

Temporal Constraint Satisfaction Problems

Roque Marín, Isabel Navarrete

Departamento de Ingeniería de la Información y de las Comunicaciones
Universidad de Murcia.
Campus de Espinardo. 30071. Murcia

e-mail: {roque,inava}@dif.um.es

In this paper we deal with the problem of temporal reasoning as a special case of a constraint satisfaction problem. We review several models, either qualitative or metric, for constraint based temporal reasoning and we show the main features of each one, such as expressivity and computational complexity. We also show some recent advances in the field of temporal reasoning with points and durations as temporal objects and finally we suggest some future works.

Problemas de Satisfacción de Restricciones Temporales

Roque Marín, Isabel Navarrete

Departamento de Ingeniería de la Información y de las Comunicaciones
Universidad de Murcia.
Campus de Espinardo. 30071. MURCIA.
{roque,inava}@dif.um.es

Resumen

En este trabajo se plantea el problema del razonamiento temporal como un caso particular de Problema de Satisfacción de Restricciones (CSP). Siguiendo esta perspectiva, se revisan los principales modelos de razonamiento temporal basados en restricciones, tanto cualitativos como métricos, comentando la potencia expresiva de cada modelo y la complejidad computacional asociada a la resolución de las tareas básicas de razonamiento temporal. Se describen también avances recientes en modelos de restricciones entre instantes y duraciones, orientados a la búsqueda de fragmentos expresivos y tratables. Por último, se analizan las tendencias en este campo y las posibles líneas de evolución futura.

Palabras clave: Razonamiento Temporal, Restricciones Temporales, Modelos de Instantes, Intervalos y Duraciones, CSP.

1. Introducción

El razonamiento temporal juega un papel importante en diversas áreas de la Inteligencia Artificial, tales como procesamiento de lenguaje natural [4], planificación [5, 19], scheduling [33, 2], diagnóstico [31, 32] o minería de datos temporales [7]. La solución usual para dotar a los programas de resolución de problemas con capacidades de gestión de información temporal consiste en construir un *razonador temporal* e integrarlo con el resolutor de problemas. El razonador temporal trata sólo con tareas de gestión e inferencia de tiempos y relaciones temporales, y no precisa ningún conocimiento sobre el dominio de la aplicación; este otro conocimiento será manejado exclusivamente por el resolutor de problemas. Esta idea, basada en una estricta separación entre conocimiento sobre el tiempo y sobre el dominio, se remite a Kahn y Gorry [18] y Allen [3].

El paradigma de los Problemas de Satisfacción de Restricciones (CSP o Constraint Satisfaction Problems) ofrece una forma sencilla e intuitiva para formalizar los modelos en que se basa un razonador temporal. De hecho, un problema de razonamiento temporal puede considerarse como un caso particular de CSP, en el que las variables representan *entidades temporales*, tales como instantes o intervalos, y las restricciones representan las *relaciones temporales* permitidas entre las entidades temporales [35].

Una ventaja del modelado de problemas temporales como CSP es que pueden resolverse mediante la amplia batería de algoritmos existentes para CSP generales. Un CSP temporal contiene la información temporal de partida suministrada sobre el problema; para inferir nueva información temporal sobre los tiempos de ocurrencia de los eventos, se aplican algoritmos basados en la propagación de restricciones, que combinan exhaustivamente la información temporal conocida.

Se han definido numerosos modelos de razonamiento temporal basados en restricciones, que se diferencian unos de otros en las entidades temporales con las que tratan y en el tipo de restricciones definidas sobre las entidades temporales. Las posibilidades en cuanto a las entidades temporales se reducen a *instantes* (también llamados *puntos*), *intervalos* o *duraciones*. En cuanto al tipo de restricciones, un modelo temporal puede manejar relaciones *cualitativas*, *métricas* o ambas. En el primer caso, lo que importa es el orden relativo de los objetos temporales, pero no la distancia temporal cuantitativa existente entre ellos. En algunas aplicaciones, en cambio, interesa mantener información cuantitativa sobre la disposición temporal de los eventos, que se perdería usando un modelo puramente cualitativo. A menudo son interesantes también los modelos híbridos, que permiten combinar información temporal cuantitativa y cualitativa.

Un aspecto importante para las aplicaciones prácticas es la complejidad computacional asociada a los distintos problemas de razonamiento temporal que se pueden plantear en estos modelos. Existen diferencias importantes en la eficiencia de los algoritmos para los distintos modelos. Obviamente, cuanto mayor es la capacidad expresiva de un modelo, mayor es la complejidad computacional asociada.

En las siguientes secciones haremos un breve repaso a algunos de los formalismos de restricciones temporales más representativos. La sección 2 se dedica al estudio de los modelos de restricciones cualitativas entre puntos y entre intervalos. En la sección 3 haremos referencia a modelos métricos. La sección 4 se centrará en los últimos desarrollos en modelos de puntos y duraciones. Cerraremos el artículo con unos comentarios sobre las principales conclusiones a que conduce esta rápida revisión de los modelos de restricciones temporales, y sobre las tendencias actuales en este campo.

2. Restricciones Temporales Cualitativas

2.1. Álgebra de Puntos

Uno de los modelos más sencillos y eficientes para razonamiento temporal con restricciones es el *álgebra de puntos* (PA) de Vilain y Kautz [41]. Dados dos puntos p_i y p_j , se pueden definir tres

relaciones binarias cualitativas disjuntas (relaciones primitivas, atómicas o básicas) entre ellos, que expresan de forma exacta y excluyente la posición relativa de los puntos en un eje temporal denso τ : el punto p_i puede ser anterior ($<$), posterior ($>$) o igual ($=$) al punto p_j . No obstante, la información disponible sobre dos puntos será a veces incompleta o indefinida y se puede expresar mediante *relaciones binarias disyuntivas* que corresponden a un subconjunto del conjunto de relaciones básicas $T = \{<, >, =\}$. Por ejemplo, la disyunción $(p_i < p_j \vee p_i = p_j)$ indica que la relación que satisfacen los puntos es $p_i \{<, =\} p_j$, o simbólicamente $p_i \leq p_j$. La relación \leq es una relación infinita que contiene el conjunto de pares de valores $\{(a_i, a_j) \in \tau^2 \mid a_i \leq a_j\}$.

Formalmente el álgebra de puntos se describe mediante la estructura $(2^T, ^{-1}, \cap, \circ)$ formada por los 8 subconjuntos de T (*PA-relaciones*), y las operaciones internas de inversión ($^{-1}$), intersección (\cap) y composición (\circ) de PA-relaciones. La relación $? \equiv \{<, >, =\}$ es la *relación universal*: contiene todas las parejas de valores $(a_i, a_j) \in \tau^2$. La relación \emptyset es la *relación nula* que no contiene ningún par de elementos del eje temporal, indicando que la información temporal es inconsistente.

Cada relación $R \in 2^T$ tiene asociada una *relación inversa* R^{-1} de forma que $p_i R p_j \Leftrightarrow p_j R^{-1} p_i$. Por ejemplo, la inversa de $<$ es $>$. La *intersección* de dos relaciones $R'_{i,j}, R''_{i,j} \in 2^T$ se obtiene como la intersección de los conjuntos de relaciones básicas que constituyen $R'_{i,j}$ y $R''_{i,j}$. La *composición* de dos PA-relaciones, $R_{i,j}$ entre los puntos p_i, p_j y $R_{j,k}$ entre los puntos p_j, p_k , es una nueva relación $R_{i,k} = R_{i,j} \circ R_{j,k}$ entre los puntos p_i, p_k *inducida* por las relaciones $R_{i,j}$ y $R_{j,k}$, de forma que: $p_i R_{i,k} p_k \Leftrightarrow p_i R_{i,j} p_j \wedge p_j R_{j,k} p_k$ (Tabla I). Por ejemplo, de la información parcial $p_1 < p_2 \wedge p_2 = p_3$ puede deducirse la relación implícita $p_1 < p_3$, que equivale a la composición de la relación $<$ entre p_1, p_2 y la relación $=$ entre p_2, p_3 .

Un caso particular de interés es el *álgebra de puntos convexa* (CPA), con las mismas operaciones del PA, pero donde se excluye la relación \neq . Los conjuntos de pares de valores que definen las CPA-relaciones son conjuntos convexos y esto supone que los algoritmos de razonamiento para este modelo sean más eficientes [40].

Supongamos un conjunto de n puntos cuya posición absoluta sobre el eje temporal es desconocida, aunque sabemos que verifican algunas relaciones binarias cualitativas entre ellos. Podemos

expresar este tipo de conocimiento temporal mediante una *red del álgebra de puntos*, o abreviadamente *PA-red*. Una PA-red puede considerarse como un CSP binario no discreto formado por un conjunto de variables $P = \{p_1, p_2, \dots, p_n\}$ que representan instantes de tiempo (siguiendo la notación habitual para problemas PA, usaremos la letra p para representar las variables temporales de este tipo de CSP, en lugar de usar las últimas letras del alfabeto); el dominio D_i de cada variable es el conjunto de números reales (i.e., el eje temporal τ) y las restricciones binarias $C_{i,j}$ vienen dadas por PA-relaciones $R_{i,j}$ entre los puntos de P . Una PA-red (o CPA-red en el caso del álgebra de puntos convexa) puede representarse mediante un *grafo de restricciones* dirigido y etiquetado, donde los nodos representan los puntos de la red y cada arco $i \rightarrow j$ se etiqueta con el símbolo de la relación binaria entre los puntos p_i y p_j . Esta representación permite adaptar fácilmente las técnicas de resolución de tareas típicas de los CSPs binarios, que en el caso del álgebra de puntos se traducen en algoritmos polinomiales eficientes. La figura 1 muestra un ejemplo sencillo de grafo de restricciones para una CPA-red de tres variables.

\circ	$=$	$<$	\leq	$>$	\geq	\neq	$?$	\emptyset
$=$	$=$	$<$	\leq	$>$	\geq	\neq	$?$	\emptyset
$<$	$<$	$<$	$<$	$?$	$?$	$?$	$?$	\emptyset
\leq	\leq	$<$	\leq	$?$	$?$	$?$	$?$	\emptyset
$>$	$>$	$?$	$?$	$>$	$>$	$?$	$?$	\emptyset
\geq	\geq	$?$	$?$	$>$	\geq	$?$	$?$	\emptyset
\neq	\neq	$?$	$?$	$?$	$?$	$?$	$?$	\emptyset
$?$	$?$	$?$	$?$	$?$	$?$	$?$	$?$	\emptyset
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset

Tabla I. Composición de PA-relaciones

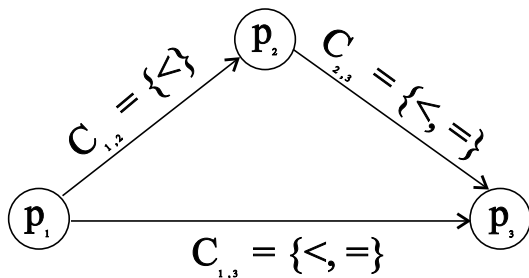


Figura 1. Grafo de restricciones de una CPA-red

Las principales tareas de razonamiento temporal que pueden interesar en aplicaciones del álgebra de puntos son las siguientes:

a) Decidir si una PA-red es **consistente** y en caso afirmativo encontrar una **solución**. El problema de la consistencia puede resolverse en $O(n^3)$ con un algoritmo de camino-consistencia [22]. Van Beek [39] propone un algoritmo más eficiente, llamado CSPAN, que determina la consistencia y encuentra una solución en $O(n^2)$. Se basa en la búsqueda de componentes fuertemente conexas en el grafo de restricciones para obtener un *grafo reducido*, donde los arcos están etiquetados solamente con las relaciones $<$, $>$ ó $?$. Si sobre el grafo reducido se lleva a cabo una ordenación topológica entonces puede obtenerse una asignación consistente para las variables (i.e., una solución) y a partir de ella un **escenario consistente** de la PA-red original.

b) Encontrar la **PA-red mínima**. Si se reducen al mínimo los dominios de las variables y las restricciones de una red, obtenemos la llamada *red mínima*, esencial para responder a consultas sobre relaciones temporales. Vilain y Kautz [41] sugirieron que el algoritmo de clausura deductiva de Allen [3], un caso especial de algoritmo de camino consistencia, es capaz de encontrar la PA-red mínima en $O(n^3)$. Van Beek [40] demostró que este algoritmo no es completo para el álgebra de puntos pero sí lo es para el álgebra de puntos convexa. Por tanto la CPA-red mínima puede obtenerse en $O(n^3)$. Para encontrar la PA-red mínima puede aplicarse un algoritmo de 4-consistencia en $O(n^4)$ [40], aunque Van Beek [39] mejora este resultado y presenta un algoritmo que calcula la PA-red mínima en $O(\max(n^3, Hn^2))$, donde H es el número de arcos de la red etiquetados con la relación \neq .

c) Obtener una **PA-relación mínima**. Para el caso de una CPA-red Van Beek [40] propone una variación del algoritmo de Dijkstra que permite calcular en $O(n^2)$ la relación mínima entre un punto y todos los demás, pero este algoritmo no es exacto para una PA-red. Para minimizar una única PA-relación R se determina la consistencia de las PA-redes resultantes de sustituir en la PA-red original la relación R por cada una de sus relaciones primitivas, obteniendo la PA-relación mínima equivalente a R en $O(n^2)$ [25].

2.2. Álgebra de Intervalos

El álgebra de intervalos (IA), introducida por Allen [3] para representar y gestionar relaciones

cualitativas entre intervalos, es uno de los modelos que más ampliamente se ha usado en la práctica y sobre el que se ha obtenido un gran número de resultados teóricos, en parte por ser el modelo formal basado en restricciones más antiguo.

Allen identificó 13 *relaciones básicas* para describir la posición relativa entre un par de intervalos: $I = \{b, m, o, s, d, f, b^{-1}, m^{-1}, o^{-1}, s^{-1}, d^{-1}, f^{-1}, eq\}$ (Tabla II). Como en el álgebra de puntos, se puede expresar conocimiento incompleto mediante *relaciones disyuntivas*, que se representan por un subconjunto de relaciones básicas. Por ejemplo la relación $x \{b, b^{-1}\} y$ expresa el hecho de que los intervalos x y y son disjuntos. Una IA-relación básica puede expresarse como una conjunción de PA-relaciones primitivas entre los puntos extremos de los intervalos. Así, la relación $x \{o\} y$ puede expresarse en el nivel de puntos como $(x^- < y^-) \wedge (x^- < y^+) \wedge (x^+ > y^-) \wedge (x^+ < y^+)$, donde x^- e y^- denotan los extremos inferiores de los intervalos x y y , respectivamente, y x^+ e y^+ los extremos superiores).

IA-RELACIÓN	POSICIÓN RELATIVA
x before (b) y	$\leftarrow x \rightarrow \quad \leftarrow y \rightarrow$
x meets (m) y	$\leftarrow x \rightarrow \leftarrow y \rightarrow$
x overlaps (o) y	$\leftarrow x \rightarrow$ $\leftarrow y \rightarrow$
x during (d) y	$\leftarrow x \rightarrow$ $\leftarrow y \rightarrow$
x starts (s) y	$\leftarrow x \rightarrow$ $\leftarrow y \rightarrow$
x finishes (f) y	$\leftarrow x \rightarrow$ $\leftarrow y \rightarrow$
x equals (eq) y	$\leftarrow x \rightarrow$ $\leftarrow y \rightarrow$

Tabla II. Relaciones básicas entre intervalos

Formalmente el *álgebra de intervalos* viene dada por la estructura $(2^I, ^{-1}, \cap, \circ_I)$ formada por los 8192 subconjuntos del conjunto de relaciones primitivas I (incluyendo la *relación universal* I y la *relación nula* \emptyset), la operación unaria *inversa* ($^{-1}$) y las operaciones binarias de *intersección* (\cap) y *composición* (\circ_I). La inversa de una IA-relación disyuntiva $R = \{r_1, \dots, r_k\}$ puede obtenerse como $R^{-1} = \{r_1^{-1}, \dots, r_k^{-1}\}$ (puesto que la operación inversa distribuye sobre la unión) y la intersección de dos relaciones $R'_{i,j}, R''_{i,j} \in 2^I$ se obtiene como la intersección de los conjuntos de relaciones básicas que constituyen $R'_{i,j}$ y $R''_{i,j}$. La composición de dos IA-relaciones, $R_{x,y}$ entre los intervalos

x, y y $R_{y,z}$ entre los intervalos y, z , es una nueva relación $R_{x,z} = R_{x,y} \circ_I R_{y,z}$ entre los intervalos x, z *inducida* por las relaciones $R_{x,y}$ y $R_{y,z}$, de forma que: $x R_{x,z} z \Leftrightarrow x R_{x,y} y \wedge y R_{y,z} z$. Para la composición de dos IA-relaciones primitivas hay que usar una *tabla de transitividad* de 13×13 elementos, que puede encontrarse en [3]. La composición de IA-relaciones disyuntivas se calcula como la unión de todas las posibles composiciones de pares de relaciones básicas: $R_{x,y} \circ_I R_{y,z} = \{r' \circ_I r'' \mid r' \in R_{x,y}, r'' \in R_{y,z}\}$ (la composición distribuye sobre la unión de relaciones). Una red del álgebra de intervalos, *IA-red*, está formada por un conjunto de variables que representan intervalos de tiempo y un conjunto de IA-relaciones entre ellos. Se representa mediante un grafo de restricciones dirigido, donde los nodos corresponden a las variables y los arcos están etiquetados con las IA-relaciones.

Desafortunadamente todos los problemas de razonamiento interesantes sobre el álgebra de intervalos son NP-duros [42]. En particular el problema de la consistencia es NP-completo [41]. Por tanto resolver estas tareas es bastante costoso, y es interesante considerar subclases del álgebra de Allen que tengan mejores propiedades computacionales. Consideremos las más significativas:

a) *Continuous endpoint subclass*, **C-IA**: conjunto de IA-relaciones que pueden expresarse como una conjunción de CPA-relaciones entre los puntos límite de los intervalos [42].

b) *Pointisable subclass*, **P-IA**: conjunto de IA-relaciones que pueden expresarse como una conjunción de PA-relaciones [42, 40, 39].

c) *ORD-Horn subclass*, **H-IA**: conjunto de relaciones del IA que puede expresarse como conjunción de restricciones de ORD-Horn sobre los puntos límite de los intervalos. Una restricción de ORD-Horn se entiende aquí como una disyunción de cualquier número de literales negativos $a \neq b$ y como mucho un literal positivo de la forma $a = b$ ó $a \leq b$ [30]. Es la única clase maximal tratable que contiene a todas las relaciones básicas del IA y cubre un 10 por ciento del álgebra completa.

Para las clases C-IA y P-IA los problemas de la consistencia, búsqueda de una solución, relación mínima, red mínima y consistencia global, se resuelven trasladando las relaciones entre intervalos a conjunciones de relaciones entre puntos, y aplicando los algoritmos referidos en la sección 2.1.

Para la clase H-IA se aplica un algoritmo de ca-

mino consistencia para decidir la consistencia en $O(n^3)$ y un algoritmo de 5-consistencia para obtener la red mínima en $O(n^5)$ [30]. Estos resultados han sido mejorados por Koubarakis [21] que decide la consistencia de un conjunto de restricciones ORD-Horn en $O(\max(n^2, Hn))$ donde H es el número de inecuaciones y la red mínima se puede obtener en $O(\max(n^4, Hn^3))$.

Existen otros conjuntos tratables. En [15] se presentan 21 clases tratables que cubren el 92 por ciento del álgebra de intervalos; 12 de ellas son maximales pero no usan todas las relaciones primitivas del IA. En [16] se caracteriza el conjunto de subálgebras del IA que son tratables y contienen ciertas relaciones básicas. La conclusión es que no hay ningún subálgebra tratable no conocida que pueda contener más de tres relaciones básicas $eq, d, o, s, f, d^{-1}, o^{-1}, s^{-1}, f^{-1}$. Eso significa que no parece que se puedan encontrar todavía nuevas subclases con mucha expresividad.

3. Restricciones Temporales Métricas

Dechter et al. [14] formalizaron el sistema de razonamiento temporal cuantitativo o métrico llamado TCSP (*Temporal Constraint Satisfaction Problem*) y un caso particular de éste, el STP (*Simple Temporal Problem*) cuya representación es completamente equivalente a los mapas temporales de Dean y McDermott [11].

Una *restricción cuantitativa* o métrica en este modelo se refiere a la distancia temporal entre variables que representan a instantes de tiempo. Una restricción se representa por un conjunto de intervalos reales $\{I_1, \dots, I_k\} = \{[a_1, b_1], \dots, [a_k, b_k]\}$. Si $k > 1$ entonces la restricción se clasifica como *disyuntiva métrica*, siendo *simple métrica* en caso contrario. Una restricción cuantitativa $C_i = \{I_1, \dots, I_k\}$ es una restricción *unaria métrica* si restringe el dominio de la variable p_i al conjunto de intervalos dado y esencialmente representa la disyunción $(a_1 \leq p_i \leq b_1) \vee \dots \vee (a_k \leq p_i \leq b_k)$. Una restricción *binaria métrica* $C_{i,j} = \{I_1, \dots, I_k\}$ restringe los valores admisibles para la distancia temporal $p_j - p_i$ y representa la disyunción $(a_1 \leq p_j - p_i \leq b_1) \vee \dots \vee (a_k \leq p_j - p_i \leq b_k)$. La *restricción universal* es $(-\infty, +\infty)$.

Un *TCSP* es una red binaria con un conjunto de variables que representan puntos de tiempo

$\{p_1, \dots, p_n\}$ y un conjunto de restricciones unarias y binarias métricas entre ellas. Con objeto de operar con restricciones cuantitativas se usa el *álgebra métrica* que consta de las operaciones de intersección de conjuntos (\cap), composición (\otimes) e inversa ($^{-1}$). Dadas dos restricciones métricas $C_{i,j}$ y $C_{j,k}$, la composición $C_{i,j} \otimes C_{j,k}$ puede ser calculada como la unión de sumas de pares de intervalos. La restricción inversa C^{-1} se obtiene como $C^{-1} = \{I_j^{-1} \mid I_j \in C\}$, donde la inversa de un intervalo $[a, b]$ es el intervalo $[-b, -a]$.

Una red TCSP puede representarse mediante un *grafo de restricciones dirigido* donde los nodos representan las variables o puntos de tiempo y los arcos están etiquetados con las restricciones métricas. En el modelo STP, cada relación entre dos puntos viene dada por un único intervalo, siendo por tanto un caso particular del TCSP. Se representa por un grafo de restricciones dirigido, de forma que si tenemos un arco $i \rightarrow j$ etiquetado por el intervalo $[20, 30]$ expresa que p_1 ocurre entre 20 y 30 unidades de tiempo antes que p_2 , esto es, $20 \leq p_2 - p_1 \leq 30$. La figura 2 muestra un ejemplo de STP.

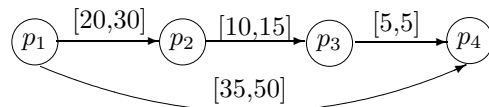


Figura 2. Grafo de restricciones de un STP

Un STP puede ser asociado con un grafo de distancias convencional G_d , que tiene los mismos vértices que el grafo de restricciones y cada restricción $[a, b]$ entre dos nodos p_i y p_j se descompone en dos restricciones: la restricción b sobre la distancia de p_i a p_j y la restricción $-a$ sobre la distancia de p_j a p_i . De esta forma, el problema puede resolverse aplicando el bien conocido algoritmo de Floyd-Warshall [1], cuyo objetivo es la búsqueda del camino más corto entre cada par de nodos de un grafo de distancias. Por tanto, la consistencia de un STP y el grafo mínimo pueden obtenerse en $O(n^3)$. Una vez que se ha obtenido el grafo mínimo, encontrar una solución requiere solo un tiempo $O(n^2)$. Se demuestra también que aplicar un algoritmo de camino consistencia como PC-1 [23] al grafo de restricciones de un STP tiene el mismo efecto que aplicar Floyd-Warshall a su grafo de distancias.

En aplicaciones a tareas de planificación, las restricciones dependen a menudo de condiciones so-

bre hechos del dominio, siendo necesario un modelo que permita la especificación de restricciones disyuntivas. El TCSP corresponde a este caso más general pero, desafortunadamente, el problema de *decidir la consistencia* en un TCSP es NP-completo [14]. De hecho la complejidad de resolver un TCSP descomponiendo éste en todos los posibles STPs es de $O(n^3k^e)$, donde k es el número máximo de intervalos por arco, y e es el número de arcos del grafo de restricciones. Por eso, en lugar de descomponer el problema, lo que se hace es ejecutar un algoritmo con retroceso sobre un meta-CSP cuyas variables son los arcos del TCSP y los dominios son los posibles intervalos. Aunque el peor de los casos en esta aproximación es también $O(n^3k^e)$ en la práctica permite utilizar técnicas de mejora del backtracking [12, 13]. También se proponen: 1) el uso de algoritmos de camino-consistencia, bien como aproximación a la red mínima, bien como preproceso para mejorar la aplicación del algoritmo con retroceso; 2) el uso de algoritmos que explotan las características topológicas de la red [34].

El algoritmo de camino consistencia PC-2 [23] aplicado a un TCSP con dominios enteros o racionales termina en $O(n^3R^3)$. Sin embargo, cuando el rango R es muy grande o cuando el dominio es el conjunto de números reales, aplicar PC-2 es problemático y puede llegar a ser impracticable porque el número de intervalos crece exponencialmente, lo que se conoce como *fragmentación*. Para evitar este problema Schwalb y Dechter [34] proponen algoritmos polinomiales de aproximación, semejantes al de camino consistencia, para decidir la consistencia y calcular la red mínima, llamados *Upper-Lower Tightening* (ULT) y *Loose Path-Consistency* (LPC).

4. Restricciones Entre Puntos y Duraciones

Hasta hace poco, no era usual incluir duraciones dentro de la ontología temporal, pero diversos autores en los últimos años han puesto de manifiesto la necesidad de establecer restricciones sobre la duración de los eventos [3, 5, 38, 20, 43, 37, 29], lo que permite modelar expresiones como “la duración d_1 es mayor que la duración d_2 ”. Consideremos el siguiente ejemplo, extendiendo al de [24]:

Ejemplo 1 *Bob, Fred y John trabajan para una*

compañía que posee una oficina local y otra principal en Los Angeles. Cuando trabajan en la oficina local, John tarda menos de 20 minutos en ir al trabajo, y Fred tarda entre 15 y 20 minutos. Dos veces por semana, John trabaja en la oficina principal, tardando 60 minutos como mínimo en llegar al trabajo. Hoy John salió de casa entre las 7:05-7:10 a.m. y Fred llegó al trabajo entre las 7:50-7:55 a.m. Bob ha tardado menos que Fred, y hoy ha salido de casa antes de las 7:45 a.m. Además, sabemos que Fred y John han coincidido en un semáforo, camino del trabajo.

En este ejemplo, aparece una restricción cualitativa sobre duraciones (“Bob ha tardado menos que Fred”) y restricciones métricas disyuntivas sobre puntos. Un razonador temporal debería poder decidir si toda esta información es consistente e inferir información adicional, como por ejemplo la hora de llegada de Bob o quién llega antes al trabajo. Esto puede hacerse con los modelos de restricciones entre puntos y duraciones, de los que existen varias versiones que difieren en expresividad y complejidad computacional [6, 26, 43, 27, 29].

El modelo PDN (*Point-Duration Network*) extiende al álgebra de puntos introduciendo variables de duraciones y relaciones cualitativas entre ellas. Llamamos *duración* a una variable, denotada como d_{ij} , que representa el tiempo transcurrido entre dos instantes p_i y p_j del eje temporal, que consideramos aquí modelado por el conjunto \mathbb{R}_0^+ . Una duración viene dada por una ecuación del tipo $d_{ij} = |p_i - p_j|$, que impone una *relación ternaria no lineal* que muestra la influencia de los puntos sobre la duración y viceversa. Dos duraciones pueden estar relacionadas mediante el mismo tipo de relaciones que se establecen en el álgebra de puntos, esto es, cualquier subconjunto de $T = \{<, >, =\}$.

Una *red cualitativa de puntos y duraciones* (PDN) es una estructura $\Sigma_{PD} = \langle N_P, N_D, Rel(P, D) \rangle$ formada por dos redes del álgebra de puntos, N_P y N_D , y un conjunto de relaciones ternarias $Rel(P, D)$ entre puntos y duraciones, donde:

a) La red N_P viene dada por un conjunto de variables $P = \{p_1, \dots, p_n\}$ que representan instantes o puntos temporales con dominio \mathbb{R}_0^+ y un conjunto $Rel(P) = \{R_{i,j} \in 2^T \mid \forall 1 \leq i, j \leq n\}$ de PA-relaciones entre puntos.

b) N_D está formada por un conjunto de variables $D = \{d_{ij} \mid p_i, p_j \in P\}$ que repre-

sentan duraciones y un conjunto $Rel(D) = \{R_{ij,km} \in 2^T \mid \forall d_{ij}, d_{km} \in D\}$ de PA-relaciones entre duraciones.

c) El conjunto de relaciones ternarias viene especificado por: $Rel(P, D) = \{(P_i, P_j, D_{ij}) \in (\mathbb{R}_0^+)^3 \mid D_{ij} = |P_i - P_j|, \forall d_{ij} \in D\}$

Nos referimos a los conjuntos $Rel(P)$, $Rel(D)$ y $Rel(P, D)$ como *PD-restricciones*. Se ha demostrado que el problema de la satisfacibilidad en PDN es NP-completo, lo que hace intratables a las demás tareas de razonamiento [26]. Por otra parte, su expresividad está limitada a relaciones cualitativas. Por ello, la actividad más reciente se ha centrado, tanto en la formalización de superclases más expresivas que el modelo PDN, como en la búsqueda de fragmentos tratables del modelo PDN y de sus superclases.

Un fragmento tratable del modelo PDN es el llamado PDN-simple, en el que sólo se permiten relaciones binarias cualitativas primitivas, es decir, sin disyunciones. En [27] se introducen condiciones necesarias y suficientes para la consistencia de una PDN-simple, y a partir de ellas se introduce un algoritmo tratable para decidir la consistencia, basado en el cálculo de componentes fuertemente conexas. Su complejidad es $O(n^2 + d^2)$, donde n es el número de instantes y d es el número de duraciones. Las soluciones pueden entonces obtenerse incrementalmente (sin retroceso) en tiempo $O(n^2 \times d)$. Sin embargo, es obvio que el modelo PDN-simple tiene una expresividad reducida.

Cabe preguntarse si existe algún modelo intermedio que mejore la expresividad del PDN-simple, sin sobrepasar la frontera de la tratabilidad. En [27] se presenta otro fragmento tratable denominado PPDN (Modelo de Puntos con Precedencia y Duraciones). En él, todas las variables de duraciones d_{ij} corresponden a duraciones entre pares de puntos ordenados $p_i < p_j$. Este requisito elimina la no-linealidad de las restricciones ternarias, ya que $d_{ij} = |p_i - p_j| = p_j - p_i$, y se ha demostrado que garantiza la tratabilidad. Merece la pena destacar que este requisito no limita demasiado la expresividad respecto al caso general PDN, pues cuando uno intenta restringir la duración de un evento es natural suponer que su duración es mayor que cero. Esto es similar a decir, en el contexto del álgebra de intervalos de Allen, que el punto de comienzo de un intervalo es estrictamente menor que su punto final. La consistencia de una red-PPDN puede establecerse en tiempo $O(n^3 + d^3)$, complejidad idéntica a la asociada a la obtención de soluciones.

Si retomamos el ejemplo 1, observamos que en él existe información métrica que no puede representarse en ninguno de los modelos cualitativos comentados hasta ahora. Una forma de superar esta limitación consiste en extender las redes-PA de puntos y de duraciones con restricciones cuantitativas sobre los dominios de las variables. Las restricciones sobre dominios son restricciones unarias, porque sólo afectan a una variable. Se obtiene así el modelo APDN (*Augmented Point-Duration Network*). Ahora, el ejemplo 1 sí puede ser completamente capturado en este formalismo. La figura 3 muestra los grafos de puntos y duraciones correspondientes al ejemplo 1; por simplicidad no se dibujan las relaciones ternarias, que están implícitas en los nombres de las variables. Este nuevo formalismo [43, 29] subsume a las redes-PA, a las redes-PA aumentadas [24], al modelo TCSP [14] y, por supuesto, al modelo PDN. Como consecuencia de ello, el modelo APDN no es tratable, sin embargo el modelo APDN-simple (sin disyunciones, ni métricas, ni cualitativas) sí lo es. En éste, la consistencia puede decidirse en tiempo $O(d \times n^2)$ y puede encontrarse una solución en $O(d \times n^3)$.

Una clase aún más general es la correspondiente al modelo MPDN (*Metric Point-Duration Network*), que extiende y subsume al APDN añadiendo restricciones binarias métricas a las restricciones unarias [28]. Las restricciones métricas, unarias o binarias, se expresan mediante intervalos reales, como en la sección 3.

Una *red métrica de puntos y duraciones* (MPDN) es una estructura $\Sigma_{PD} = \langle N_P, N_D, Rel(P, D) \rangle$ formada por dos TCSP, N_P and N_D , y un conjunto de restricciones ternarias $Rel(P, D)$, donde:

a) N_P viene dada por un conjunto de variables $P = \{p_1, \dots, p_n\}$ que representan instantes de tiempo y toman valores en \mathbb{R}_0^+ , y un conjunto de restricciones unarias y binarias métricas entre puntos.

b) N_D está formada por un conjunto $D = \{d_{ij} \mid p_i, p_j \in P\}$ de variables de duración sobre \mathbb{R}_0^+ y un conjunto de restricciones unarias y binarias métricas entre duraciones.

c) $Rel(P, D)$ es un conjunto de restricciones ternarias dadas por tripletas de valores reales, (P_i, P_j, D_{ij}) , para los puntos y duraciones, que satisfacen la ecuación de la distancia euclídea $d_{ij} = |p_i - p_j|$, para cada duración d_{ij} .

Se ha propuesto un caso particular tratable llamado MPDN simple [28], en el que las restricciones métricas no admiten disyunciones y todas las duraciones lo son con precedencia. Para este fragmento, la consistencia puede ser determinada en tiempo $O(n^3 + d^3)$, por lo que el modelo supone un buen compromiso entre expresividad y complejidad.

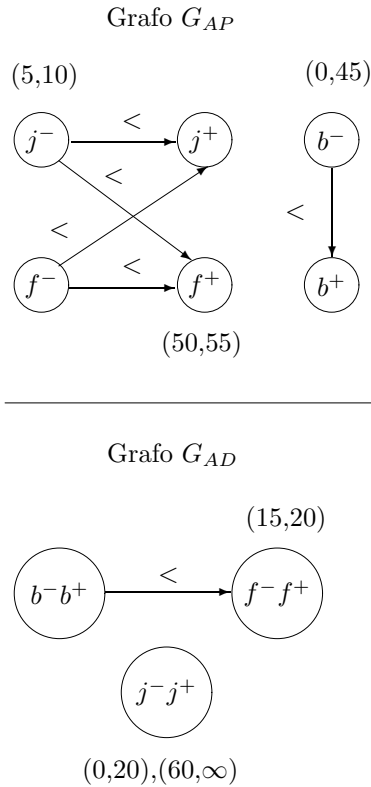


Figura 3. Grafos de puntos y duraciones

5. Conclusiones

En este trabajo se ha planteado el problema del razonamiento temporal como un caso particular de Problema de Satisfacción de Restricciones (CSP), en el que las variables representan a entidades temporales (puntos, intervalos o duraciones) y las restricciones corresponden a relaciones temporales entre esas entidades, cualitativas o métricas. Siguiendo esta perspectiva, se han revisado los principales modelos de razonamiento temporal basados en restricciones. Este campo, al igual que la inteligencia artificial en general, ha evolucionado más como una ingeniería que como una ciencia. En cualquier ingeniería, cobran es-

pecial relevancia los límites prácticos. Por ello, la perspectiva de esta revisión se ha centrado en los aspectos de ingeniería, proporcionando una visión general de la expresividad de los distintos tipos de modelos para razonamiento temporal, e intentando identificar las fronteras entre tratabilidad e intratabilidad computacional. Estos aspectos son clave para la aplicabilidad a problemas prácticos de tamaño real.

A pesar de los avances experimentados en las dos últimas décadas en la formalización de modelos de restricciones temporales, quedan muchas cuestiones por resolver. Estas cuestiones centran las tendencias actuales y hacen que siga siendo un campo de investigación muy activo. Así, por ejemplo, la búsqueda de fragmentos tratables y expresivos es un problema abierto, especialmente para el caso de modelos híbridos que integran distintos tipos de entidades temporales y relaciones temporales, como es el caso de los modelos de puntos y duraciones.

Otro tema de gran interés es el desarrollo de razonadores temporales prácticos, completos y eficientes que puedan ser reutilizados en distintos dominios de aplicación. Ejemplos en esta línea son los razonadores LaTeR (Layered Architecture for Temporal Reasoning) [9] o FuzzyTIME (Fuzzy Temporal Information Management Engine) [10], ambos orientados a modelos tratables, y dotados de lenguajes de interacción muy expresivos. Cuando se trabaja con modelos no tratables, es necesario que el razonador temporal combine algoritmos de preprocesamiento y algoritmos con retroceso, buscando siempre un compromiso entre el tiempo de computación dedicado a ambas fases. Este tipo de soluciones prácticas requiere de razonadores inteligentes, capaces de analizar los costes computacionales de las distintas técnicas de preprocesamiento, y de elegir la combinación de técnicas más apropiada para cada problema particular. En esta línea, en [8] se presenta un razonador temporal que aplica reglas heurísticas para seleccionar los algoritmos aproximados más rentables en problemas de redes de restricciones temporales disyuntivas borrosas. Las reglas heurísticas se obtuvieron tras un análisis experimental exhaustivo en el que se caracterizó el comportamiento de una amplia variedad de algoritmos para unos 40.000 problemas de restricciones con distintos valores de parámetros como conectividad, rango y amplitud de las restricciones, o número de variables.

La búsqueda de técnicas para la obtención de so-

luciones parciales en problemas sobrerrestringidos es otra línea que puede tener un impacto práctico interesante en aplicaciones de planificación o secuenciación sujetos a restricciones temporales. El desarrollo de métodos de abstracción temporal de datos y, particularmente, de técnicas para gestionar eventos repetitivos y patrones temporales complejos es un ejemplo más de línea en la que se trabaja activamente.

Finalmente, cabe citar aplicaciones emergentes del razonamiento temporal basado en restricciones a campos como la minería de datos temporales o la web semántica, que pueden experimentar un gran desarrollo en los próximos años.

Agradecimientos

Este trabajo ha sido subvencionado por las ayudas TIC2000-0873-C02-02 y FEDER2003-UNMU-E006 del MCyT español, así como por la ayuda PB/46/FS/02 de la Fundación Séneca de la Región de Murcia.

Referencias

- [1] A. Aho, J. Hopcroft, and J. Ullman. *The design and analysis of computer algorithms*. Addison-Wesley Publishing Company, 1974.
- [2] M. Alfonso. *Un modelo de integración de técnicas de clausura y CSP de restricciones temporales: Aplicación a problemas de scheduling*. PhD thesis, Departamento de Ciencia de la Computación e Inteligencia Artificial. Universidad de Alicante, Alicante, España, 2001.
- [3] J. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.
- [4] J. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–154, 1984.
- [5] J. Allen and J. Koomen. Planning using a temporal world model. In *Proceedings of IJCAI-83*, pages 741–747, 1983.
- [6] F. Barber. A metric time-point and duration-based temporal model. *SIGART Bulletin*, 3:30–49, 1993.
- [7] C. Bettomo. Mining temporal relationships with multiple granularities in time sequences. *IEEE Data Engineering Bulletin*, 21(3):32–38, 1998.
- [8] A. Bosch, M. Torres, and R. Marin. Reasoning with disjunctive fuzzy temporal constraint networks. In *Proceedings of TIME-2002*, pages 36–43, 2002.
- [9] V. Brusoni, L. Console, and P. Terenziani. Efficient query answering in LaTeR. In *Proceedings of TIME95*, pages 121–128, 1995.
- [10] M. Campos, A. Carceles, J. Palma, and R. Marin. A general purpose fuzzy temporal information management engine. In *Proceedings of EURASIA-ICT 2002, Workshop FM-IPS*, pages 93–97, 2002.
- [11] T. Dean and D. McDermott. Temporal data base management. *Artificial Intelligence*, 32:1–55, 1987.
- [12] R. Dechter. Enhancement schemes for constraint processing: backjumping, learning, and cutset decomposition. *Artificial Intelligence*, 41(3):273–312, 1990.
- [13] R. Dechter and I. Meiri. Experimental evaluation of preprocessing algorithms for constraint satisfaction problems. *Artificial Intelligence*, 68:211–241, 1994.
- [14] R. Dechter, I. Meiri, and J. Pearl. Temporal constraint networks. *Artificial Intelligence*, 49:61–95, 1991.
- [15] T. Drakengren and P. Jonsson. Twenty-one large tractable subclasses of allen’s algebra. *Artificial Intelligence*, 93:297–319, 1997.
- [16] T. Drakengren and P. Jonsson. A complete classification of tractability in allen’s algebra relative to subsets of basic relations. *Artificial Intelligence*, 106:205–219, 1998.
- [17] I. Hamlet and J. Hunter. A representation of time for medical expert systems. *Lecture Notes in Medical Informatics*, 33:112–119, 1987.
- [18] K. Kahn and G. Gorry. Mechanizing temporal knowledge. *Artificial Intelligence*, 9:87–108, 1977.
- [19] H. Kautz. A formal theory of plan recognition and its implementation. In J. Allen, H. Kautz, R. Pelavin, and J.D. Tenenber, editors, *Reasoning about Plans*, pages 69–125. Morgan Kaufmann, 1991.

- [20] M. Koubarakis. Dense time and temporal constraint with \neq . In *Proceedings of KR-92*, pages 24–35, 1992.
- [21] M. Koubarakis. Tractable disjunctions of linear constraints. *Lecture Notes in Computer Science*, 1118:297–307, 1996.
- [22] P. Ladkin and R. Maddux. The algebra of constraint satisfaction problems and temporal reasoning. Technical report, Kestrel Institute, Palo Alto, CA, 1988.
- [23] A. Mackworth. Consistency in networks of relations. *Artificial Intelligence*, 8(1):99–118, 1977.
- [24] I. Meiri. Combining qualitative and quantitative constraints in temporal reasoning. *Artificial Intelligence*, 87:343–385, 1996.
- [25] I. Meiri and J. Pearl. Faster constraint satisfaction algorithms for temporal reasoning. Technical Report R-151, Computer Science Department, University of California, Los Angeles, CA, 1990.
- [26] I. Navarrete and R. Marin. Qualitative temporal reasoning with points and durations. In *Proceedings of IJCAI-97*, pages 1454–1559, 1997.
- [27] I. Navarrete, R. Marin, M. Cardenas, and M. Balsa. Un nuevo modelo tratable y eficiente para razonamiento temporal cualitativo con puntos y duraciones. *Inteligencia Artificial*, (Premio ‘Jose Cuenca’ al mejor artículo de CAEPIA’99/TTIA’99), 9:12–19, 2000.
- [28] I. Navarrete, A. Sattar, and R. Marin. Deciding consistency of a point-duration network with metric constraints. In *Proceedings of TIME-ICTL2003*, pages 147–154, 2003.
- [29] I. Navarrete, A. Sattar, R. Wetprasit, and R. Marin. On point-duration networks for temporal reasoning. *Artificial Intelligence*, 140:39–70, 2002.
- [30] B. Nebel and H. Burckert. Reasoning about temporal relations: A maximal tractable subclass of allen’s interval algebra. *Journal of the ACM*, 42(1):43–66, 1995.
- [31] K. Nokel. Temporally distributed symptoms in technical diagnosis. In *Proceedings of IJCAI-89*, pages 1255–1260, 1989.
- [32] J. Palma and R. Marin. Modelling contextual meta-knowledge in temporal model based diagnosis. In *Proceedings of ECAI 2002*, pages 407–411, 2002.
- [33] A. Isli S. Belhadji. Temporal constraint satisfaction techniques in job shop scheduling problem solving. *Constraints: An International Journal*, 3:203–212, 1998.
- [34] E. Schwalb and R. Dechter. Processing disjunctions in temporal constraint networks. *Artificial Intelligence*, 93:29–61, 1997.
- [35] E. Schwalb and L. Vila. Temporal constraints: a survey. *Constraints*, 2-3:129–150, 1998.
- [36] F. Song and R. Cohen. The interpretation of temporal relations in narrative. In *Proceedings of AAAI-88*, pages 745–750, 1988.
- [37] S. Staab. From binary temporal relations to non-binary ones and back. *Artificial Intelligence*, 128(1-2), 2001.
- [38] E. Tsang. The consistent labeling problem in temporal reasoning. In *Proceedings of AAAI-87*, pages 251–255, 1987.
- [39] P. van Beek. Reasoning about qualitative temporal information. *Artificial Intelligence*, 58:297–326, 1992.
- [40] P. van Beek and R. Cohen. Exact and approximate reasoning about temporal relations. *Computational Intelligence*, 6:132–144, 1990.
- [41] M. Vilain and H. Kautz. Constraint propagation algorithms for temporal reasoning. In *Proceedings of the National Conference on Artificial Intelligence (AAAI-86)*, pages 377–382, 1986.
- [42] M. Vilain, H. Kautz, and P. van Beek. Constraint propagation algorithms for temporal reasoning: A revised report. In D. Weld and J. de Kleer, editors, *Readings in qualitative reasoning about physical systems*, pages 373–381. Morgan Kaufmann, 1989.
- [43] R. Wetprasit and A. Sattar. Temporal reasoning with qualitative and quantitative information about points and durations. In *Proceedings of AAAI-98*, pages 656–663, 1998.