

Metodología Semicualitativa para Razonar sobre Sistemas Dinámicos *Semiquantitative Methodology to Reasoning about Dynamic Systems*

J. A. Ortega¹, R. M. Gasca¹, J. Torres¹, M. Toro¹, L. González¹, F. Velasco¹, C. Angulo²

¹Universidad de Sevilla. ²Universidad Politécnica de Cataluña
Av. Reina Mercedes s/n 41012 – Sevilla (Spain)
e-mail: ortega@lsi.us.es

Artículo recibido en enero 25, 2000; aceptado en noviembre 05, 2004

Resumen

En este artículo se propone una metodología para razonar sobre los modelos semicualitativos construidos para sistemas dinámicos con conocimiento cualitativo y cuantitativo. La información cualitativa de estos sistemas puede componerse de: operadores cualitativos, etiquetas cualitativas, funciones de bandas y funciones continuas cualitativas. Se presenta un formalismo para incorporar esta información a los modelos. La metodología propuesta permite estudiar no sólo del régimen estacionario, ampliamente estudiado en la literatura, sino que además posibilita realizar un estudio del régimen transitorio de los sistemas. Se presenta también un estudio teórico sobre la validez de las conclusiones obtenidas en la metodología. Los comportamientos del sistema se pueden obtener automáticamente aplicando algoritmos de clustering y se expresan mediante un conjunto de reglas jerárquicas obtenidas mediante algoritmos genéticos. La metodología se ha aplicado a un par de modelos, siendo uno el modelo de dos estanques interconectados y otro un modelo de crecimiento logístico donde se ha incorporado un retraso en el bucle de realimentación.

Palabras clave: Modelos semicualitativos, Representación del conocimiento, Razonamiento cualitativo.

Abstract

In this article a methodology to reason over semiquantitative models built for dynamic systems with qualitative and quantitative knowledge is proposed. The qualitative information of these systems can be composed of: qualitative operators, qualitative labels, bands functions and qualitative continuous functions. A formalism to incorporate this information to the models is presented. The proposed methodology allows to study not just about the stationary regime, widely studied in the literature, but also it makes possible to carry out a study of the transitory regime of the systems. It also presents a theoretical study about the validity of the conclusions obtained in the methodology. The behaviours of the system can be obtained automatically applying clustering algorithms and are expressed through a set of hierarchical rules obtained from genetics algorithms. The methodology has been applied to a couple of models, one of them is the interconnected pools model and the other a logistical growing model where a delay in the feedback curls has been incorporated.

Keywords: Semiquantitative models, Knowledge representation, Qualitative reasoning.

1 Introducción

En ciencia e ingeniería se estudian sistemas cuyo conocimiento está normalmente compuesto de información cualitativa, cuantitativa, y mezcla de las dos, que se denomina información semicualitativa. Todo este conocimiento debe ser tenido en cuenta a la hora de estudiar los modelos construidos para estos sistemas. Estos modelos deben aportar diferentes niveles de abstracción numérica, desde la descripción puramente cualitativa (Kuipers, 1994), semicuantitativa (Kay, 1996; Berleant y Kuipers, 1997); (Ortega *et al.*, 1998), numérica intervalar (Vescovi *et al.*, 1995; Corliss, 1995) y cuantitativa.

La mayoría de las veces en los sistemas dinámicos no se considera el conocimiento cualitativo para el estudio y simulación de los mismos. En tal caso se trabajan con modelos cuantitativos, cuyo estudio es bien conocido.

Si bien, si se considera el conocimiento cualitativo del sistema, se han desarrollado en la bibliografía diferentes aproximaciones: transformación de relaciones no lineales a lineales a trozos, el método de MonteCarlo, distribuciones de probabilidad, métodos numéricos con intervalos reales, relaciones causales (Bousson y Travé-Massuyès, 1994), conjuntos difusos (Bonarini y Bontempi, 1994), programación lógica con restricciones (Hickey, 1994), y combinación de todos los niveles de abstracción cualitativo y cuantitativo (Kay, 1996). En este trabajo los sistemas dinámicos estudiados poseen cierta clase de conocimiento cualitativo.

En relación con los sistemas dinámicos, hay que decir que hay bastante bibliografía que estudia el régimen estacionario de un sistema dinámico, sin embargo, el estudio del régimen transitorio también es bastante importante. Por ejemplo, en los sistemas industriales de producción es interesante el estudio del transitorio para perfeccionar su rendimiento o para conocer

el tiempo de recuperación tras fallos en su funcionamiento. Con la metodología descrita en este artículo se puede estudiar cualquier instante del sistema, o sea su régimen transitorio y estacionario.

Por otra parte, las técnicas de obtención de conocimiento en bases de datos (KDD) (Adriaans y Zantinge, 1996) son apropiadas para adquirir información de los datos almacenados en una base de datos.

El término KDD se utiliza para referirse a todo el proceso de obtención de información útil en datos. El problema de la extracción de conocimiento en bases de datos involucra varios pasos, desde la manipulación de los datos, para su preprocesamiento y transformación hasta la investigación y razonamiento sobre los mismos mediante técnicas procedentes de las matemáticas y de la estadística. Aunque el problema de extraer conocimiento de los datos (u observaciones) no es nuevo, su automatización en el contexto de las bases de datos abre muchos nuevos problemas aún sin resolver.

KDD ha evolucionado, y continúa evolucionando, en múltiples áreas de investigación tales como bases de datos, aprendizaje mediante computadores (machine learning), obtención de patrones, inteligencia artificial y razonamiento con conocimiento incompleto, adquisición de conocimiento en sistemas expertos, visualización de datos, etc. Los sistemas KDD aplican teorías, métodos y algoritmos de todos estos campos.

El término minería de datos (*data mining*) se ha usado mayoritariamente por estadísticos, e investigadores en el campo de las bases de datos, si bien últimamente está siendo utilizado por el mundo de los negocios. La minería de datos es uno más de los pasos del proceso de KDD. Los otros pasos adicionales son la preparación, selección y limpieza de los datos, incorporación de conocimiento apropiado y la interpretación de los resultados que se derivan. Una descripción detallada de estos pasos pueden encontrarse en (Fayyad *et al.*, 1996) (Ortega *et al.*, 2000).

Por otro lado, bases de datos históricas, temporales y espacio-temporales han sido profusamente estudiadas en la bibliografía (Agrawal *et al.*, 1995). Se han utilizado en aplicaciones financieras, comerciales y sistemas de producción. También en bases de datos científicas con series temporales procedentes de medidas de sensores, por ejemplo, en datos sobre el tiempo, geológicos, medioambientales, etc. En este artículo, estamos interesados en bases de datos temporales correspondientes a la evolución de sistemas dinámicos semicualitativos.

La teoría y las herramientas de las bases de datos proporcionan la infraestructura necesaria para almacenar, acceder y manipular los datos. Mezclando de manera adecuada, minería de datos, series temporales, y bases de datos se propone en este artículo una metodología para estudiar sistemas dinámicos que evolucionan en el tiempo.

La idea de la metodología se puede resumir en la siguiente frase: "*cuando un modelo semicualitativo se transforma en una familia de modelos cuantitativos, cada uno de estos modelos tiene un comportamiento cuantitativo diferente, sin embargo entre si pueden responder a comportamientos cualitativos similares*". De esta manera, la metodología propuesta va a permitir estudiar un modelo con conocimiento cuantitativo y cualitativo como si se tratara de un conjunto de modelos cuantitativos, ampliamente estudiados. De esta manera, podremos estudiar el comportamiento de estos modelos tanto en el régimen transitorio como en el estacionario, siendo esta otra de las principales aportaciones de este trabajo.

La metodología propuesta transforma un modelo semicualitativo en una familia de modelos cuantitativos. La simulación de cada modelo cuantitativo genera una trayectoria en el espacio de fase. Estos comportamientos cuantitativos se almacenan en una base de datos. Se propone un lenguaje de consulta/clasificación sobre esta base de datos. Las consultas permiten conocer las propiedades cualitativas de las trayectorias de la base de datos. La clasificación permite etiquetar los diferentes comportamientos cualitativos del sistema. Este comportamiento se expresa por medio de un conjunto de reglas cualitativas jerárquicas.

La obtención automática de los patrones de comportamiento se basa en algoritmos de clustering (Ester *et al.*, 1996; Guha *et al.*, 1998) pero aplicados sobre series temporales. La manera de describir estos comportamientos es mediante un conjunto de reglas jerárquicas obtenidas mediante algoritmos genéticos (Aguilar *et al.*, 1998).

Si bien, estos últimos procedimientos no se abordan en este artículo, y su tratamiento se deja para posteriores trabajos.

Este artículo se distribuye como sigue: en primer lugar se presenta de manera esquemática la metodología y se define el concepto de modelo semicualitativo. A continuación se muestra el tipo de conocimiento cualitativo y sus técnicas de la transformación. Seguidamente se explica el algoritmo para la generación de la base de datos, y se define la sintaxis abstracta y la semántica del lenguaje de consulta/clasificación. Se realiza además, un estudio teórico sobre la validez de las conclusiones obtenidas del modelo y su extrapolación al sistema real. La metodología desarrollada se recoge de manera esquemática en la figura 1.

La metodología se aplica a un par de modelos. El primero se trata de un modelo con dos tanques interconectados, que se trata de un ejemplo bastante estudiado en la bibliografía. El otro ejemplo es de un modelo de crecimiento logístico con un retraso. La metodología aquí desarrollada se va a aplicar a un sistema real de producción industrial de acero (Ortega *et al.*, 1999).

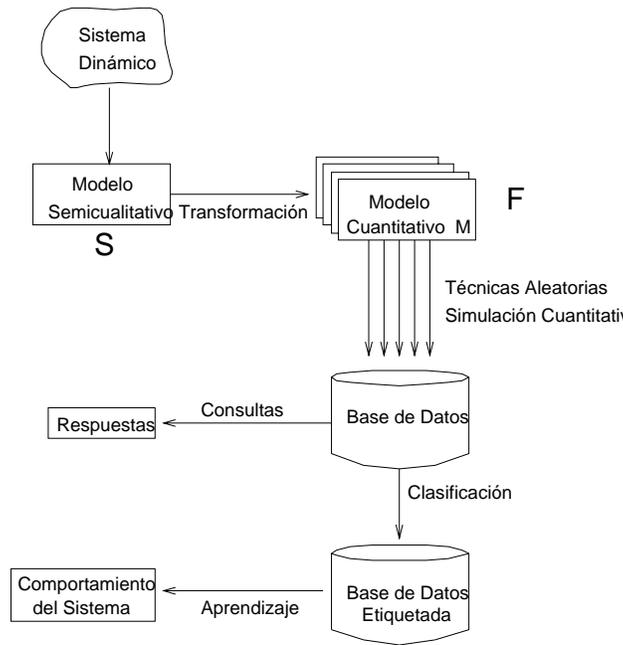


Fig. 1. Metodología propuesta

2 Metodología

Sea un sistema dinámico del que se tiene cierto tipo de conocimiento cualitativo. A partir de él y mediante técnicas de modelado se construye un modelo semicualitativo S .

Sobre el conocimiento cualitativo de este modelo se aplican una serie de reglas de transformación cuyo resultado es una familia F de modelos cuantitativos. El tipo de conocimiento cualitativo de estos sistemas, así como sus reglas de transformación se describen de manera detallada en la sección 5 de este artículo.

Para elegir cada uno de los modelos cuantitativo M de esta familia $M \in F$ se obtiene mediante técnicas estocásticas. Estos modelos se simulan de manera cuantitativa generando una trayectoria cuantitativa. Cada trayectoria se forma con los valores de las variables de estado desde su valor inicial hasta su valor final y los valores de los parámetros del sistema. Por lo tanto, cada trayectoria contiene para las variables de estado, sus valores en régimen transitorio y en estacionario.

Estos comportamientos de los modelos cuantitativos se almacenan en una base de datos de trayectorias cuantitativas T . Se propone un lenguaje para realizar preguntas sobre las propiedades cualitativas del conjunto de trayectorias incluidas en la base de datos. Estas trayectorias también pueden ser clasificadas según algún criterio, obteniéndose entonces una base de datos etiquetada.

Los diferentes patrones cualitativos de comportamiento del sistema se pueden obtener automáticamente a partir de la base de datos etiquetada aplicando algoritmos de aprendizaje basados en algoritmos genéticos, si bien estos últimos no se explican en este artículo, dejando su descripción para futuros trabajos.

3 Modelo Semicualitativo

Los modelos semicualitativos correspondientes a sistemas dinámicos con conocimiento cualitativo y cuantitativo que se tratan en este artículo se formulan con operadores aritméticos, operadores relacionales, funciones predefinidas (ln , exp , sen), literales numéricos y por supuesto por conocimiento cualitativo que se describe posteriormente.

Sea un modelo semicualitativo S obtenido de un sistema dinámico con conocimiento cualitativo, formulado como sigue

$$\Phi(\dot{x}, x, q, t), \quad x(t_0) = x_0, \quad \Phi_0(q, x_0) \quad (1)$$

Siendo $x \in \mathfrak{R}^n$ el conjunto de variables de estado del sistema, q los parámetros, t el tiempo, \dot{x} la variación de las variables de estado con el tiempo, Φ_0 las restricciones con las condiciones iniciales del sistema, y Φ las restricciones que definen el sistema y que dependen de \dot{x}, x, q, t . Siendo Φ y Φ_0 ecuaciones $f(x,y) = 0$ e inecuaciones $f(x,y) > 0$, que por claridad se va a representar de manera abreviada como se ha hecho en 1.

Cuando la metodología se aplica, las ecuaciones del sistema dinámico (1) se transforman en un conjunto de restricciones entre las variables, parámetros e intervalos. En este artículo, estamos interesados en aquellos sistemas que pueden expresarse como (2) después de que las reglas de transformación se hayan aplicado

$$\begin{cases} \dot{x} = f(x, p, t), & x(t_0) = x_0, \\ p \in I_p, & x_0 \in I_0 \end{cases} \quad (2)$$

Donde p incluye los parámetros del sistema y los nuevos parámetros obtenidos por medio de las reglas de la transformación que posteriormente explicaremos en este trabajo, f es una función obtenida aplicando las reglas de transformación, e I_p, I_0 son intervalos reales. La ecuación (2) es una familia F de sistemas dinámicos dependientes de p y x_0 .

4 Conocimiento Cualitativo

Los sistemas dinámicos que se estudian en este trabajo pueden tener conocimiento cualitativo en los parámetros, en las condiciones iniciales y en las funciones que forman el sistema de ecuaciones diferenciales. Este conocimiento se puede expresar mediante operadores y etiquetas cualitativas, funciones de banda y funciones continuas cualitativas.

La representación del conocimiento cualitativo se va a realizar mediante operadores que tienen asociados intervalos reales. Esta representación tiene como principal ventaja el permitir la fácil integración del conocimiento cualitativo y cuantitativo. (Gasca *et al.*, 1996). También posibilita la incorporación del conocimiento de los expertos en la definición del rango de las variables y los parámetros cualitativos del sistema. Hay técnicas desarrolladas sobre análisis de intervalos (Moore, 1979; Alefeld y Herzberger, 1983) para solucionar los problemas de ecuaciones diferenciales con variables y parámetros que son intervalos.

4.1 Operadores Cualitativos

Los *operadores cualitativos* se utilizan para describir características cualitativas de los parámetros y condiciones iniciales del sistema. Cada operador op cualitativo se define mediante un intervalo I_{op} donde se puede decir que la magnitud tiene esa etiqueta cualitativa. El intervalo para cada magnitud es suministrado por los expertos, de acuerdo con su conocimiento sobre el sistema.

Los operadores cualitativos se clasifican de acuerdo con el número de expresiones a los que se apliquen en unarios y binarios, que se representarán por U y B respectivamente. Es necesario que exista entre estos operadores una cierta relación. Esto es, no se pueden definir para una misma magnitud sus operadores unarios y binarios de manera arbitraria sino que deben guardar entre sí una cierta relación para que sea consistente su definición (Agell *et al.*, 2000).

4.1.1 Operadores Unarios

Un operador unario se utiliza para representar conocimiento en orden de magnitud absoluto sobre las características cualitativas del sistema dinámico: *el caudal es alto, la población de bacterias es baja*, etc. son ejemplos claros de este tipo de conocimiento.

Cada magnitud cualitativa que aparece en el sistema debe tener sus propios operadores unarios. No es lo mismo una magnitud que mida temperaturas cuyos valores pueden fluctuar en un rango de decenas que otra por ejemplo que mida bacterias en un sistema químico donde seguramente se hable de millones de unidades.

Sea U_x el conjunto de operadores unarios para una magnitud x , por ejemplo $U_x = \{NG_x, NM_x, NP_x, AP0_x, PP_x, PM_x, PG_x\}$ representa para x el conjunto de etiquetas cualitativas *muy negativo, moderadamente negativo, ligeramente negativo, ...*. La semántica de estos operadores así como sus reglas de transformación se explican en la siguiente sección de este artículo.

El objetivo de estas etiquetas es el permitir incorporar en la definición del problema el conocimiento que se pueda tener sobre la magnitud de una cierta expresión o variable del sistema. Este conjunto de etiquetas no siempre tiene porque ser el mismo. Ahora bien, si es habitual en la bibliografía, cuando se habla de etiquetas cualitativas unarias el tener un conjunto similar en número y semántica al que aquí se ha seleccionado (Travé-Massuyès *et al.*, 1997) (Agell *et al.*, 2000). En

cualquier caso, si se quisiera se podría tener mayor número de etiquetas en la definición del sistema. El efecto que esto tendría sería aumentar la expresividad en el planteamiento del problema, pero por contra, sería necesario redefinir la relación de este nuevo conjunto de etiquetas cualitativas unarias con las binarias.

4.1.2 Operadores Binarios

Los operadores binarios se utilizan para representar relaciones entre magnitudes del problema: *la altura del caudal del primer tanque es algo mayor que la del segundo tanque, la población inicial de bacterias es comparable a la final*. Para cada magnitud cualitativa del problema hay que definir su propio conjunto de operadores binarios.

Los operadores binarios se clasifican en:

- Operadores relacionados con la diferencia $\geq, =, \leq$ que representan las relaciones *mayor o igual, igual y menor o igual* entre las magnitudes a los que se apliquen.
- Operadores relacionados con el cociente $\{«<, < \sim, \sim <, > \sim, >», Cl, Vo, Ne, Di\}$. Estas etiquetas cualitativas representan relaciones del tipo *mucho menor que, moderadamente menor que*,

Una definición más detallada de la utilidad de estos operadores y de su definición se puede encontrar en (Mavrouniotis y Stephanopoulos, 1990) (Travé-Massuyès *et al.*, 1997). Asimismo, es necesario que esta definición de los operadores binarios se haga de manera que sea coherente con la hecha para los operadores unarios (Agell *et al.*, 2000)

4.2 Funciones de Banda

Una función de banda $y = g(x)$ (figura 2) representa la familia de funciones incluidas entre dos funciones reales, una superior $\bar{g} : \mathfrak{R} \rightarrow \mathfrak{R}$ y otra inferior $\underline{g} : \mathfrak{R} \rightarrow \mathfrak{R}$. Se denotan mediante la tupla

$$\langle \underline{g}(x), \bar{g}(x), I \rangle, \quad \forall x \in I : \underline{g}(x) \leq \bar{g}(x) \quad (3)$$

Donde I es el rango de definición de g , y x es la variable independiente.

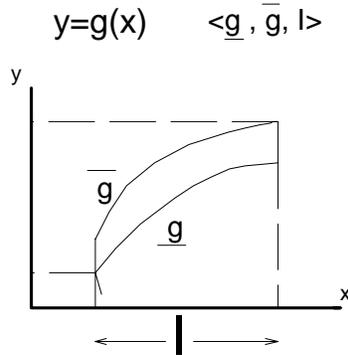


Fig. 2. Función de banda

El significado de cada función de banda es que para cualquier valor $\forall x \in I$ el valor de la función es un valor intervalar $y = [\underline{g}(x), \bar{g}(x)]$. Conforme a la definición heurística de la metodología propuesta, el valor que se elegirá para esta función, al realizar su evaluación cuantitativa será un valor elegido arbitrariamente dentro de ese intervalo.

4.3 Funciones Continuas Cualitativas

Una función continua cualitativa $y = h(x)$ (figura 3) representa una restricción entre los valores de x, y según las propiedades de h . Se denotan por

$$y = h(x), \quad h \equiv \{P_1, s_1, P_2, \dots, s_{k-1}, P_k\} \quad (4)$$

Siendo P_i puntos de la función, cada uno se define por el par de marcas cualitativas (d_i, e_i) , siendo d_i la marca asociada a la x y e_i a y . Estos puntos están separados por el signo s_i de la derivada en el intervalo entre dos puntos consecutivos. El signo s_i es $+$ si en el intervalo correspondiente la función es monótona creciente, $-$ si es decreciente, ó 0 si es constante. Una función cualitativa monótona es un caso particular de estas funciones donde el signo es siempre el mismo $s_i = \dots = s_{k-1}$.

$$y=h(x) \quad h= \{ P_1, +, P_2, +, P_3, +, P_4, -, P_5, -, P_6 \}$$

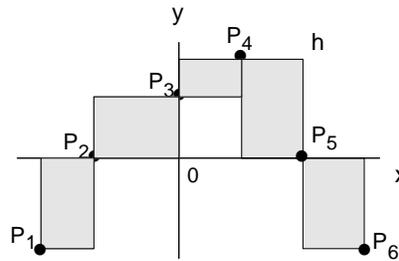


Fig. 3. Función cualitativa continua

La interpretación cualitativa (figura 4.a) para cada $P_i=(d_i, e_i)$ de $y=h(x)$ es

$$\begin{cases} \text{si } x = d_i \Rightarrow y = e_i \\ \text{si } d_i < x < d_{i+1} \Rightarrow \begin{cases} s_i = + \Rightarrow e_i < y < e_{i+1} \\ s_i = - \Rightarrow e_i > y > e_{i+1} \\ s_i = 0 \Rightarrow y = e_i \end{cases} \end{cases}$$

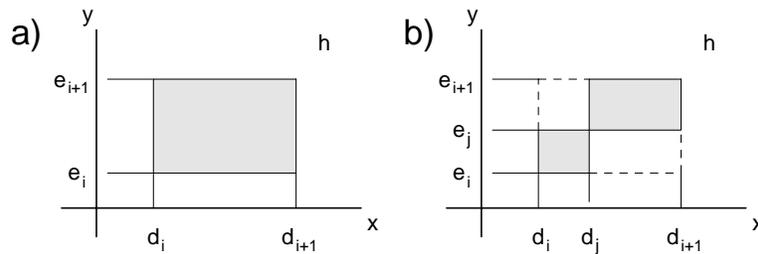


Fig. 4. Interpretación cualitativa de una función

5 Técnicas de Transformación

Para abordar la transformación de un modelo semicualitativo S en una familia de modelos cuantitativos F se aplican las reglas de transformación al conocimiento cualitativo de S que a continuación se detalla.

El modelo resultante de la transformación tras aplicar las reglas posteriormente descritas tiene la forma:

$$\begin{cases} \dot{x} = f(x, p, t), \quad x(t_0) = x_0, \\ p \in I_p, \quad x_0 \in I_0 \end{cases} \quad (5)$$

Esta es una familia de modelos cuantitativos dependiente de los parámetros, de las condiciones iniciales y de la transformación de las funciones cualitativas continuas.

5.1 Operadores Cualitativos Unarios

Sea U_x un conjunto de operadores unarios para una magnitud x , $U_x = \{NG_x, NM_x, NP_x, AP0_x, PP_x, PM_x, PG_x\}$ que representan para dicha magnitud las etiquetas cualitativas *muy negativo*, *moderadamente negativo*, *ligeramente negativo*, *aproximadamente cero*, *ligeramente positivo*, *moderadamente positivo* y *muy positivo* respectivamente.

La regla de transformación de cada operador unario $op_u \in U_x$ es

$$op_u(e) \equiv \begin{cases} e - r = 0, \\ r \in I_u \end{cases} \quad (6)$$

Donde r es una nueva variable que se introduce en el sistema e I_u el intervalo ligado a op_u definido de acuerdo con la parametrización de la tabla 1.

En esta tabla, α y β es una información que debe suministrar el experto para cada magnitud cualitativa del sistemas, así como el valor infinito v_∞ .

Operador op_u	Intervalo I_u
<i>NG</i>	$[-v_\infty, -\beta]$
<i>NM</i>	$[-\beta, -\infty]$
<i>NP</i>	$[-\alpha, 0]$
<i>AP0</i>	$[-\alpha/2, \alpha/2]$
<i>PP</i>	$[0, \alpha]$
<i>PM</i>	$[\alpha, \beta]$
<i>PG</i>	$[\beta, v_\infty]$

Tabla 1. Operadores cualitativos unarios

Los operadores unarios consecutivos se pueden combinar para obtener otros nuevos operadores. En este caso el intervalo asociado al nuevo operador será el que resulte de unir los intervalos.

Tanto la definición de los operadores, como la manera de obtener los intervalos se han definido de acuerdo con (Travé-Massuyès *et al.*, 1997).

5.2 Operadores Cualitativos Binarios

Las reglas de transformación de estos operadores dependen de su tipo.

- Operadores relacionados con la diferencia se les aplican las siguientes reglas de transformación

$e_1 = e_2 \equiv e_1 - e_2 = 0$
$e_1 \leq e_2 \equiv \begin{cases} e_1 - e_2 - r = 0 \\ r \in [-\infty, 0] \end{cases}$
$e_1 \geq e_2 \equiv \begin{cases} e_1 - e_2 - r = 0 \\ r \in [0, +\infty] \end{cases}$

Tabla 2. Operadores binarios relacionados con la diferencia

- Operadores relacionados con el cociente cuya regla de transformación es

$$op_b(e_1, e_2) \equiv \begin{cases} e_1 - e_2 * r = 0, \\ r \in I_b \end{cases} \quad (7)$$

Siendo r una nueva variable e I_b un intervalo ligado a op_b de acuerdo con su etiqueta cualitativa, tal y como se describe en la tabla 3.

Operador b	Intervalo I_b
«	$[0, Pre]$
-<	$[Pre, (1+Pre)^{-1}]$
~<	$[(1+Pre)^{-1}, 1]$
≈	$[(1+Pre)^{-1}, 1+Pre]$
>~	$[1, 1+Pre]$
>-	$[1+Pre, 1/Pre]$
»	$[1+Pre, v_{\infty}]$
Cl	$[(\epsilon+\delta)/2, 2-(\epsilon+\delta)/2]$
Vo	$[1-\delta*\epsilon, 1+\delta*\epsilon]$
Ne	$[-\delta*\epsilon, \delta*\epsilon]$
Di	$[-v_{\infty}, \delta] \wedge [2-\delta, +v_{\infty}]$

Tabla 3. Operadores binarios relacionados con el cociente

Donde los operadores «, -<, ~<, ≈, >~, >-, » , Cl, Vo, Ne, Di representan respectivamente las etiquetas cualitativas *mucho menor que, moderadamente menor que, ligeramente menor que, aproximadamente igual a, ligeramente mayor que moderadamente mayor que, mucho mayor que, cercano a, comparable con, despreciable respecto a* y finalmente *distante de*.

La tabla 3 se ha dividido en dos partes. El primer grupo de operadores, desde « hasta », son operadores parametrizados por Pre (parámetro de tolerancia). Su valor debe ser determinado por los expertos y representa una cantidad "mucho menor que uno" para la relación de órdenes de magnitud. La elección de los valores de los intervalos se ha definido de acuerdo con (Mavrovouniotis y Stephanopoulos, 1990).

En cuanto al segundo grupo de operadores, desde Cl hasta Di , los intervalos asociados se parametrizan de acuerdo con dos valores ϵ, γ . El parámetro ϵ es un infinitésimo positivo y su valor debe ser escogido teniendo en cuenta esta premisa. Estos parámetros deben verificar entre sí la siguiente relación de orden como se demuestra en (Agell *et al.*, 2000):

$$0 < \epsilon < \delta < 1 - \delta * \epsilon < 1 \quad (8)$$

La demostración de esta relación se puede encontrar en el trabajo de Ortega (Ortega *et al.*, 2000).

Por lo tanto a la hora de definir estos operadores es necesario que el experto suministre estos valores que sirven para parametrizar de manera adecuada la tabla 3 y de esta manera obtener la transformación de los operadores binarios que aparecen en el problema.

Los sistemas de transformación propuestos se basan en parte en contar con la información de un experto conocedor del sistema. En cualquier caso, para que los sistemas de transformación propuestos sean coherentes deberían de definirse respetando las relaciones de orden propuestas en la ecuación 8. Este es por tanto un tema abierto de investigación, ya que si se pudiera de alguna manera realizar una transformación más automática, siempre conseguiríamos una mayor garantía del proceso de transformación.

5.3 Funciones de Banda

La regla de la transformación que se aplica a una función de banda tal y como se define en (3) es la siguiente:

$$g(x) = \alpha \underline{g}(x) + (1 - \alpha) \overline{g}(x) \quad \text{con } \alpha \in [0, 1] \quad (9)$$

Donde α es una nueva variable que se introduce en el sistema, de manera que si $\alpha=0 \Rightarrow g(x) = \overline{g}(x)$ y si $\alpha=1 \Rightarrow g(x) = \underline{g}(x)$ y cualquier otro valor de α perteneciente al intervalo $[0, 1]$ representa un valor comprendido entre $\underline{g}(x)$ y $\overline{g}(x)$. Esta regla de transformación, permite transformar la función de banda en una función dependiente de las funciones de las que parte la función de banda, así como de una variable α perteneciente al intervalo $[0, 1]$, de manera que cuando esta variable tome un valor en dicho intervalo, se tendrá una instancia particular de la función de banda. Así pues, para cada selección de α se tiene una función de banda diferente.

5.4 Funciones Cualitativas Continuas

Las reglas de transformación de una función continua cualitativa se aplican en tres pasos: normalización, extensión y transformación.

a) Normalización. La definición de la función se *completa* y *homogeneiza* usando las siguientes propiedades de continuidad:

1. una función que cambia de signo entre dos puntos consecutivos pasa a través de un punto cuyo valor en la función es 0
 2. una función cuya derivada cambia de signo entre dos puntos consecutivos, pasa por un punto cuya derivada es cero.
- La definición de cualquier función (4) se completa siempre con los puntos extremos $(-\infty, +\infty)$, los puntos de corte con los ejes, y aquellos donde cambia la derivada (un máximo o un mínimo de h).

b) Extensión. La definición de estas funciones se enriquece mediante un proceso automático de incorporación de nuevas marcas. Esto se hace para disminuir la incertidumbre en la definición de una función, ya que como puede observarse el área del rectángulo se reduce (figura 4.b).

Sea k el número de puntos de una función cualitativa continua h . Si se añade un nuevo punto entre cada dos puntos consecutivos de h , entonces $2k-1$ puntos definen h . Cuando este proceso se repite i veces, el número de puntos que definen a la función cualitativa es

$$2^i k - \sum_{j=0}^{i-1} 2^j \quad (10)$$

Este número es importante para el algoritmo **Seleccionar H** que más adelante se explica.

c) Transformación. Una función cualitativa continua

$$y = h(x), \quad h \equiv \{P_1, s_1, P_2, \dots, s_{k-1}, P_k\} \quad (11)$$

Donde cada punto $P_i=(d_i, e_i)$ se transforma en un conjunto de funciones cuantitativas H . Cada función de H cumple las restricciones de la función cualitativa h . Se aplica el algoritmo **Seleccionar H** para obtener H . Este algoritmo divide h en sus segmentos.

Un *segmento* es una sucesión de puntos consecutivos $\{P_m, \dots, P_n\}$ de manera que todos tienen igual signo de derivada y no contienen un punto que corte un eje. Los segmentos dividen una función en regiones monótonas. El algoritmo propuesto aplica técnicas estocásticas para escoger cada función cuantitativa de H . Estas técnicas son similares al método de MonteCarlo, sin embargo, los valores obtenidos deben satisfacer las restricciones de h .

La heurística seguida aplica una distribución aleatoria uniforme para obtener los valores para cada marca. Los valores obtenidos deben verificar las mismas relaciones de orden que existían entre sus respectivas marcas.

6 Generación de la Base de Datos de Trayectorias

Aplicando las reglas de transformación a S se obtiene una familia F de modelos cuantitativos. Esta familia depende de un conjunto de parámetros intervalares p y de las funciones H . Cada modelo particular M de F se selecciona por medio de técnicas estocásticas, y se simula de manera cuantitativa aplicando cualquiera de las técnicas de simulación existentes (Euler, Runge-Kutta, ...). Esta simulación genera una trayectoria r que se almacena en la base de datos T . Cada trayectoria se forma con los valores de los parámetros y el conjunto de series temporales con los valores que cada variable toma desde el instante inicial hasta el final. Los siguientes algoritmos se aplican para obtener T .

Seleccionar Modelo(F)

para cada parámetro o variable intercalar de F seleccionar un valor en su intervalo para cada función h de F seleccionar una función cuantitativa H

Generación de base de datos T

$T := \{ \}$

para $i = 1$ a N

$M := \text{Seleccionar Modelo}(F)$
 $r := \text{Simulación Cuantitativa}(M)$
 $T := T \cup r$

Donde N es el número de simulaciones a realizar y se define de acuerdo con la sección 8.

7 Lenguaje de Consulta/Clasificación

Una vez que la base de datos con las simulaciones del modelo ha sido obtenido es necesario contar con un lenguaje que permita realizar consultas sobre el sistema, además de poder clasificar los diferentes comportamientos del sistema. El objetivo de esta sección es precisamente describir dicho lenguaje de consulta/clasificación.

El lenguaje debe permitir representar cualquier consulta que conforme a la dinámica de sistemas se hace sobre un sistema dinámico. Por lo tanto se debe permitir consultar tanto del régimen estacionario como del transitorio del sistema. Estas consultas se deben formular utilizando operadores lógicos, aritméticos y relacionales. Así como con predicados predefinidos y que son comúnmente utilizados en el análisis de los sistemas como son el del equilibrio o el ciclo límite.

Este lenguaje se amplía con la posibilidad de clasificar las trayectorias con etiquetas cualitativas, y así obtener una base de datos clasificada, a la que se le podrán aplicar técnicas de aprendizaje para describir los comportamientos del sistema.

7.1 Sintaxis Abstracta

Sea T el conjunto de todas las trayectorias r almacenadas en la base de datos.

7.1.1 Consultas

La sintaxis abstracta de las consultas que se pueden realizar con el lenguaje propuesto es la siguiente:

$Q : \forall r \in T \bullet [r, P]$ $ \exists r \in T \bullet [r, P]$ $ \mathbb{N} r \in T \bullet [r, P]$ $ [r, P]$	$P : P_b$ $ P \wedge P$ $ P \vee P$ $ \neg P$
$P_b : P_d$ $ f(L(F))$ $ \forall^r t : F \bullet F$ $ \exists^r t : F \bullet F$	$P_d : EQ$ $ CL$
$F : F_b$ $ F \& F$ $ F F$ $!F$	$F_b : e_b$ $ e \in I$ $ u(e)$ $ b(e, e)$

Tabla 4. Sintaxis abstracta del lenguaje

Una consulta Q es: un cuantificador $\forall, \exists, \mathbb{N}$ aplicado sobre T , o una consulta básica $[r, P]$ que evalúa a *true* cuando la trayectoria r verifica la propiedad P . El operador \forall se utiliza cuando se quiere ver si todas las propiedades verifican la consulta básica, mientras que el \exists se utiliza si se quiere consultar si alguna de las trayectorias verifican la consulta. El \mathbb{N}

sirve para saber el número de trayectorias que han verificado la consulta. Una propiedad P se puede formular mediante la composición de otras propiedades utilizando operadores lógicos \wedge, \vee, \neg . El resultado es la aplicación de los operadores entre las propiedades parciales.

Una propiedad básica P_b puede ser: una propiedad predefinida P_d , una función booleana f aplicada sobre una lista L de puntos o intervalos que verifican una fórmula F , o un cuantificador \forall, \exists aplicado a los valores de una trayectoria particular durante un tiempo t . Este tiempo puede ser: un instante, un rango de tiempo, un operador temporal predefinido, o la lista de tiempos donde se verifica una fórmula F .

Una propiedad predefinida P_d es aquella cuya formulación es automática. Corresponden normalmente con consultas que se realizan sobre los sistemas dinámicos y que su formulación es automática (Ortega *et al.*, 1998). Hay dos predefinidas: EQ que se verifica cuando la trayectoria acaba en un equilibrio estable; y CL cuando acaba en un ciclo límite.

Las consultas elementales sobre los sistemas dinámicos se expresan mediante fórmulas. En el lenguaje propuesto una fórmula F puede componerse de otras fórmulas combinadas por medio de operadores booleanos $\&, |, !$. Una fórmula básica F_b puede ser una expresión booleana e_b , o una expresión numérica e perteneciente a un intervalo, o un operador cualitativo unario u o binario b .

Clasificación

Una regla de la clasificación se plantea como un conjunto de consultas básicas con etiquetas, y posiblemente otras expresiones e_{n1}, e_{n2}, \dots

$$[r, P] \Rightarrow A, e_{n1}, \dots \quad [r, P_B] \Rightarrow B, e_{n2}, \dots \quad \dots$$

Una trayectoria r se clasifica con la etiqueta δ si verifica la propiedad P_δ .

7.2 Semántica

La semántica de cada instrucción de este lenguaje se transforma en una consulta sobre la base de datos. Una consulta $[r, P]$ es *true* si la trayectoria r verifica la propiedad P .

La semántica de una consulta con un cuantificador depende de este cuantificador. Si es \forall , evalúa a *true* cuando todas las trayectorias $r \in T$ verifican P . Si es \exists devuelve *true* cuando hay al menos una trayectoria $r \in T$ que verifica P . Si es N entonces se devuelve el número de trayectorias de T que cumplen P .

Sea $\forall t : F_1 \bullet F_2$ una propiedad básica, que será *true* si durante el tiempo que F_1 se satisface, todos los valores de r verifican F_2 . Si el cuantificador es \exists , entonces es *true* cuando al menos un valor de r que satisface F_1 , también satisface F_2 . Para evaluar una fórmula F , es necesario sustituir sus variables para sus valores. Estos valores se obtienen de T .

Sea $[r, P_A] \Rightarrow A, e_{A1}$ una regla de la clasificación. Una trayectoria $r \in T$ se clasifica con la etiqueta A si verifica propiedad P_A . El resultado de evaluar e_{A1} para esta trayectoria también se guarda en la base de datos.

8 Estudio Teórico sobre las Conclusiones

Cuando se afirma al aplicar la metodología que "hay un comportamiento del sistema que verifica la propiedad P " o que "todos los comportamientos del sistema verifican la propiedad P ", surge la pregunta: ¿son las conclusiones obtenidas aplicables al sistema real? La respuesta se encuentra en esta sección.

Realmente la pregunta que interesa hace es: ¿cual es la condición necesaria para afirmar que todos los comportamientos del sistema verifican una propiedad P ?

Sea Δ el espacio de trayectorias del sistema, y sea Ω el espacio de las trayectorias de Δ que verifican P (figura 5). Sea $Vol(s)$ el volumen de un espacio s . Estamos interesados en saber ¿cual es la condición que debe verificarse para asegurar que $Vol(\Delta) = Vol(\Omega)$? De forma esquemática, esta cuestión puede plantearse como ¿cual es la condición que debe cumplirse para que la siguiente implicación sea cierta?

$$\forall r \in T \bullet [r, P] \stackrel{\alpha}{\Rightarrow} \forall r \in \Delta \bullet [r, P] \quad (12)$$

Siendo α el grado de confianza. Necesitamos de técnicas estadísticas para responder a esta pregunta.

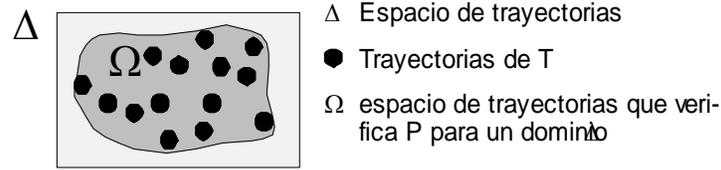


Fig. 5. Espacio de trayectorias

Sea p la probabilidad de que una trayectoria r verifique una propiedad Q y $q = 1 - p$

$$p = \frac{Vol(\Omega)}{Vol(\Delta)} \quad (13)$$

$\alpha = 0.05$	
$p = 0.6$	$n = 50$
$p = 0.8$	$n = 100$
$p = 0.98$	$n = 1000$
$p = 0.998$	$n = 10000$
$\alpha = 0.01$	
$p = 0.5$	$n = 200$
$p = 0.9$	$n = 1000$
$p = 0.99$	$n = 10000$
$p = 0.9999$	$n = 10^6$

Tabla 5. Ejemplos de relación α , n y p

Sea x una variable aleatoria, de forma que para n trayectorias el valor x es n si las $n-1$ primeras trayectorias verifican Q , y la n -ésima no. Sea α el grado de confianza. La expresión

$$\alpha = P(x > n) \quad (14)$$

Es la probabilidad que las n primeras trayectorias verifican Q y hay una trayectoria en lo que queda de Δ que no la verifica.

Teorema: La probabilidad p verifica que

$$p \geq 1 - \frac{1}{n\alpha} \quad (15)$$

Su demostración se encuentra en el apéndice A de este artículo.

La interpretación de este teorema es: *dato un grado de confianza α , si se quiere asegurar que una propiedad Q se verifica para un sistema dinámico con una probabilidad p , es necesario al menos obtener n trayectorias que la verifiquen.*

La siguiente tabla recoge varios ejemplos para p y n siendo $\alpha = 0.05$ y $\alpha = 0.01$:

De manera análoga, una consulta $\exists r \in \Delta \bullet [r, P]$ se puede siempre formular como $\neg \forall r \in \Delta \bullet [r, \neg P]$, aplicando una propiedad del cálculo de predicados, y por lo tanto, el estudio anterior sirve para este cuantificador. No obstante al ser una pregunta que sobre si existe un comportamiento que verifica una propiedad, basta simplemente con encontrar uno para poderlo afirmar, no obstante, siempre se pueden aplicar un razonamiento similar al realizado anteriormente.

9 Ejemplos de Aplicación

9.1 Modelo de Estanques Interconectados

Sea el sistema dinámico formado por dos tanques interconectados, tal y como se muestra en la figura 6. Se trata de un modelo bastante estudiado en la bibliografía y cuyo comportamiento es bastante conocido. En este caso artículo se añade a la definición de este modelo dinámico cierto conocimiento cualitativo.

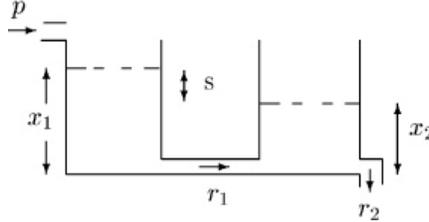


Fig. 6. Sistema dinámico de tanques interconectados

Las ecuaciones Φ que definen este modelo semicualitativo S son:

$$\left\{ \begin{array}{l} r_1 = h_1(s), r_2 = h_2(x_2), \\ s = x_1 - x_2, \frac{dx_1}{dt} = p - r_1 \\ \frac{dx_2}{dt} = r_1 - r_2 \\ h_1 \equiv \{(-\infty, -\infty), +, (0,0), +, (+\infty, +\infty)\}, \\ h_2 \equiv \{(-\infty, -\infty), +, (0,0), +, (+\infty, +\infty)\} \end{array} \right. \quad (16)$$

En este modelo p es un parámetro que indica el flujo que entra en el sistema. Las variables de estado son x_1, x_2 que denotan la altura del primer y segundo tanque respectivamente. Existe además una serie de variables intermedias en el sistema: s es la diferencia de alturas que existe entre ambos tanques, r_1 es el flujo que pasa entre el primer tanque y el segundo y r_2 es el flujo de salida del sistema. Estas dos variables r_1, r_2 son dependientes de las funciones cualitativas h_1, h_2 respectivamente. Estas dos funciones son monótonas crecientes que pasan por el origen.

Las condiciones iniciales del sistema se describen mediante la siguiente información cualitativa

$$\Phi_0 \equiv \left\{ \begin{array}{l} PM_x(x_1), \\ PM_x(x_2), \\ PM_p(p) \end{array} \right. \quad (17)$$

Estos operadores unarios cualitativos han sido definidos por los expertos como:

$$\langle x, 10, 20 \rangle, \quad \langle p, 10, 20 \rangle$$

Las alturas de ambos estanques inicialmente toman la etiqueta cualitativa *moderadamente positivo* al igual que el parámetro p . Ambos operadores se han definido con el mismo rango, ahora bien no hay que confundir que dichos operadores corresponden a magnitudes distintas, siendo necesario definirlos de manera separada. Para poder transformar estos

operadores en intervalos es necesario que los expertos definan el concepto de infinito para ambas magnitudes: $\infty_x = 200$ y $\infty_p = 100$.

Este modelo responde a la definición de un modelo semicualitativo, ya que tiene conocimiento cualitativo y cuantitativo que viene descrito tanto en la definición del sistema como en el estado inicial del mismo.

En este sistema interesa conocer

1. Si siempre alcanza un equilibrio estable.
2. Si en dicho equilibrio siempre la altura del primer tanque es mayor que la del segundo en la parte final de la simulación.
3. Clasificar la base de datos con tres etiquetas:
 - *Etiq1*, aquellas trayectorias donde la altura del primer tanque en tiempo final de la simulación es *moderadamente positivo*, y al principio de la simulación es *ligeramente mayor* que la del segundo tanque, estando esta relación entre magnitudes definida mediante $\langle ,r,100,0.1 \rangle$.
 - *Etiq2*, la trayectorias donde la altura del primer tanque es *moderadamente positiva* en la parte media de la simulación
 - *Etiq3*, las trayectorias que no se hayan clasificado con alguna de las etiquetas anteriores.

El modelo S se transforma en la familia de modelos cuantitativos F , que se describe mediante las ecuaciones:

$$\begin{cases} r_1 = H_1(s), r_2 = H_2(x_2), H_1, H_2, \\ s = x_1 - x_2, x_1 = p - r_1, \\ x_2 = r_1 - r_2, \\ x_1, x_2 \in [10,20] p \in [10,20] \end{cases} \quad (18)$$

Los intervalos de las variables de estado x_1, x_2 y del parámetro p se han definido aplicando las técnicas de transformación a los operadores unarios. Las funciones cualitativas h_1, h_2 mediante el algoritmo **GenerarH(h)** se transforman en H_1, H_2 respectivamente.

Sobre esta familia de modelos F (19) se aplica el algoritmo **GenerarBD(F)** para obtener la base de datos de trayectorias T .

Las consultas que se desean hacer sobre T utilizando el lenguaje se plantean mediante las siguientes instrucciones:

$$1. \forall r \in T \bullet [r, EQ]$$

En este caso se está interesado en conocer si todas las trayectorias de T terminan en un equilibrio estable. La respuesta a esta consulta es que todas las trayectorias verifican el predicado, por lo que se puede concluir que el sistema siempre acaba en un equilibrio estable y no presenta ciclos límites.

$$2. \forall r \in T \bullet [r, EQ \wedge \forall t \in RF \bullet x_1 > x_2]$$

La consulta intenta descubrir si las trayectorias que alcanzan el equilibrio en la parte final de la simulación RF se verifica que la altura del primer tanque x_1 es mayor que la del segundo x_2 . La respuesta a esta consulta de nuevo es afirmativa, es decir, todas las trayectorias la verifican.

3. La clasificación propuesta se plantea mediante las reglas:

$$\begin{cases} [r, (\forall t \in RF \bullet PM_x(x_1)) \\ \wedge (\forall t \in RI \bullet x_1 > \sim x_2)] \Rightarrow Etiq1 \\ [r, (\forall t \in RM \bullet PM_x(x_1))] \Rightarrow Etiq2 \\ [r, Re sto] \Rightarrow Etiq3 \end{cases}$$

Por simplificar, se ha puesto en la última regla de clasificación el operador *Resto*. Éste se formula como la negación de la disyunción de las otras propiedades, esto es:

$$\text{Resto} \equiv (\neg P_1 \vee \neg P_2) \quad (19)$$

Los resultados que se han obtenido con esta metodología para el régimen estacionario del sistema coinciden con otros aparecidos en la bibliografía que estudian este sistema. En el caso de la metodología propuesta en este artículo, además de conclusiones sobre el estacionario, podemos realizar consultas sobre el régimen transitorio. Un estudio más detallado de la aplicación de esta metodología a este modelo, se puede consultar en (Ortega, 2000). En la figura 7 se muestra los resultados obtenidos para el modelo propuesto.

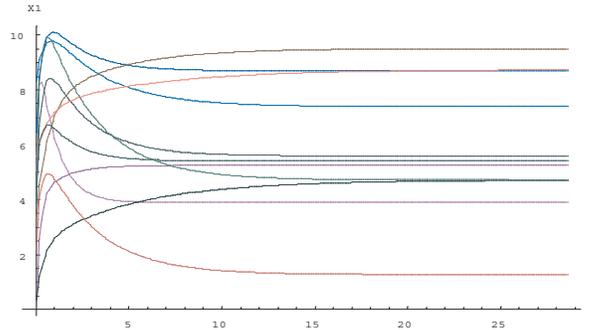


Fig. 7. Formas de alcanzar el equilibrio en los tanques

9.2 Modelo Logístico con Retraso

En el mundo real es muy común encontrar procesos de crecimiento donde una fase inicial de crecimiento exponencial es seguida por otra fase de acercamiento asintótico a un valor de saturación (figura 8.a). A estos procesos se le dan los nombres genéricos: logístico, sigmoideal, o *s-shaped*.

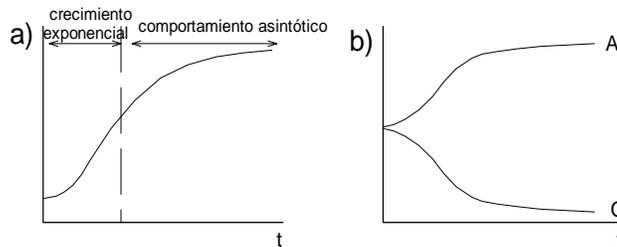


Fig. 8. Modelo con crecimiento logístico

Estos modelos se han estudiado bastante en literatura. Abundan en procesos naturales, sociales y socio-tecnológicos. Por citar algunos, aparecen en: la evolución de las bacterias, la extracción de mineral, el crecimiento de la población mundial, desarrollos económicos, las curvas de aprendizaje, determinados fenómenos dentro de una población, como rumores o epidemias, nuevos productos que se introducen en el mercado, etc. Estos sistemas responden a un patrón de comportamiento bimodal con dos atractores (figura 8.b): *A* crecimiento normal, y *O* decadencia y extinción.

Cuando se añade un retraso en la realimentación el comportamiento estacionario del sistema permanece inalterado, ahora bien el transitorio del mismo sufre alteraciones generando un nuevo tipo de comportamiento no contemplado en los modelos logísticos.

Sean las ecuaciones diferenciales de un modelo *S* semicualitativo correspondiente a un sistema logístico con un retraso en el ciclo de realimentación. Las ecuaciones del modelo son

$$\Phi \equiv \begin{cases} \dot{x} = x(nr - m), y = \text{delay}_\tau(x), \\ x > 0, r = h(y), \\ h \equiv \{(-\infty, -\infty), +, (d_0, 0), +, \\ (0, 1), +, (d_1, e_0), -, \\ (1, 0), -(+\infty, -\infty)\} \end{cases} \quad (20)$$

Siendo n el factor de crecimiento, m el factor de decrecimiento, y h una función cualitativa con un máximo en (x_1, x_2) (figura 9).

$$h \equiv \{(-\omega, -\omega), +, (d_0, 0), +, (0, 1), +, (d_1, e_0), -, (+\omega, -\omega)\}$$

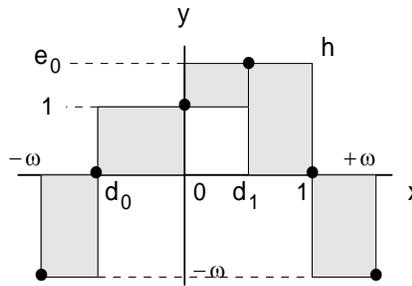


Fig. 9. Función cualitativa h

Las condiciones iniciales del sistema son

$$\Phi_0 \equiv \begin{cases} X_0 \in [LP_x, MP_x], LP_x(m), \\ LP_x(n), \tau \in MP_\tau, VP_\tau \end{cases} \quad (21)$$

Donde LP, MP, VP son operadores unarios cualitativos para las variables x, τ . Sobre este sistema se quiere conocer:

1. si siempre se alcanza un equilibrio
2. si hay un equilibrio cuyo valor no es cero
3. si todas las trayectorias con equilibrio cercano a cero, lo alcanzan sin oscilar
4. clasificar la base de datos de acuerdo con los patrones de comportamientos del sistema

Se aplica la metodología propuesta a este modelo. En primer término, se transforma S mediante la aplicación de las reglas de transformación

$$\begin{cases} x = x(nr - m), y = delay_\tau(x), \\ x > 0, r = H(y), H, \\ x_0 \in [0, 3], m, n \in [0, 1], \tau \in [0.5, 10] \end{cases} \quad (22)$$

Donde H se ha sido obtenido aplicando el algoritmo *Seleccionar H* sobre h , y los intervalos se definen de acuerdo con el conocimiento de los expertos. Con el algoritmo *Generación de Base de datos T* se obtiene la base de datos de trayectorias. Las preguntas propuestas se formulan con el lenguaje como sigue:

1. $\forall r \in T \bullet [r, EQ]$
2. $\exists r \in T \bullet [r, EQ \wedge \exists t : t \approx t_f \bullet !AP0_x(x)]$
3. $\forall r \in T \bullet \left[r, EQ \wedge \exists t : t \approx t_f \bullet \right. \\ \left. AP0_x(x) \wedge length(x=0) = 0 \right]$

Siendo $AP0_x$ un operador unario de x .

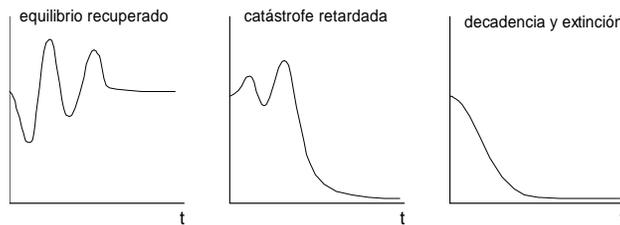


Fig. 10. Modelo de crecimiento logístico con retraso

La lista de puntos donde $x = 0$ es la lista con los máximos y mínimos de la trayectoria. Si su longitud es 0, entonces no hay oscilaciones.

4. La base de datos se clasifica con las etiquetas

$$[r, EQ \wedge \text{length}(x = 0) > 0 \wedge \exists t : t \approx t_f \bullet AP0_x(x)] \Rightarrow \text{equilibrio recuperado},$$

$$[r, EQ \wedge \text{length}(x = 0) > 0 \wedge \exists t : t \approx t_f \bullet AP0_x(x)] \Rightarrow \text{catástrofe retardada},$$

$$[r, EQ \wedge \exists t : t \approx t_f \bullet AP0_x(x)] \Rightarrow \text{decadencia y extinción}$$

Estos corresponden con los tres posibles comportamientos del sistema (figura 10). Los resultados obtenidos están en concordancia con aquellos que se obtienen cuando se realiza un razonamiento matemático (Karsky *et al.*, 1992; Aracil *et al.*, 1997).

10 Conclusiones y Trabajo Futuro

En este artículo se presenta de manera resumida una nueva metodología para automatizar el análisis de sistemas dinámicos con conocimiento cuantitativo y cualitativo. Se ha presentado una forma de incorporar esta información cualitativa a sistemas dinámicos semicualitativos. Los modelos para estos sistemas se construyen a partir de restricciones, operadores cualitativos, funciones aritméticas, de banda y cualitativas.

La metodología se basa en reglas de transformación, aplicación de técnicas estocásticas, simulación cuantitativa, generación de una base de datos de trayectorias y definición de un lenguaje de consulta/clasificación. La simulación se realiza por medio de técnicas estocásticas, guardándose los resultados en una base de datos que se puede clasificar con el lenguaje propuesto. Sobre la base de datos etiquetada se aplican algoritmos genéticos para obtener conclusiones sobre el sistema dinámico, si bien su detalle se deja para futuros trabajos.

Sobre los sistemas estudiados, los resultados obtenidos con la metodología expuesta coinciden con otros aparecidos en la bibliografía, por lo que parece apropiado continuar su estudio y aplicación a otros sistemas semicualitativos.

En el futuro, queremos enriquecer el lenguaje propuesto con: operadores para comparar trayectorias entre sí, operadores temporales entre distintos tiempos de una misma trayectoria, más tipo de ecuaciones,... En cuanto a los sistemas dinámicos estamos interesados en abordar los sistemas con restricciones y la incorporación de nuevos tipos de conocimiento cualitativo y la posibilidad que en el sistema existan múltiples escalas de tiempo.

Para finalizar decir, que una compañía metalúrgica está interesada en modificar su sistema de producción de acero aplicando la metodología que se ha expuesto en este artículo. En futuros trabajos, describiremos este sistema en detalle y las conclusiones obtenidas.

Apéndice A

En este apéndice se demuestra el siguiente teorema:

Teorema: Sea p la probabilidad de que una trayectoria r verifique una propiedad Q , el número de trayectorias, y α el grado de confianza. La probabilidad p verifica que:

$$p \geq 1 - \frac{1}{n\alpha} \quad (23)$$

Demostración:

La esperanza matemática $E[x]$ de una variable aleatoria x se define como

$$E[x] = \sum_{n=1}^{\infty} np^{n-1}q \quad (24)$$

Realizando manipulación simbólica y sustituyendo la suma geométrica por su valor se obtiene

$$E[x] = \frac{q}{p} \sum_{n=1}^{\infty} np^n = \frac{q}{p} \frac{p}{(1-p)^2} = \frac{1}{1-p} \quad (25)$$

Por otro lado, y aplicando la desigualdad de Chebyshev

$$E[x] = \sum_{x=1}^{\infty} xp(x) \geq \sum_{x=n+1}^{\infty} np^{n-1} \quad (26)$$

sustituyendo el sumatorio por su valor

$$E[x] = nP(x > n) \quad (27)$$

por lo tanto se obtiene que

$$\frac{E[x]}{n} \geq P(x > n) \quad (28)$$

sustituyendo $E[x]$ por su valor obtenido en (26), y si se aplica (14),

$$\frac{1}{n(1-p)} \geq P(x > n) = \alpha \quad (29)$$

operando de manera simbólica, se demuestra el teorema

$$\frac{1}{n\alpha} \geq 1-p \Rightarrow p \geq 1 - \frac{1}{n\alpha} \quad (30)$$

En la bibliografía existen otras aproximaciones para intentar justificar la idoneidad de aplicar el método de Monte Carlo a la hora de simular un conjunto de sistemas que evolucionan en el tiempo. En (Lagoa y Barmish, 2001) se realiza una aproximación intentando demostrar la selección que se hace, sin embargo, focalizándolo en consideraciones sobre la robustez de la distribución de Monte Carlo que se selecciona. En nuestro caso, demostramos que nuestra aproximación es independiente del modelo de selección que se haya elegido para aplicar el método de selección de valores, esto es, la distribución particular elegida.

Referencias

1. **Adriaans P., y Zantinge D.**, "Data Mining" Addison Wesley Longman, 1996.
2. **Agell N., Prats F., Sánchez M., Travé-Massuyès L.**, "A gateway from absolute to relative order of magnitude", *The 14th International Workshop on Qualitative Reasoning*, 2000, Morelia (México).
3. **Aguilar J., Riquelme J., y Toro M.**, "Decision queue classifier for supervised learning using rotated hyperboxes", en *Lecture Notes in Artificial Intelligence* n° 1484, 1998, pp. 326-336.
4. **Agrawal R., Lin K.I., Sawhney H.S., y Shim K.**, "Fast similarity search in presence of noise, scaling, and translation in time series databases", en *The 21th VLDB Conference*, 1995, Zurich (Switzerland).
5. **Alefeld G. y Herzberger J.**, "Introduction to Interval Computations", *Academic Press*, New York. 1983.
6. **Aracil J., Ponce E., Pizarro L.**, "Behavior patterns of logistic models with a delay", en *Mathematics and computer in simulation* 44, 1997, pp. 123-141.
7. **Berleant D. y Kuipers B.J.**, "Qualitative and quantitative simulation: bridging the gap" en *Artificial Intelligence* 95, 1997, pp. 215-255.
8. **Bonarini A. y G. Bontempi G.**, "Qua.SI.: Qualitative Simulation approach for fuzzy models", en *Proceedings of ESM'94*, Barcelona (España) 1994, pp. 420-424.
9. **Bousson K. y Travé-Massuyès L.**, "Putting more numbers in the qualitative simulator CA-EN", en *Proceedings of the 2nd International Conference on Intelligent Systems Engineering*, 1994.

10. **Corliss G.F.**, "Guaranteed error bounds for ordinary differential equations", en *Theory of Numerics in Ordinary and Partial Differential Equations*. Oxford University Press. 1995.
11. **Ester M., Kriegel H.-P., Sander J., y Xu X.**, "A density-based algorithm for discovering clusters in large spatial database with noise", en *International Conference on Knowledge Discovery in Databases and Data Mining (KDD-96)*, Portland (USA), 1996, pp. 226-231.
12. **Fayyad U., Piatetsky-Shapiro G., y Smyth P.** "The KDD process for extrating useful knowledge from volumes of data", en *Communications of the ACM* Vol 39 nº 11, 1996, pp. 27-34.
13. **Gasca R.M., Toro M., Ortega J.A.** "Propagación de restricciones integrando conocimiento cualitativo y cuantitativo", en *Boletín de AEPIA* núm 6, 1996, pp. 23-30.
14. **Gasca R.M.** "Razonamiento y Simulación en Sistemas que integran conocimiento cualitativo y cuantitativo". *Ph.D. Thesis*, Universidad de Sevilla (España), 1998.
15. **Guha S., Rastogi R., y Shim K.**, "CURE: And efficient clustering algorithm for large databases", en *Proceedings ACM SIGMOD International Conference Management of Data* New York (USA), 1998, pp. 73-84.
16. **Hickey T.J.**, "CLP(F) and Constrained ODEs". *Technical Report* 1994.
17. **Karsky M., Dore J.-C., y Gueneau P.**, "Da la possibilité d'apparition de catastrophes diffèrès", en *Ecodecision* nº 6, 1992.
18. **Karypis G.; Han E.-H.; y Kumar V.**, "Chameleon: Hierarchical clustering using dynamic modeling"
19. *Computer* nº 32 (8), pp. 68-75, 1999.
20. **Kay H.**, "Refining imprecise models and their behaviors". *Ph.D. Thesis* University of Texas (USA), 1996.
21. **Kuipers B.J.**, "Qualitative reasoning. Modeling and simulation with incomplete knowledge". *The MIT Press*, Cambridge, Massachussets, 1994.
22. **Lagoa C.M y Barmish B.R.** "Distributionally Robust Monte Carlo Simulation: A Tutorial Survey" Technical Report EECS 01-01 Electrical Engineering and Computer Science Department Case Western Reserve University, 2001
23. **Mavrouniotis M.L. y Stephanopoulos G.**, "Formal order-of-magnitude reasoning", en *Readings in Qualitative Reasoning about Physical Systems*, 1990, pp. 323-336.
24. **Moore R.E.**, "Methods and Applications of Interval Analysis" *SIAM Studies in Applied Mathematics* Philadelphia,(PA) 1979.
25. **Ortega J.A., Gasca R.M. y Toro M.**, "Automatización del análisis y simulación de modelos cualitativos mediante restricciones". *Actas de la VII Conferencia Asociación Española para Inteligencia Artificial*}, pp. 85-94, Málaga (España), 1997.
26. **Ortega J.A., Gasca R.M. y Toro M.**, "Including qualitative knowledge in semiquantitative dynamical systems", en *Lecture Notes in Artificial Intelligence* 1415, 1998, pp. 329-335.
27. **Ortega J.A., Gasca R.M. y Toro M.** "Obtaining semi-qualitative behaviors of dynamic systems by means of quantitative simulations". *Workshop Model based systems and qualitative reasoning of the Thirtieth European Conference on Artificial Intelligence*, Brighthon (UK), pp. 21-26, 1998.
28. **Ortega J.A., Gasca R.M. y Toro M.**, "A semiquantitative methodology for reasoning about dynamic systems". *The 13th International Workshop on Qualitative Reasoning* Loch Awe (Scotland), pp. 169-177, 1999a.
29. **Ortega J.A., Gasca R.M. y Toro M.**, "A new methodology for reasoning about semiquantitative dynamical systems". *The 3rd IMACS / IEEE International Multiconference on: Circuits, Systems, Communications and Computers - Software and hardware for the 21th Century* Athens (Greece), pp. 33-38, 1999b.
30. **Ortega J.A., Gasca R.M. y Toro M.**, "Behavior patterns of semiquantitative dynamic systems by means of quantitative simulations", en *The 16th International Joint Conference on Artificial Intelligence*. Qualitative and model based reasoning for complex systems and their control, Estocolmo (Suecia), 1999, pp. 42-48.
31. **Ortega J.A.**, "Patrones de comportamiento temporal en modelos semicualitativos con restricciones". *Tesis doctoral*, Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla, 2000.
32. **Travé-Massuyès, Dague Ph. y Guerrin F.**, "Le raisonnement qualitatif pour les sciences de l'ingénieur". *Hermes Ed.* Paris 1997.
33. **Vescovi M., Farquhar A., e Iwasaki Y.** "Numerical interval simulation: combined qualitative and quantitative simulation to bound behaviors of non-monotonic systems" en *Proceedings 14th International Joint Conference on Artificial Intelligence*, 1995, pp. 1806-1812.



Juan A. Ortega. *Obtuvo el grado de Doctor en Informática en 2000 por la Universidad de Sevilla (España). Desde 1992 es profesor en el Departamento de Lenguajes y Sistemas Informáticos en la Universidad de Sevilla. Sus áreas de interés son la simulación semicualitativa y la obtención de patrones temporales en modelos que evolucionan en el tiempo.*



Rafael M. Gasca. *Obtuvo el grado de Doctor en Informática en 1998 por la Universidad de Sevilla (España). Desde 1991 es profesor en el Departamento de Lenguajes y Sistemas Informáticos en la Universidad de Sevilla. Su investigación se centra en la diagnosis de sistemas y en seguridad informática.*



Jesús Torres. *Obtuvo el grado de Doctor en Informática en 1997 por la Universidad de Sevilla (España). Desde 1991 es profesor en el Departamento de Lenguajes y Sistemas Informáticos en la Universidad de Sevilla. Sus áreas de interés son la ingeniería de requisitos, la orientación a objetos y los sistemas distribuidos.*



Miguel Toro. *Obtuvo el grado de Doctor Ingeniero Industrial en 1987 por la Universidad de Sevilla (España). Desde 1985 es profesor en el Departamento de Lenguajes y Sistemas Informáticos en la Universidad de Sevilla. Sus investigaciones se centran en la simulación de sistemas dinámicos y en la ingeniería del software.*



Luis González Abril. *Es Profesor Asociado en el Dpto. De Economía Aplicada I de la Universidad de Sevilla. Es licenciado en Matemáticas por la Universidad de Sevilla y Dr. en Ciencias Económicas por la Universidad de Sevilla. Sus líneas de investigación son: máquinas de aprendizajes y bifurcaciones de sistemas dinámicos con aplicación a problemas económicos.*



Francisco Velasco Morente. *Es Profesor Titular de Universidad en el Dpto. de Economía Aplicada I de la Universidad de Sevilla. Es licenciado en Matemáticas por la Universidad de Sevilla y Dr. en Ciencias Matemáticas por la Universidad de Sevilla. Sus líneas de investigación son: bifurcaciones de sistemas dinámicos continuos y discretos y modelos de control óptimo, ambas con aplicación a problemas económicos.*



Cecilio Angulo Bahón. *Es Profesor Lector en el Dpto. de Ingeniería de Sistemas y Automática de la Universidad Politécnica de Cataluña. Es licenciado en Ciencias Matemáticas por la Universidad de Barcelona y Dr. en Ciencias por la Universidad Politécnica de Cataluña. Sus líneas de investigación son: aprendizaje automático y sistemas inteligentes de control.*