacia $-\infty$ (o sea, son mayores que algún $\delta^t (\gamma^{-1})^*$ para cierto t); esta situación no debería tener significado práctico ya que los daters $d(k)$ deberían tomar el valor $-\infty$ antes de algún evento número k_0 que constituye el origen de la numeración de los eventos (en particular, la valuación debe ser finita);
- contienen elementos con líneas verticales hacia $+\infty$ (o sea, son mayores que algún $\gamma^k \delta^*$ para cierto k); esta situación corresponde a daters que toman valor infinito después de cierto número k , o sea, no ocurren eventos después del número k (no hay vivacidad).

El otro problema con los mínimos representativos es que no son estables para la suma o multiplicación (a diferencia de los máximos representativos), o sea, después de una suma o una multiplicación de dos elementos en $\mathcal{M}_m^{\text{áx}}[[\gamma, \delta]]$, aunque esta operación haya sido efectuada con mínimos representativos, el resultado en general no está en la forma mínima. Una información más detallada sobre mínimos representativos se encuentra en [1, Theorem 5.20].

De todos modos, realizar cálculos en $\mathcal{M}_m^{\text{áx}}[[\gamma, \delta]]$ implica manipular elementos de $\mathbb{B}[[\gamma, \delta]]$, aplicando las reglas habituales de suma y multiplicación de series de potencias, *más las reglas de (5.7c)*.

Ejercicio 5.9. Demostrar que ambos lados de las ecuaciones (5.7c) son equivalentes.

Geoméricamente, un pixel en la posición (k, t) domina cualquier pixel situado sobre la misma recta horizontal hacia la derecha y sobre la misma línea vertical debajo de él. Por transitividad, también domina cualquier pixel en el cono sudeste del cual es vértice (este cono puede pensarse como la “sombra” del pixel: cualquier pixel en la sombra de un dado pixel “desaparece”).

Ejercicio 5.10. Verificar nuevamente que e es la misma que γ^* , $(\delta^{-1})^*$, y $\gamma^*(\delta^{-1})^*$. La primer equivalencia da una justificación de lo que se hizo considerando e como equivalente a γ^* y llamándolo impulso.

Ejercicio 5.11. Verificar las simplificaciones que se usaron en §3.2.4,

- $4\gamma \leq 5$ o bien $\gamma\delta^4 \leq \delta^5$;
- $4\gamma \leq 8$ o bien $\gamma\delta^4 \leq \delta^8$;
- $1\gamma^2 \leq 5\gamma$ o bien $\gamma^2\delta \leq \gamma\delta^5$.

Ejercicio 5.12. Recomenzar los cálculos de §3.2.4 y del Ejemplo 3.3 utilizando solamente la notación $\mathcal{M}_{\text{in}}^{\text{ax}}[\gamma, \delta]$ (desde el principio, cuando las ecuaciones del sistemas fueron obtenidas) y utilizar todas las reglas de simplificación correspondientes para alcanzar el resultado esperado.

5.3.3.3. Interpretación informal de $\mathcal{M}_{\text{in}}^{\text{ax}}[\gamma, \delta]$

Otra mirada intuitiva al álgebra $\mathcal{M}_{\text{in}}^{\text{ax}}[\gamma, \delta]$ es la siguiente. Un “pixel” (k, t) brinda una “pieza de información” sobre eventos, es decir,

el evento número k no ocurre antes que en tiempo t .

Tales piezas de información, ubicadas en las transiciones de entrada, “circulan” en el GET y son “desplazadas” por los arcos que siguen: por ejemplo, un arco (o sea un lugar), con κ marcas iniciales y un tiempo de espera de τ introduce un desplazamiento de κ en el dominio de los eventos y de τ en el dominio temporal. Algebraicamente, corresponde a multiplicar los monomios $\gamma^k\delta^t$ (representando la pieza de información) por el operador de desplazamiento $\gamma^\kappa\delta^\tau$.

En las transiciones, las piezas de información son “recolectadas” o “sumadas” (traen información sobre el disparo de la transición en el cual son recolectadas). En este proceso las reglas de simplificación de (5.7c) pueden ser interpretadas por las siguientes consideraciones: si se recolectan dos piezas de información sobre el mismo tipo de evento (disparo de una transición), una con el par (k, t) , la otra con el par (k', t') tal que

$$k' \geq k, \quad t' \leq t,$$

o sea, el pixel (k', t') está en la “sombra” de (k, t) , entonces la pieza anterior puede descartarse ya que tiene menos información que la otra con respecto al mismo tipo de evento. Esto se ve claramente en la interpretación informal hecha en la primer frase de esta subsección sobre el significado del pixel en (k, t) .

Por último, al final de este proceso en el cual las piezas de información fluyen desde las entradas hacia las salidas siendo desplazadas por los lugares y recolectadas y simplificadas (o comprimidas) por las transiciones, se tiene una colección de piezas de información en cada transición de salida (que se codifica en un elemento de $\mathcal{M}_{\text{in}}^{\text{ax}}[\gamma, \delta]$). Lo que ocurre en las salidas se supone que es la secuencia de eventos *más temprana* compatible con tales piezas de información.

Capítulo 6

Comportamiento asintótico de grafos de eventos temporizados autónomos y optimización de recursos

Los objetivos de este capítulo son

- repasar los resultados fundamentales sobre el comportamiento de sistemas “autónomos”, es decir, sistemas para los cuales las entradas están definidas por reglas de feedback, con lo cual pueden evolucionar de manera autónoma;
- demostrar el impacto práctico de estos resultados en problemas de diseño como optimización de recursos.

Los resultados matemáticos son sólo enunciados (las demostraciones pueden encontrarse en [1]) y son ilustrados con ejemplos simples pertenecientes al área de la manufactura.

6.1. Motivación: modelización de Sistemas de Manufactura Flexible (SMF)

6.1.1. Ejemplo de jobshop

Los “workshops” involucran máquinas y partes. Hacemos la distinción entre “flowshops” y “jobshops”:

flowshop: cada tipo de parte visita las máquinas en el mismo orden; pero algunos tipos de partes pueden saltar algunas máquinas;

jobshop: cada tipo de parte visita las máquinas en su propio orden (algunas máquinas pueden ser saltadas).

Consideremos un jobshop con 3 máquinas M_1, M_2, M_3 produciendo 3 tipos de partes P_1, P_2, P_3 . Las siguientes son las secuencias seguidas para la producción de cada parte:

$$\begin{aligned}P_1 &: M_1 \rightarrow M_2 \rightarrow M_3 , \\P_2 &: M_3 \rightarrow M_2 , \\P_3 &: M_1 \rightarrow M_3 .\end{aligned}$$

Más aún, debe respetarse una relación de producción de 1/1/2 (o sea, un 25 % de P_1 y P_2 debe producirse al mismo tiempo que un 50 % de P_3). Una forma de forzar esta relación es repetir la secuencia P_1, P_2, P_3, P_3 . El orden en el cual las partes de esta secuencia básica son tomadas por las máquinas se denomina *scheduling*. La Figura 6.1 corresponde al siguiente scheduling:

$$\begin{aligned} M_1 &: P_1 \rightarrow P_3 \rightarrow P_3 , \\ M_2 &: P_1 \rightarrow P_2 , \\ M_3 &: P_1 \rightarrow P_2 \rightarrow P_3 \rightarrow P_3 . \end{aligned}$$

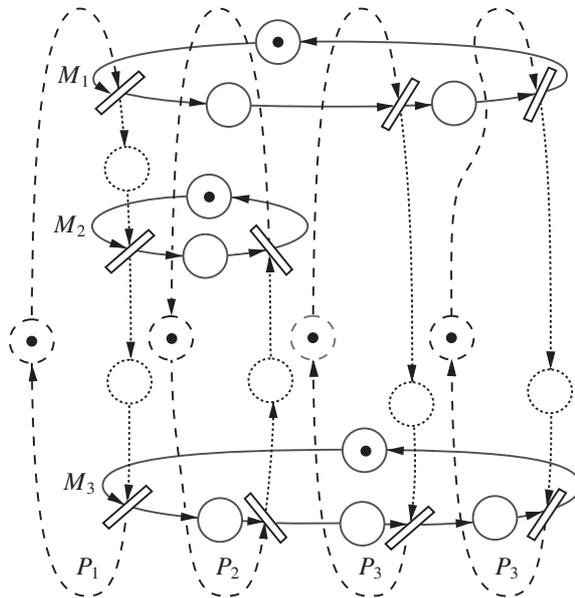


Figura 6.1: Un jobshop con 3 máquinas y 3 tipos de partes en proporción 1/1/2

6.1.2. Ecuaciones de lazo cerrado (con daters), sistema autónomo

Como se vio en los capítulos previos, usando daters, un grafo de eventos temporizado puede describirse por medio de las siguientes ecuaciones generales en el álgebra máx-plus:

$$\begin{aligned} x(k) &= A_0x(k) \oplus A_1x(k - 1) \oplus \dots \oplus B_0u(k) \oplus B_1u(k - 1) \oplus \dots , \\ y(k) &= C_0x(k) \oplus C_1x(k - 1) \oplus \dots \oplus D_0u(k) \oplus D_1u(k - 1) \oplus \dots , \end{aligned}$$

y si además, las entradas están determinadas por las salidas siguiendo las reglas de feedback:

$$u(k) = K_0y(k) \oplus K_1y(k - 1) \oplus \dots , \tag{6.1}$$

las ecuaciones resultantes son:

$$x(k) = (A_0 \oplus B_0 K_0 C_0)x(k) \oplus (A_1 \oplus B_1 K_1 C_1)x(k-1) \oplus \dots \quad (6.2)$$

En §3.1.5–3.1.6, fue mostrado cómo estas ecuaciones pueden llevarse a la forma canónica:

$$\xi(k) = A\xi(k-1) \quad (6.3)$$

para algún vector de estado ξ y una matriz A (luego la llamaremos A por simplicidad).

Observación 6.1. En el ejemplo presentado en la Figura 6.1, los arcos en feedback tienen todos una única marca; por lo tanto sólo la matriz K_1 en (6.1) es distinta de cero, y las componentes de esta matriz (diagonal) son iguales a los correspondientes tiempos ubicados en los lugares de los arcos en feedback que unen transiciones de entrada y de salida. La interpretación física es la siguiente:

- la cantidad de marcas en los lugares en feedback puede interpretarse como la cantidad de paletas (que acarrea cada tipo de parte) disponibles en la entrada del workshop en el instante inicial (en el instante inicial puede haber otras paletas ya en proceso en el sistema, el número total de paletas para un tipo determinado de parte es igual al número de marcas que se encuentran en el circuito, que se obtiene siguiendo el recorrido de esa parte desde el comienzo hasta el final del workshop y de regreso hasta el comienzo a través de los arcos en feedback — recordemos que en cualquier circuito de un GET, el número de marcas es constante).
- los tiempos colocados en los lugares de los arcos en feedback pueden interpretarse como el tiempo total necesario para quitar una parte terminada de la paleta correspondiente, que esta paleta regrese al comienzo del workshop y que una nueva parte sin procesar sea ubicada en la misma paleta.

6.2. Comportamiento a lazo cerrado (o autónomo)

6.2.1. Interpretación de un autovalor

Supongamos que existe z (vector columna no nulo) y λ (escalar), tales que, en $\mathbb{R}_{\text{máx}}$ (pero no en $\overline{\mathbb{R}_{\text{máx}}}$!),

$$Az = \lambda z .$$

Si $x(0) = z$, luego $x(k) = A^k x(0) = \lambda^k x(0) = \lambda x(k-1)$. Esto es un régimen “periódico” o “regular”: para cada tipo de evento, dos eventos sucesivos están separados por un tiempo λ que es la *inversa del “throughput”*. El “throughput” está dado por el número de partes de un determinado tipo procesadas por unidad de tiempo.

6.2.2. Preguntas

Se discutirán ahora las siguientes cuestiones:

1. ¿ Existencia del par (λ, z) ?
2. ¿ Unicidad de λ (o del throughput) y de z (forma del “régimen periódico”) ?
3. ¿ Qué ocurre si se comienza desde algún $x(0)$ diferente de un autovector ?

6.2.3. El caso del autovalor nulo

Existe $z \neq \varepsilon$ tal que $Az = \varepsilon$ (esto es, $\lambda = \varepsilon$) si y sólo si el GET asociado con (respectivamente el grafo de precedencia de) A tiene transiciones de salida (respectivamente, nodos que son sumideros, o sea, sin sucesores). Entonces los elementos no nulos de z corresponden a esas transiciones (respectivamente, nodos).

De aquí en más asumiremos que esta situación no se presenta.

6.3. El caso irreducible

6.3.1. Unicidad del autovalor

Recordemos el Teorema 3.23 en [1].

Teorema 6.2. *Si la matriz A es irreducible, es decir si el grafo de precedencia asociado a A es fuertemente conexo, existe un único autovalor (pero posiblemente varios autovectores). Este autovalor es igual al “máximo ciclo medio” del grafo, es decir, el valor máximo de las siguientes cantidades (llamadas ciclo medio) calculadas para todos los circuitos del grafo: “tiempo total a lo largo del circuito dividido por la longitud del circuito (número de arcos)”.*

6.3.2. Autovectores

En el caso irreducible, puede demostrarse que ningún componente de los autovectores puede ser ε . Los autovectores no son necesariamente únicos (ver [1, Ejemplo 3.25]). Un resultado más preciso se obtiene mirando la estructura del *grafo crítico* (ver [1, Definición 3.94]). Esencialmente, el grafo crítico está compuesto por todos los circuitos críticos y un circuito crítico es aquel cuyo ciclo medio es maximal.

Para una descripción precisa del autoespacio, ver Teorema 3.101 en [1].

Nota. Puede siempre considerarse que el autovalor es e reemplazando la matriz A por $B = \lambda^{-1}A$.

6.3.3. Comportamientos transitorio y asintótico

Si no se comienza desde un estado $x(0)$ que es igual a un autovector, ¿qué es lo que pasa?. En cualquier caso, tenemos $x(k) = A^k x(0)$. O sea que el comportamiento asintótico de $x(k)$ puede ser más bien explicado a través del comportamiento asintótico de las potencias sucesivas A^k de A .

Hay resultados generales que relacionan el comportamiento asintótico de A^k con el llamado *proyector espectral* (ver [1, §3.7.3–3.7.4]). Nos limitaremos aquí a resultados más simples sobre la *ciclicidad* de A .

Para una definición de *ciclicidad asintótica* y *ciclicidad*, ver [1, Definición 3.111]. Esencialmente, una matriz A se dice d -cíclica si existe un entero K tal que

$$\forall k \geq K, \quad A^{k+d} = \lambda^d A^k.$$

El entero K es la longitud del comportamiento transitorio.

Para el resultado principal, ver [1, Teorema 3.112]. Esencialmente, en el caso irreducible, A es siempre d -cíclica, y se conoce de qué forma la ciclicidad d depende de la estructura del grafo crítico.

6.3.4. Ejemplo de un caso patológico

Puede ocurrir un comportamiento transitorio arbitrariamente largo. Esto ocurre cuando existe un circuito que es *casi* crítico. Por ejemplo, con

$$\begin{pmatrix} -\eta & -1 \\ e & e \end{pmatrix},$$

donde η es un número positivo *pequeño*, se tiene que

$$\begin{pmatrix} -\eta & -1 \\ e & e \end{pmatrix}^k = \begin{pmatrix} \max(-k\eta, -1) & -1 \\ e & e \end{pmatrix}.$$

El régimen “estacionario” se alcanza cuando $-k\eta \leq -1$, o sea, $k \geq 1/\eta$ que es un número grande con η cercano a $0 = e$.

6.4. El caso reducible

6.4.1. Repaso

Para una matriz triangular en bloques con sólo *dos* bloques, se debe resolver

$$\begin{aligned} A_{11}z_1 &= \lambda z_1, \\ A_{21}z_1 \oplus A_{22}z_2 &= \lambda z_2. \end{aligned}$$

Sea $\tilde{\lambda}_1$, respectivamente $\tilde{\lambda}_2$, el único autovalor de A_{11} , respectivamente A_{22} , y \tilde{z}_1 , respectivamente \tilde{z}_2 , los autovectores asociados.

- $\tilde{\lambda}_2$ es siempre autovalor de A (usar el autovector $(\varepsilon \quad \tilde{z}_2)^\top$).
- $\tilde{\lambda}_1$ es un autovalor de A si $\tilde{\lambda}_1 \geq \tilde{\lambda}_2$: de hecho, la primera ecuación es satisfecha con \tilde{z}_1 ; resolviendo la segunda ecuación en z_2 se tiene

$$z_2 = (\tilde{\lambda}_1^{-1} A_{22})^* (\tilde{\lambda}_1^{-1} A_{21} \tilde{z}_1) ,$$

que es finita bajo la condición dada.

Intuitivamente,

- si el sistema anterior es más rápido (o sea, $\tilde{\lambda}_1 < \tilde{\lambda}_2$), el sistema posterior va a imponer su velocidad al sistema completo y $\tilde{\lambda}_2$ es el único autovalor del sistema completo;
- en la situación opuesta, el sistema anterior va a imponer su velocidad al sistema posterior (o sea, $\tilde{\lambda}_1$ es un autovalor); pero $\tilde{\lambda}_2$ puede también ser observada con $z_1 = \varepsilon$, lo que implica que el sistema anterior produjo previamente un número infinito de marcas.

El caso general es de complejidad combinatoria; pero esencialmente la idea básica es la misma que antes (ver la tesis de Gaubert [8]).

6.4.2. Aplicación: respuesta en frecuencia de GET

En la teoría clásica de sistemas, se considera una entrada igual a una función seno con una frecuencia ω como input de un sistema lineal estable: después de un período transitorio, la salida del sistema es también una función seno con la misma frecuencia ω ; pero amplificada y trasladada. En el marco de GET, el análogo de una entrada con frecuencia fija puede producirse como la salida de un GET simple como el lazo dibujado en negro en Figura 6.2: este lazo envía 2 marcas cada 3

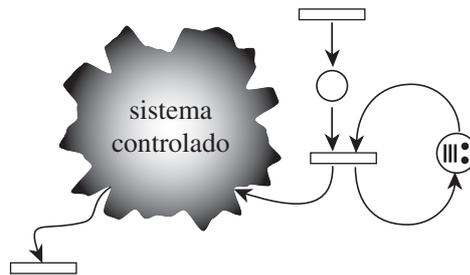


Figura 6.2: Un GET con un input a velocidad constante

unidades de tiempo, siempre y cuando sea alimentado con una infinidad de marcas en su entrada. En el marco de los daters, la frecuencia correspondiente es $3/2$ que es en verdad el tiempo promedio entre 2 eventos o simplemente el autovalor del sistema.

Sin entrar en detalles, la representación dater del sistema completo en Figura 6.2 tiene la forma triangular por bloques considerada en §6.4.1. De esto y de los resultados de §6.4.1, puede esperarse que:

- si el autovalor del sistema posterior (downstream) es menor que la “frecuencia” $3/2$ (el sistema posterior es más rápido que su frecuencia), entonces la salida del sistema completo tendrá la frecuencia $3/2$;
- si el autovalor del sistema posterior es mayor que la frecuencia $3/2$ (el sistema posterior es más lento que su frecuencia) entonces la “frecuencia” del sistema posterior será observada a la salida.

Pueden encontrarse resultados más detallados sobre frecuencias de respuestas de GET en [1, §5.8].

6.5. Autovalores generalizados

En la forma canónica (6.3), cada lugar (o arco) tiene exactamente una marca en el marcado inicial, de tal forma que el número de arcos del circuito es el mismo que el número de marcas. En la forma general de (6.2), los lugares tienen $0, 1, 2 \dots$ marcas.

Consideremos

$$x(k) = A_0 x(k) \oplus A_1 x(k-1) \oplus \dots \oplus A_n x(k-n), \quad (6.4a)$$

y supongamos que A_0 no contiene circuitos. La forma explícita es

$$x(k) = A_0^* (A_1 x(k-1) \oplus \dots \oplus A_n x(k-n)). \quad (6.4b)$$

La forma canónica es

$$\xi(k) = \mathcal{A} \xi(k-1), \quad (6.4c)$$

con

$$\xi(k) = \begin{pmatrix} x(k-n+1) \\ x(k-n+2) \\ \vdots \\ x(k) \end{pmatrix} \text{ y } \mathcal{A} = \begin{pmatrix} \varepsilon & e & \varepsilon & \dots & \varepsilon \\ \varepsilon & \varepsilon & e & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \vdots & \vdots & \vdots & & e \\ A_0^* A_n & \dots & \dots & \dots & A_0^* A_1 \end{pmatrix}.$$

Con la colección de matrices (A_0, A_1, \dots, A_n) , asociamos la matriz de polinomios

$$\mathbf{A}(\gamma) \stackrel{\text{def}}{=} \bigoplus_{i=0}^n \gamma^i A_i.$$

Para una matriz de polinomios $\mathbf{A}(\gamma)$, un *autovalor generalizado* es un número $\lambda \neq \varepsilon$ tal que existe un vector z no trivial que satisface

$$z = \mathbf{A}(\lambda^{-1})z .$$

Aquí, el valor numérico λ^{-1} (igual a $-\lambda$ en notación usual) es sustituido por la variable formal γ . En otras palabras, z es un autovector común asociado con la matriz numérica $\mathbf{A}(\lambda^{-1})$ y el autovalor e .

Esto es entonces equivalente a decir que λ es un autovalor generalizado de la matriz de polinomios asociada con una de las formas (6.4) (nótese que la última lleva a la definición estándar de un autovalor de \mathcal{A}).

Finalmente, el autovalor generalizado es el *máximo cociente del tiempo de espera total a lo largo del circuito dividido el número total de marcas en el mismo circuito, sobre todos los circuitos del grafo*.

6.6. Algoritmos para calcular autovalores y/o autovectores

Algoritmo de potencias: básicamente, calcular las sucesivas potencias de A (debe haber ciclicidad) y, para autovectores, calcular $(\lambda^{-1}A)^+$ (que es lo mismo que $(\lambda^{-1}A)^*$ para columnas relevantes). Un buen truco para calcular B^* es calcular los *cuadrados* sucesivos de $e \oplus B$. Además, como $B = \lambda^{-1}A$ tiene circuitos con pesos negativos, excepto los críticos que tienen peso $e = 0$, si B es $n \times n$, entonces B^k para $k \geq n$ no puede contribuir a B^* (o sea, se puede terminar con B^{n-1}).

Expansión de la matriz transferencia: la respuesta al impulso está compuesta por los sucesivos “coeficientes de Markov” $CA^k B$.

Algoritmo de Karp: sólo para autovalores. Se basa en el teorema de Karp ([1, Teorema 2.19]).

Algoritmo de Howard: Una adaptación de un algoritmo inicialmente diseñado para resolver ecuaciones de programación dinámica.

6.7. “Periodicidad” de la respuesta al impulso, racionalidad, realizabilidad

Ver Cap. 5, §2.4 de [10].

6.8. Estabilizabilidad y optimización de recursos

6.8.1. Estabilidad y control feedback

En este párrafo, se discute brevemente los siguientes resultados: un sistema que es *estructuralmente controlable y observable* puede ser *estabilizado* mediante

feedback *sin alteraciones de performance*, lo cual se asemeja a los resultados existentes en la teoría clásica de sistemas. Se definen, en primer lugar, los siguientes conceptos.

Estabilidad: un GET autónomo se dice *estable* si todas las transiciones operan en igual proporción, es decir:

$$\lim_{k \rightarrow +\infty} \frac{x_i(k)}{k}$$

tiene el mismo valor para todas las transiciones x_i . Esto garantiza que las marcas son producidas en todos lados a la misma velocidad con lo cual no se acumularán entre ningún par de transiciones.

Controlabilidad estructural: un GET se dice *estructuralmente controlable* si cualquier transición interna puede ser alcanzada por un camino proveniente de alguna transición de entrada.

Observabilidad estructural: Un GET es *estructuralmente observable* si cualquier transición interna es el origen de un camino que llega a alguna transición de salida.

Un GET puede estabilizarse fácilmente haciendo que su grafo sea fuertemente conexo. De acuerdo con §6.3, el sistema tendrá un único autovalor y todas las transiciones operarán, asintóticamente, a la misma velocidad. Es fácil ver que un sistema estructuralmente controlable y observable, y por lo tanto estable, puede hacerse fuertemente conexo conectando transiciones de entrada y salida apropiadas mediante arcos en feedback.

Sin embargo, este procedimiento puede crear nuevos circuitos que aumenten el autovalor del sistema en lazo cerrado con respecto a los autovalores del sistema en lazo abierto. Por lo tanto, los circuitos en feedback pueden lentificar el sistema con respecto a su velocidad “natural”, y la estabilidad se obtendría al costo de un deterioro del rendimiento. Afortunadamente esta desventaja puede evitarse de la siguiente manera: supongamos que un nuevo circuito ha sido creado mediante arcos en feedback resultando crítico; agregando suficientes marcas en los arcos en feedback (lugares), este circuito dejará de ser crítico (ya que el denominador del cociente mencionado al final de §6.5 aumentará), y, por lo tanto, se vuelve a la situación en la que sólo los circuitos originalmente en el sistema en lazo abierto pueden llegar a ser críticos. De esta manera se ha recuperado el rendimiento del sistema original con el agregado de la estabilidad.

Claramente, agregar marcas significa agregar recursos con el consecuente aumento del costo. De aquí surge la consigna de agregar *la mínima cantidad posible de marcas*. Esto se discutirá en la sección siguiente.

6.8.2. Aplicación a SMF

Objetivo: obtener la máxima velocidad posible, la de *cuello de botella*,¹ con el mínimo número de recursos.

- En un *flowshop*, la máxima velocidad se logra agregando suficientes “paletas”² de cada tipo, *independientemente del scheduling de partes en las máquinas*. De hecho, en un *flowshop*, si el circuito crítico no es horizontal, existen al menos un arco vertical descendente y uno ascendente; pero los únicos arcos verticales ascendentes son arcos en feedback, y agregar marcas en dichos arcos corresponde a aumentar el número de paletas.
- En un *jobshop*, esto es cierto si uno puede agregar “paletas” en alguna posición inicial apropiada en la línea de producción (aquí con *partes parcialmente manufacturadas*). De hecho, en el caso de un *jobshop*, un circuito crítico puede tener arcos verticales ascendentes y descendentes en la parte interna del grafo ya que las marcas circulan en el workshop en cualquier dirección. Por lo tanto, para eliminar los circuitos críticos no horizontales es necesario ser capaz de agregar marcas en los arcos verticales internos que correspondan a piezas a ser procesadas.
- Luego, hay un desacoplamiento completo entre *performance óptima* y *scheduling óptimo*. El rol de *scheduling óptimo* es minimizar el número de “paletas” requerido para alcanzar *performance óptima*. Esto puede verse como un problema de *programación lineal entera*; pero también pueden imaginarse buenas soluciones heurísticas. Por ejemplo, se puede comenzar con el mínimo número de paletas, determinar el circuito crítico, si no es puramente horizontal, eliminarlo agregando marcas en algún arco vertical y continuar este proceso hasta que finalmente se obtenga un circuito crítico puramente horizontal.

6.8.3. Ejemplo

El ejemplo ilustrado en la Figura 6.3 es un simple *jobshop* con dos máquinas y dos tipos de piezas. La máquina más lenta es M_1 con un tiempo de ciclo total de 11, mientras que para la M_2 es 9. La pieza P_1 siempre debe pasar por la máquina M_2 antes que por la M_1 , mientras que P_2 debe hacerlo en el orden inverso. Para cada tipo de pieza existe inicialmente una paleta. Las posiciones de las marcas en el mercado inicial determinan el inicio de las operaciones de las máquinas y las piezas.

La diferencia entre las cuatro situaciones (a), (b), (c) y (d) en Figura 6.3 es la prioridad de cada pieza para ser procesada en cada máquina (*scheduling*).

¹es decir, el peso máximo de los circuitos *horizontales* en Figura 6.1, que corresponden a la *máquina más lenta* en el *jobshop*

²Ciclos correspondientes a cada pieza en la Figura 6.1 son los ciclos verticales.

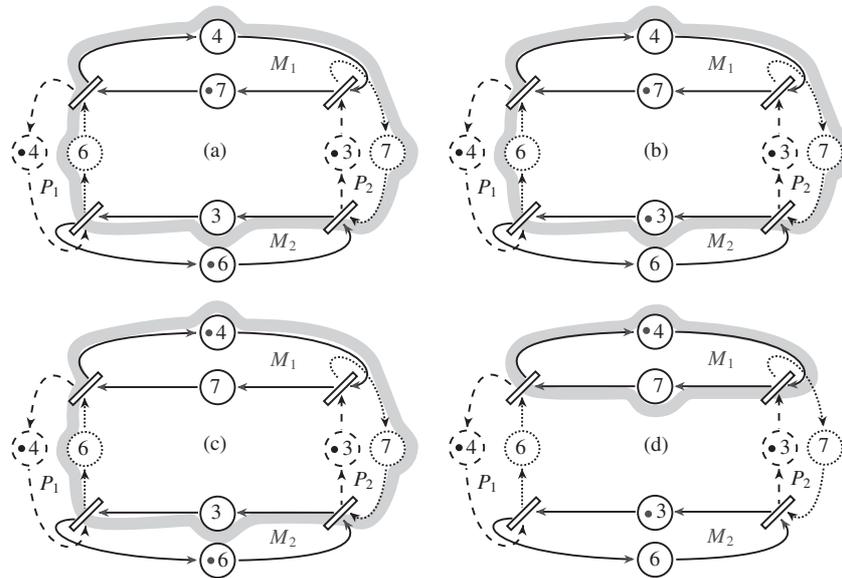


Figura 6.3: Un jobshop con 2 máquinas y 2 partes

- En (a), M_1 debe comenzar con P_1 y M_2 con P_2 ; en este caso, obtenemos un *deadlock* perfecto, materializado por el circuito resaltado en gris en el cual *no* hay marcas y el peso total es positivo: esto corresponde a un autovalor *infinito*. Para evitar este *deadlock*, se puede agregar una marca en el arco vertical que tiene un tiempo de espera de 6: esto corresponde a agregar una paleta de tipo P_1 con una pieza P_1 ya procesada en M_2 ; pero todavía no en M_1 .
- En (b), tanto M_1 como M_2 deben comenzar con P_1 ; en este caso, el autovalor correspondiente al circuito crítico (resaltado en gris) es 20. Para eliminar este circuito, podemos nuevamente agregar una marca en la misma posición como en el caso anterior.
- En (c), M_1 y M_2 deben comenzar con P_2 ; en este caso, el circuito crítico (resaltado en gris) es el mismo que en (b) y puede aplicarse el mismo remedio.
- Finalmente en (d), M_1 debe comenzar con P_2 y M_2 debe comenzar con P_1 ; en este caso, el circuito crítico (resaltado en gris) corresponde a la máquina más lenta M_1 y no se necesitan paletas suplementarias para obtener el mejor rendimiento alcanzable. Este es el único scheduling para el cual se consigue el mejor rendimiento, con sólo dos paletas, una para cada tipo de pieza. Sin embargo, en los otros casos también es posible lograr el mejor rendimiento agregando recursos suficientes.

Epílogo

El Capítulo 4 de este curso, que fue dedicado a la Teoría de Residuación, es sin duda un capítulo denso. Cuando el curso fué dictado en abril de 2000 se presentaron varios ejemplos que mostraron cómo esta importante teoría se utiliza en la solución de algunos problemas. Entre los ejemplos a los que hacemos referencia estuvieron:

- el cálculo de las últimas entradas de un GET descrito por daters que producen salidas no mayores a los valores prefijados de las mismas;
- la síntesis en feedback usando matrices transferencia (ver la tesis de B. Cotentenceau [4]);
- la teoría geométrica de semimódulos (análogos a los espacios vectoriales en el álgebra máx-plus), proyecciones sobre semimódulos, etc., todo lo cual corresponde a trabajos recientes cuyas referencias pueden encontrarse en las siguientes direcciones Web:
 - <http://www-rocq.inria.fr/scilab/quadrat/>
 - <http://www-rocq.inria.fr/scilab/cohen/>
 - <http://amadeus.inria.fr/gaubert/gaubert.html>

Una presentación de algunas herramientas especiales de software (“max-plus toolbox” de Scilab) fue realizada por Elina Mancinelli y Pablo Lotito en ocasión del dictado del curso. Las direcciones de Web arriba mencionadas también contienen información sobre las dichas herramientas.

Bibliografía

- [1] F. Baccelli, G. Cohen, G.J. Olsder, J.-P. Quadrat. *Synchronization and Linearity — An Algebra for Discrete Event Systems*. Wiley, New-York, 1992.
- [2] G. Birkhoff. *Lattice Theory*. Amer. Math. Soc. Coll. Pub., Providence, 1967 (3rd Ed.).
- [3] T.S. Blyth, M.F. Janowitz. *Residuation Theory*. Pergamon Press, Oxford, 1972.
- [4] B. Cottenceau. *Contribution à la commande de systèmes à événements discrets : synthèse de correcteurs pour les graphes d'événements temporisés dans les dioïdes*. Thesis dissertation, ISTIA, Université d'Angers, France, 1999.
- [5] F. Di Cesare, G. Harlakis, J.M. Proth, M. Silva, F.B. Vernadat. *Practice of Petri Nets in Manufacturing*. Chapman & Hall, London, 1993.
- [6] J. Dresel, J. Esparza. *Free choice Petri Nets*. Cambridge University Press, 1995.
- [7] P. Dubreil, M.L. Dubreil-Jacotin. *Leçons d'Algèbre Moderne*. Dunod, Paris, France, 1964.
- [8] S. Gaubert. *Théorie Linéaire des Systèmes dans les Dioïdes*. Thesis dissertation, École des Mines de Paris, Paris, 1992.
- [9] E. Menguy. *Contribution à la commande des systèmes linéaires dans les dioïdes*. Thesis dissertation, ISTIA, Université d'Angers, France, 1997.
- [10] J.-P. Quadrat. *Semi-Anillos en Matemática Aplicada*. CUADERNO 28, Instituto de Matemática Beppo Levi, Universidad Nacional de Rosario, Argentina, 1999.
- [11] C. Reutenauer and H. Straubing. Inversion of matrices over commutative semiring. *Journal of Algebra*, vol. 88, pp. 350–360, 1984.
- [12] D.E. Rutherford. Inverses of Boolean matrices. *Proceedings of the Glasgow Mathematical Association*, vol. 6, pp. 49–53, 1963.
- [13] R. Valette. *Les Réseaux de Petri*. Cours, Janvier 1999 (<http://www.laas.fr/~robert>).
- [14] J.H. Wedderburn. Boolean linear associative algebra. *Annals of Mathematics*, vol. 35, pp. 185–194, 1934.

Serie: CUADERNOS del Inst. de Matemática “B. Levi”

NUMEROS APARECIDOS:

1. G. STAMPACCHIA, “Introducción a las ecuaciones en derivadas parciales e inecuaciones variacionales”.
2. C. PROCESI, “Teoría de Galois elemental”.
3. U. MOSCO, “Transformada de Fourier y distribuciones”.
4. P. BERNHARD, “Introducción a la teoría de Control Optimo”.
5. P. FAURRE, “Teoría de Control, Filtrado estadístico e Introducción al Control Estocástico”.
6. “Encuentro Latinoamericano sobre Métodos Teóricos y Numéricos de la Física-Matemática”, E. Rofman (Ed.), Rosario, 2–7 de julio de 1973 :
 - U. D’AMBROSIO, “Superficies generalizadas no cálculo das variaciones”, p. 5–24.
 - L.A. MEDEIROS, “Ecuaciones diferenciales en derivadas parciales”, p. 25–58.
 - E. ROFMAN, “Desigualdades variacionales. Existencia y aproximación numérica, de soluciones”, p. 59–86.
7. A. BENSOUSSAN, “Teoría moderna de Control Optimo”.
8. C. BAIOCCHI, “Movimiento de un fluido en medios porosos. Enfoque variacional, cuasi-variacional y numérico”.
9. R. GLOWINSKI, “Introducción a la aproximación de inecuaciones variacionales elípticas”.
10. H. SAMARA, “Resolución numérica de ecuaciones diferenciales. Formulación variacional del método TAU”.
11. “Seminario sobre el Problema de Stefan y sus Aplicaciones”, D.A. Tarzia (Ed.), Rosario, 4–8 de julio de 1983, Volumen I :
 - D.A.TARZIA, “Introducción al Seminario sobre el problema de Stefan y sus aplicaciones”, p. 5–32.
 - D.A.TARZIA, “Problemas unidimensionales de conducción del calor con frontera móvil”, p. 33–61.
 - J.E. BOUILLET, “Algunos resultados sobre ecuaciones de tipo parabólico”, p. 63–74.
 - J-E. BOUILLET, “Soluciones autosemejantes con cambio de fase”, p. 75–104.
 - L.T. VILLA, “El problema inverso de Stefan”, p. 105–131.
 - L.T. VILLA, “Problemas optimales en la conducción del calor con cambio de fase, p. 133–142.
 - R.H. NOCHETTO, “Una introducción general a la resolución numérica del problema de Stefan unidimensional”, p. 143-166.
 - G.G. GARGUICHEVICH - M.C. SANZIEL, “Una introducción general a la resolución aproximada del problema de Stefan unidimensional”, p. 167–177.

12. "Seminario sobre el Problema de Stefan y sus Aplicaciones", D.A. Tarzia (Ed.), Rosario, 4-8 de julio de 1983, Volumen 2 :
 - D.A. TARZIA, "Soluciones exactas del problema de Stefan unidimensional", p. 5-36.
 - D.A. TARZIA, "Estudios teóricos en el problema de Stefan unidimensional a una fase", p. 37-85.
 - P.P. MARANGUNIC, "Algunos conceptos básicos del análisis funcional", p. 87-118.
 - H.R. BERTORELLO, "Sobre la solidificación de aleaciones binaria", p. 119-142.
 - L. CRIVELLI - S. IDELSOHN - C. BAIGORRIA, "Análisis de la fusión en barras de combustible nuclear originada por sobrecargas térmicas accidentales", p. 143-170.
 - E.E. ALANIS, "Problemas de cambio de fase en la acumulación de calor a partir de energía solar", p. 171-177.
 - M.B. STAMPELLA, "Sobre la determinación de coeficientes térmicos desconocidos de un material a través de un problema con cambio de fase", p. 179-193.
13. A. FASANO, "Las zonas pastosas en el problema de Stefan", en "II Seminario sobre el Problema de Stefan y sus Aplicaciones", D.A. Tarzia (Ed.), Rosario, 13-17 de octubre de 1986, Volumen 1.
14. "II Seminario sobre el Problema de Stefan y sus Aplicaciones", D.A. Tarzia (Ed.), Rosario, 13-17 de octubre de 1986, Volumen 2:
 - H. BERTORELLO, "Termodinámica del cambio de fase con aplicación a transformaciones sólido-líquido", p. 9-27.
 - J.E. BOUILLET, "Comparación de soluciones de ecuaciones parabólicas", p. 29-44.
 - D.A. TARZIA, "Estudios teóricos básicos en el problema de Stefan unidimensional a dos fases, p. 45-75.
 - D.A. TARZIA, "El problema de Stefan multidimensional a una fase", p. 77-99.
 - L.T. VILLA, "La ecuación de la difusión y su aplicación a problemas de frontera libre", p. 101-116.
 - N. AGUILERA, "El método de Alt para el problema del dique poroso", p. 117-128.
 - G.G. GARGUICHEVICH, "Las inecuaciones variacionales elípticas", p. 129-138.
 - P.R. MARANGUNIC, "Distribuciones y espacios de Sobolev", p. 139-165.
 - E. ZARANTONELLO, "Teoría espectral cónica, p. 167-169.
15. P.L. GEORGE, "Introducción a la utilización del Software Científico MODULEF".
16. J.L. MENALDI, "Programación dinámica estocástica".
17. "III Seminario sobre Problemas de Frontera Libre y sus Aplicaciones", D.A. Tarzia (Ed.), Rosario, 11-15 de octubre de 1988, Volumen 1:
 - J.E. BOUILLET, "Comportamiento asintótico en ecuaciones de conducción-difusión degeneradas", p. 7-13.
 - C. CORTAZAR - M. ELGUETA, "Sobre localización de soluciones de la ecuación de medios porosos", p. 15-27.
 - G-G. GARGUICHEVICH - D.A. TARZIA, "Sobre un problema de Stefan a dos fases con fuente de energía interna", p. 29-44.
 - M. KORTEN, "Soluciones autosemejantes de ecuaciones de difusión-conducción generalizadas", p. 45-53.

- R. MASCHERONI, "Métodos simplificados para la predicción de tiempos de congelación y descongelación de alimentos", p. 55–71.
- D.A. TARZIA, "Comportamiento asintótico exponencial en la ecuación de medios, porosos con absorción", p. 73–86.
- C.V. TURNER, "Difusión de un solvente en un polímero no homogéneo", p. 87–104.
- K. TWARDOWSKA, "A free boundary value problem in solidification of binary alloys. Deterministic and stochastic approach", p. 105–119.
- L.T. VILLA, "Problemas de frontera libre y móvil en procesos de ingeniería química", p. 121–132.
- D.A. TARZIA, "El caso estacionario de] problema de Stefan a dos fases y problemas relacionados", p. 133–151.
18. M. PRIMICERIO - R. GIANNI, "La filtración en medios porosos", en "III Seminario sobre Problemas de Frontera Libre y sus Aplicaciones", D.A. Tarzia (Ed.), Rosario, 11–15 de octubre de 1988, Volumen 2.
19. J.C. CESCO, "Economía, matemática", en "Reunión Anual de la Unión Matemática Argentina", Rosario, 11–14 de octubre de 1989.
20. L. QUINTAS, "Teoría de juegos", en "Reunión Anual de la Unión Matemática Argentina", Rosario, 11–14 de octubre de 1989.
21. E. GENTILE, "Forma Normal de Jordan", en "Reunión Anual de la Unión Matemática Argentina", Rosario, 11–14 de octubre de 1989.
22. M. UGHI, "On the porous media equation with either source or absorption".
23. "IV seminario sobre Problemas de Frontera Libre y sus Aplicaciones", D.A. Tarzia (Ed.), Rosario, 14–18 de diciembre de 1992, Volumen 1 :
- J.C. REGINATO - D.A. TARZIA, "Free boundary models for root growth of crops", p. 1–26.
- E. PARDO, "Modelado térmico de soldaduras GMA", p. 27–33.
- M.K. KORTEN, "Fronteras libres en soluciones de la ecuación $u_t = \Delta(u - 1)_+$ ", p. 35–37.
- M.C. SANZIEL - D.A. TARZIA, "Problemas con multifases", p. 39–51.
- L.T. VILLA, "Modelos de Frontera Libre-Móvil en sistemas de difusión reacción química gas-sólido", p. 53–71.
- L.R. BERRONE - G.G. GARGUICHEVICH, "Sobre un problema estacionario de Stefan con fuente de energía", p. 73–80.
- J.I. ETCHEVERRY, "Análisis de Modelos de Ablación de Polímeros", p. 81–92.
- V.A. KUZ, "Dynamic study of evaporation of droplets", p. 93–102,
- D.F. DELMASTRO - A. CLAUSSE, "Oscilaciones de frontera de entalpía en flujos de dos fases", p. 103–119.
24. "IV Seminario sobre Problemas de Frontera Libre y sus Aplicaciones" D.A. Tarzia (Ed.), Rosario, 14–18 de diciembre de 1992. Volumen 2:
- L.R. BERRONE, "Rango temporal de validez de modelos que involucran a la ecuación del calor-difusión", p. 1–36.
- C.V. TURNER, "El problema de Stefan unidimensional para el líquido sobreenfriado", p. 37–49.
- R.B. MASCHERONI, "Transferencia de calor con cambio de fase en sistemas multidimensionales: uso de factores de forma en métodos aproximados", p. 51–66.

- M. STORTI, "Numerical modeling of ablation phenomena as two-phase Stefan problem", p. 67–88.
- C.B. LEDERMAN, "Un problema de frontera libre en elasticidad", p. 89–96.
- J.E. BOUILLET, "Fronteras libres en ecuaciones de conducción con coeficientes discontinuos", p. 97–102.
- D.A. TARZIA, "Numérica analysis for the heat flux in a mixed elliptic problem with solution of non-constant sign", p. 103–120.
25. J-C. REGINATO - D.A. TARZIA, "Toma de agua y crecimiento de raíces de cultivos", p. 1–13.
- R-H. MASCHERONI, "Dos problemas de frontera libre acoplados: solidificación con simultánea sublimación", p. 15–26.
- J.I. ETCHEVERRY, "On a numerical method for solving a boundary value problem for $u_t = \alpha(u)_{xx}$ ", p. 27–34.
- D.A. TARZIA - C.V. TURNER, "El problema de Stefan a una fase para el líquido sobreenfriado con una condición de contorno convectiva", p. 35–45.
- L.T. VILLA - O.D. QUIROGA, "Modelos de frontera libre en transformaciones fluido -sólido reactivo", p. 47–62.
- V-A. KUZ, "One-dimension model for dynamic condensation of a vapor", p. 63–73.
- A. BOUCIGUEZ - L. SARAVIA, "Análisis térmico y solución numérica al problema de transferencia de calor en dos fases", p. 75–84.
- D.A. TARZIA - C.V. TURNER, "El comportamiento asintótico para un problema de Stefan a una fase con una condición de contorno convectivo", p. 85–92.
26. C. LEDERMAN - N. WOLANSKI, "Un problema de frontera libre de la teoría de combustión I", p. 1–13.
- C. LEDERMAN - N. WOLANSKI, "Un problema de frontera libre de la teoría de combustión II", p. 15–25.
- A.R. GONZALEZ - D.A. TARZIA, "Determination of thermal coefficients in semi-infinite materials with mushy zone in phase-change process", p. 27–48.
- J.I. ETCHEVERRY, "Modelado numérico de la fusión y evaporación de materiales mediante haces de electrones", p. 49–58.
- A.C. BRIOZZO - D.A. TARZIA, "Solución exacta de un problema de frontera libre de flujos en un medio saturado-no saturado con difusividad variable", p. 59–72.
- P.R. MARANGUNIC, "El problema de Stefan simétrico a dos fases con sobre-enfriamiento", p. 73–88.
- D.A. TARZIA, "Sobre el caso estacionario del problema de Stefan-Signorini con condiciones de contorno mixtas", p. 89–102.
27. A. JEAN-MARIE, "Introducción a la teoría matemática de colas de espera", (en preparación).
28. J.P. QUADRAT, "Semi-anillos en matemática aplicada", 1999.

CUADERNOS
DEL INSTITUTO DE MATEMATICA "BEPPO LEVI"

Páginas tipeadas provistas por autores

UNR EDITORA

Editorial de la Universidad Nacional de Rosario
Urquiza 2050 - (S2000AOB) Rosario - Santa Fe

República Argentina

Octubre 2001

Nota : Para intercambiar esta serie con publicaciones similares, dirigirse a:

**CUADERNOS del Instituto de Matemática ì Beppo Levi,
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario,
Avda. Pellegrini 250,
2000 Rosario, Argentina.**

Note : To exchange this serie with similar publications, please write to:

**CUADERNOS del Instituto de Matemática ì Beppo Levi,
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario,
Avda. Pellegrini 250,
2000 Rosario, Argentina.**