

Control de una columna de destilación binaria utilizando agentes inteligentes

Arturo González Sánchez¹
David Romero Romero²

¹ Programa de Investigación en Reactores y Procesos,
Instituto Mexicano del Petróleo (IMP).

² Sección de Estudios de Posgrado e Investigación (SEPI),
Escuela Superior de Ingeniería Mecánica y Eléctrica (ESIME),
Instituto Politécnico Nacional (IPN).
MÉXICO.

Teléfono: 9175 8415¹, 5729 6000, ext. 54784²

correo electrónico: agonzal@imp.mx¹, dromero@ipn.mx²

Recibido el 24 de agosto de 2007; aceptado el 3 de marzo de 2008.

1. Resumen

Se presenta un *sistema multi-agente* dedicado a tareas de control de procesos. El objetivo es aplicar la tecnología de agentes inteligentes a una estrategia de control, para que el control sea flexible, autoorganizado y autónomo. También se hace una comparación contra una estrategia de control avanzado para mostrar las ventajas de los agentes de software como controladores. Se escogió una columna de destilación binaria como proceso de prueba debido a las severas interacciones, respuesta no lineal, grandes tiempos muertos y gran tamaño. Este proceso es muy común en las industrias de refinación del petróleo y petroquímica. Las capacidades de los agentes pueden explotarse para obtener una estrategia de control capaz de sobreponerse a estas características del proceso. Como resultado se obtiene una estrategia de control capaz de sobreponerse a las interacciones y adaptable a la respuesta no lineal del proceso, que no se logra con el control convencional.

Palabras clave: control inteligente, agentes de software, columna de destilación binaria.

2. Abstract (Control of a Binary Distillation Column Utilizing Intelligent Agents)

A Multi-Agent System dedicated to process control tasks is presented. The objective is to apply the concepts of agent technology to a control strategy to make it flexible, self-coordinated and autonomous. Also a comparison against an advanced control system is made in order to show the advantages of software agents as controllers. A binary distillation column was selected as a test process because of its severe interactions, nonlinear response and long dead times. This process is common in oil refining and petrochemical industries. Agent capabilities

can be exploited to obtain a control strategy capable to overcome the process features like interactions and nonlinear response.

Key words: intelligent control, software agents, binary distillation column.

3. Introducción

En el contexto de control de procesos, la torre de destilación es uno de los procesos caracterizados por tener respuesta no lineal, considerables retardos de tiempo y fuertes interacciones entre sus variables [1], además es un sistema de gran tamaño. Esto constituye el obstáculo más difícil de salvar para el diseño de su control.

En esencia, las características de estos procesos hacen que la estabilidad de la planta sufra graves deterioros cuando operan bajo un esquema de control de lazos independientes, ya que ocurre una perturbación interna producida por variables que interactúan, mientras son ajustadas por sus respectivos controladores. Estas peculiaridades del proceso han impulsado la búsqueda de soluciones generalmente englobadas bajo el concepto de *control multivariable* [2].

Al aplicar la estrategia de control basada en agentes de software [3], dándole a cada agente su respectiva variable de control (VC) y su correspondiente variable manipulada (VM), se aprovechan las ventajas de diseñar un controlador que utiliza técnicas de inteligencia artificial para: *a)* aprender a tomar sus decisiones *b)* también aprender a manejar su VC así como el efecto que tienen las demás variables en el entorno sobre ésta, *c)* ser autónomo en sus decisiones por lo que no depende de alguien para construir y mantener su modelo así como realizar sus tareas, y *d)* coordinarse con los demás a través de mensajes, solicitudes y comandos. Con estas propiedades una estrategia de control puede anticiparse al efecto de todas las demás variables, incluyendo las perturbaciones externas.

El diseño del agente propuesto como controlador se basa en una arquitectura de tres capas; una planificadora basada en un *proceso de decisión de Markov* (PDM), encargada de decidir a qué tarea dedicarse en función del estado del proceso; una capa reactiva basada en redes neuronales que actúa simultáneamente como el sistema de percepción y acción del agente; una capa de coordinación basada en una heurística propia de los autores, que se encarga de las negociaciones con los demás agentes que se presenta el apartado 4.3.

El objetivo de este trabajo es determinar las ventajas de emplear agentes de software en vez del esquema de control regulatorio basado en dispositivos PID y un esquema de control avanzado.

Diversos ejemplos de agentes aplicados en la industria [4], reflejan el alcance de esta tecnología [5] [6] [7] [8] [9]. No obstante, en control de

procesos al nivel regulatorio las aplicaciones son escasas debido a las restricciones de tiempo que este tipo de control demanda. No obstante, la arquitectura propuesta resuelve con una red neuronal como máquina de inferencia del agente.

4. Arquitectura del agente

El agente diseñado para controlar una variable del proceso, se desenvuelve en un entorno constituido por los dispositivos de control y medición digitales, las computadoras del *sistema de control distribuido* (SCD) y la red de comunicaciones Ethernet. Dicho entorno es emulado por el simulador de proceso con el que el agente interactúa en este trabajo.

Se considera que en el sistema de control los dispositivos que pueden dar cabida a los agentes de software son los instrumentos de medición y control de la planta gracias a las capacidades de cómputo y almacenamiento con que cuentan. Así, dentro del SCD un agente tiene acceso a todas las variables del proceso pudiendo seleccionar aquellas de su interés para predecir sus efectos sobre la VC a su cargo.

Los agentes desarrollados para esta investigación son programas escritos en lenguaje C++ que se ejecutan en el *sistema operativo Windows* (SOW). Se basan en las definiciones del entorno de desarrollo *Java Agent Development Framework* (JADE) [10].

No obstante lo anterior, este proyecto se beneficia de las funciones que ofrece el SOW para la simulación del *sistema multi-agente* (SMA) ya que se utiliza como entorno, en el que pueden convivir los diferentes agentes de control y el simulador del proceso.

Simulador y agentes comparten un protocolo de comunicación basado en el KQML [11] para intercambiar información y comandos incluso a través de una red Ethernet, emulando la distribución del SCD.

Las decisiones del agente se reflejan como el tiempo que destina a las tareas que "sabe hacer", las cuales son módulos intercambiables codificados en archivos DLL.

La arquitectura del agente, representada en la Fig. 1, se distribuye en tres capas encargadas de: a) la planificación de actividades, b) el control de la VC asignada al agente y c) coordinación con los demás.

4.1. Capa de planificación reactiva

Las decisiones sobre qué tarea es más importante realizar en función del estado del proceso se manifiestan como el tiempo que el agente le asigna a cada tarea. Asimismo, el modelo de decisión debe considerar que en algún momento ciertas tareas pueden no ser necesarias por lo que no se planifican o incluso si están en ejecución sean suspendidas.

El PDM es la representación formal de la toma de decisiones que el agente utiliza para saber qué hacer ante las situaciones que se le presentan.

El objetivo del PDM es determinar la mejor acción en cada posible estado para maximizar la recompensa final. Para lo cual existen diferentes métodos de aprendizaje reforzado [12], en los que el objetivo es aprender del entorno interactuando con él sin tener un modelo inicial del mismo, ni de la función de recompensas.

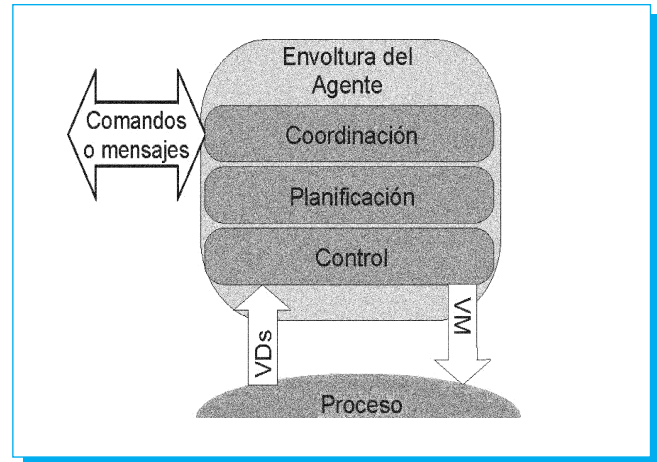


Fig. 1. Capas de la arquitectura empleada.

El método de aprendizaje reforzado utilizado en este trabajo es el llamado *Q-learning* [13]. Este método construye un mapa de los estados en que puede hallar el entorno así como la función de recompensa por las acciones que ejerce. Una vez que el agente construye su matriz Q durante una etapa de entrenamiento inicial es capaz de seleccionar la tarea y el tiempo destinado a realizarla para cumplir su objetivo: Mantener su VC en el *punto de ajuste* (PA) aún en presencia de perturbaciones en el entorno, identificadas como variables de disturbio (VD). Un PDM se compone de los elementos siguientes:

- Un conjunto finito de estados, S .
- Un conjunto finito de acciones asociadas a cada estado, $A(s)$.
- Una función de transferencia que define una distribución de probabilidad (DP), también definida como una función de probabilidad de transición, $P: S \times A \rightarrow DP$. La DP que describe el entorno posee la propiedad de Markov: dada la acción y el estado actual, el siguiente estado es independiente de todos los estados y acciones previas.

$$P(s_t | s_{t-1}, a) \quad (1)$$

El estado actual captura todo lo relevante sobre el entorno para predecir lo que será el siguiente estado.

$$P(s_t | s_{t-1}, a) = P(s_t | s_{t-1}, s_{t-2}, \dots, s_1, a) \quad (2)$$

- Una función de recompensa al tomar la acción a en el estado s , $R: S \times A \rightarrow R$.
- Una política estocástica de Markov, $\pi: S \times A \rightarrow [0, 1]$. Es la descripción formal de una conducta, representa la probabilidad de seleccionar la acción a en el estado s . Puede verse como la suma de las recompensas descontadas esperadas para llegar al estado al estado s_{t+1} . También puede verse como un plan reactivo generado en respuesta a la situación actual.

En cada episodio el PDM evalúa la política π no estacionaria:

$$V_{\pi, t}(s) = R(s, \pi_t(s)) + \gamma \sum_{s'} P(s' | s, \pi_t(s)) V_{\pi, t-1}(s') \quad (3)$$

Donde γ es el factor de descuento, si su valor es cercano a cero, el agente tiende a considerar solamente las recompensas inmediatas; si es cercano a uno, considera con mayor peso recompensas futuras, disponiéndose a retardar la recompensa.

La política óptima es la que obtiene la mayor recompensa escogiendo la acción óptima sin conocer P valiéndose de la función Q :

$$Q(s, a) \tag{4}$$

Donde $V^* \pi_t = \max_a Q(s, a)$, a' son las acciones posibles desde el estado s para ir a s' . Lo que recursivamente define la actualización de Q como:

$$Q(s) = R(s | a_t) + \gamma \max_{a'} Q(s_{t+1}, a') \tag{5}$$

Las acciones que el agente debe aprender para ir al estado estable desde los otros estados del proceso se representa en la tabla 1, de estados y recompensas también llamada matriz R .

La tabla 1 es la matriz R de estados y recompensas. Los estados en la columna del lado izquierdo representan al "estado actual" y los estados en la línea superior representan los estados a los que se puede ir. Los guiones en algunas celdas señalan que no se puede pasar entre esos estados. La matriz R modela el sistema de recompensas. En la Fig. 2 se representan los estados en los que el proceso puede estar, las acciones que el agente debe tomar y los eventos del proceso que lo hacen cambiar de estado.

Los estados que el agente identifica son los siguientes:

Estable (E): Es el estado objetivo en el que las variables del proceso no cambian con el tiempo o lo hacen variando alrededor del PA con una pequeña amplitud y un período muy grande. Existe algún grado de desviación aceptable dentro del $\pm 1\%$.

Estable con desviación (ED): Aunque estable, la VC está desviada del PA.

Transitorio (T): Ocurre cuando cambian las condiciones de operación. Algunas de las variables se desplazan de un valor a otro hasta que el proceso llega a otro estado estable.

Inestable (I): Se presentan oscilaciones sostenidas por diferentes causas, externas o internas. Asimismo, en la Fig. 2 las flechas que señalan la transición de un estado a otro indican que la transición ocurre por las causas siguientes:

af: La tarea para este arco es la de control prealimentado.

ar: La tarea es corregir utilizando retroalimentación.

ae: Detección de parámetros erróneos.

p o d: Las transiciones etiquetadas como "p o d" son causadas principalmente por el proceso debido a cambios en las condiciones de operación o disturbios.

Tabla 1. Matriz R de recompensas en la transición de estados.

R	E	E_D	I	T
E	10	-	-	0
E_D	10	1	-	0
I	-	-	0	3
T	10	4	0	0

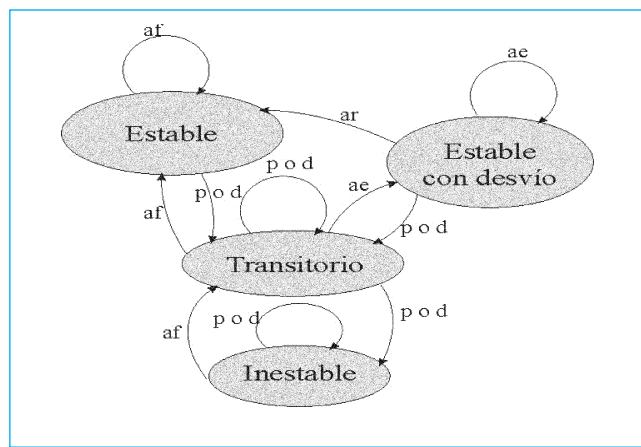


Fig. 2. Estados en los que se puede hallar la VC debido a las acciones del agente y a las perturbaciones internas y externas.

Las tareas que el agente puede realizar son:

Control prealimentado: Aplica su modelo de control prealimentado si el estado del proceso es E , T o I . En los dos últimos el PDM aumenta la frecuencia de los muestreos.

Control retroalimentado: Se utiliza si el estado del proceso es ED . Corregir a pasos la VM hasta acercar la VC al PA y activa la tarea de pos-entrenamiento.

Pos-entrenamiento: Debe capturar la combinación actual de VD, VM y reasignar el valor de VC que obtiene con esa combinación. También capturar la correlación corregida y reentrenar su red.

Coordinación: Esta tarea se activa cuando hay dos o más agentes que necesitan coordinarse para hallar un punto de balance. Administra los acuerdos entre ellos ya que éstos continuamente pueden crearse o destruirse, además de que las negociaciones son asíncronas y aleatorias.

Comunicación: Además de lo importante que es para la coordinación, esta tarea se encarga de obtener datos del proceso, enviar comandos y manejar la cola de mensajes que el agente recibe.

4.2. Capa de control de proceso

La capa de control del proceso consiste de una *red neuronal artificial* (RNA) por retropropagación [14] que da al agente habilidades de cumplir las restricciones de tiempo que demanda el control regulatorio. Como máquina de inferencia la RNA permite al agente: *a)* medir las variables del entorno (pudiendo ser un número considerable de ellas), *b)* inferir directamente su acción de control y transmitirla al elemento final en el mismo instante de muestreo, *c)* consumir muy poco tiempo en obtener una inferencia y *d)* seguir almacenando correlaciones en la misma estructura de conocimiento sin necesitar cada vez más espacio. Nótese que la acción de control no puede basarse en un mecanismo de búsqueda o toma de decisiones mediante la evaluación de reglas, porque dichos mecanismos podrían consumir más tiempo del que el agente tiene asignado para su tarea.

Con el fin de evitar la complejidad y los retardos en el entrenamiento de la red que pudieran impedirle al agente cumplir las restricciones de tiempo, así como evitar un crecimiento desmedido en las

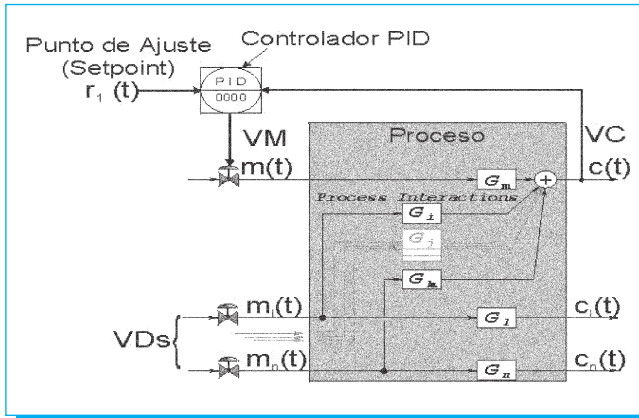


Fig. 3. Una sola variable controlada afectada por la VM y las VD.

dimensiones de las matrices de pesos, la RNA se utiliza como clasificador de patrones. Así, lo que almacena son los estados de balance entre las VDs, el PA y la VM. Por lo que se emplea una RNA entrenada para clasificar los patrones de VDs que se le presentan a la entrada, de modo que a cada combinación de VD la RNA da a su salida el valor de VM que les impide afectar a la VC. En este caso la RNA retropropagada es sencilla, eficiente y cumple con las necesidades del problema sin necesidad de una red más compleja.

Considerando el proceso multivariable de la Fig. 3, donde un lazo retroalimentado trata de corregir las perturbaciones causadas por las VD. Suponiendo que el proceso fuera lineal, la respuesta de la VC puede representarse en Laplace (en este apartado $s = \sigma + j\omega$) como la suma de las entradas cada una afectada por la función de transferencia entre sí y la salida.

$$VC(s) = VM(s)G_m(s) + \sum_{i=1}^n VD_i(s)G_i(s) \quad (6)$$

Donde: $G_m(s)$ es la función de transferencia entre la VM y la VC. $G_i(s)$ es la función de transferencia entre la VD_i y la VC.

Cada función $G(s)$ en (6), bien puede representarse por retardos de primer orden más tiempo muerto.

Suponiendo que la variación de las VD y en concordancia la variación de la VM son escalones de amplitud y duración infinitesimal, que al unirlos en el tiempo producen una trayectoria, puede considerarse a cada entrada como su función escalón en el dominio de Laplace, $VM(s) = m/s$ o $VD(s) = l_i/s$. Si se considera el estado de balance alcanzado una vez que ha transcurrido el efecto dinámico de las variables que afectan a la VC del agente; es decir, cuando el efecto dinámico de la VM y de cada VD transcurre:

$$\lim_{s \rightarrow 0} s \frac{m}{s} \frac{K_m e^{-sd_m}}{\tau_m s + 1} = mK_m \quad (7)$$

$$\lim_{s \rightarrow 0} s \frac{l_i}{s} \frac{K_i e^{-sd_i}}{\tau_i s + 1} = l_i K_i$$

Donde: K_m es la ganancia entre controlador y elemento de control. m es la magnitud del escalón en la VM. K_i es la ganancia que tiene cada

entrada i respecto a la VC. l_i es la variación escalón de la perturbación i . Por lo que en estado de balance la VC resultante es:

$$VC = mK_m + \sum_{i=1}^n l_i K_i \quad (8)$$

Para construir el modelo del controlador a RNA, el agente se organiza para capturar los ejemplos que modelan las acciones de control como se representa en la Fig. 4.

De modo que para cada escalón a lo largo de la trayectoria de un disturbio debe producirse una trayectoria de la acción de control que lo compensa y que aún cuando sea no lineal la RNA puede producirla.

$$m = v_{PA}(r) + \sum_{i=1}^n v_i(l_i) \quad (9)$$

Donde: $v_i(l_i)$ son perturbaciones no controladas por otros agentes. Cada función v_i , "nu", en (9) es la contribución de los nodos ocultos a la señal de salida del nodo lineal que produce la VM. Pueden verse como el procesamiento que sufre cada señal al atravesar la red, siendo esas señales el PA y las demás perturbaciones VD_j .

$$VC = [v_{PA}(r) + \sum_{i=1}^n v_i(l_i)] K_m + \sum_{i=1}^n l_i K_i \quad (10)$$

El entrenamiento de la red con los ejemplos del esquema de la Fig. 4 ajusta las funciones v para que la contribución de cada entrada a la red haga $v_i(l_i)K_m + l_i K_i = 0$, asimismo se obtiene seguimiento al punto de ajuste $VC = v_{PA}(r)K_m$. Por lo que, con las perturbaciones canceladas al tiempo que ocurren, la función de la v_{PA} es capaz de llevar el proceso al punto de ajuste deseado:

$$v_{PA}(r) = \frac{PA}{K_m} \quad (11)$$

De este modo aún fuera de las regiones habituales de operación, un agente puede reentrenar su RNA con los datos de los estados operativos que encuentre. Eventualmente puede obtener la respuesta para una gama de combinaciones de las VD, o inferirlas gracias a la capacidad de generalización de la RNA, sin necesitar modificar la estructura de la base de conocimiento. Además de estimar los

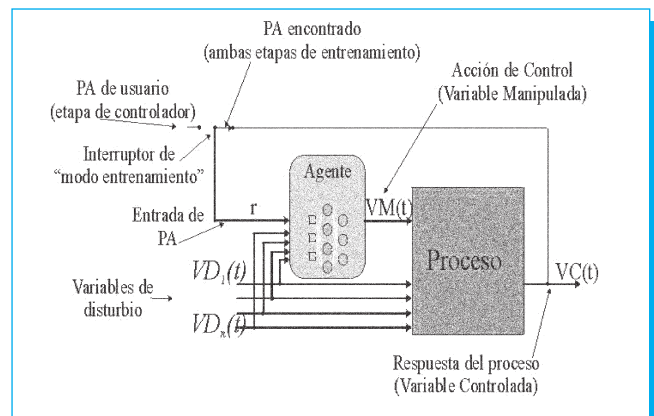


Fig. 4. Empleo de la RNA como máquina de inferencia de las acciones de control.

parámetros K_m y K_i de (10) para simular la respuesta de su VC antes de iniciar una negociación.

Dado que el entrenamiento de la RNA es hecho por el agente mientras interactúa con el proceso, debe hacerse tan rápidamente como sea posible, por lo que se emplea el método de Levenberg y Marquardt [15] por ser más eficiente y rápido que el método de retropropagación convencional. Cada vez que un nuevo estado es capturado el entrenamiento se realiza hasta obtener un error de 0.1 o hasta alcanzar un máximo de 500 épocas.

4.3. Capa de coordinación

En principio, la organización de los agentes es heterárquica debido a que cada entidad permanece controlando su VC sin depender de lo que hagan las demás, al menos mientras no surja una condición que las obligue a depender unas de otras. Ya sea por las condiciones del proceso o por causas humanas, la coordinación entre agentes se vuelve necesaria cuando:

- Un agente va a realizar ajustes a su variable controlada, por lo que avisa a los demás antes de ejercer alguna acción.
- Dos o más agentes controlan variables que se afectan mutuamente en forma antagónica; es decir, cualquier acción que uno haga le afecta al otro y viceversa.
- Un agente no tiene las condiciones de operación para alcanzar su PA, pero alguna de las variables que considera como VD es controlada por otro u otros, por lo que puede "pedir ayuda" negociando algún valor de PA de aquel agente.

Se requiere que cada agente en el entorno conozca a los demás, para lo cual cada uno construye para sí un directorio con los datos de los demás como la VC que tiene a cargo y dirección de Internet (IP) con objeto de comunicarse con ellos en algún momento.

Es importante tomar en cuenta que el instrumento de negociación de un agente, en este entorno, es su VC. Además debe considerarse el caso de que una de las variables que causa problemas no esté en poder de ninguno de los agentes. Para solucionarlo, puede asignarse a un nuevo agente. No obstante, cuando una variable problemática está fuera del alcance de la estrategia de control, los agentes deben adaptarse a ésta.

Dado que algunas variables son más importantes que otras, es necesario que cada agente conozca la importancia de su variable en función del estado del proceso. Esto permite al conjunto jerarquizar sus acciones cuando así se requiere. La noción de su importancia se da al agente mediante la asignación de una prioridad a cada variable. Así, un agente con mayor prioridad asumir el liderazgo en ciertas condiciones, como cambios de PA o disturbios que afectan a una sección específica; mientras que en otras puede ser subordinado, logrando así que el esquema de control se reconfigure dinámicamente. El módulo de coordinación procesa transacciones; es decir, los ciclos de petición/concesión, con otros agentes. Como el procedimiento de petición es independiente del de concesión, cada agente sigue por separado e independientemente dichos procedimientos, lo que le permite hacer diferentes negocios en forma asincrónica y aleatoria. No obstante, cuando ya se acordó con uno o varios agentes y cualquiera de los involucrados decide terminar el acuerdo, las negociaciones se

renuevan con los que se quedan, dándoles aviso y repitiendo las transacciones.

Los acuerdos dependen principalmente de las prioridades asignadas a cada variable y el factor de "máxima desviación aceptable" (MDA) respecto al PA, que permite al agente desviarse del PA en forma temporal y como consecuencia de una negociación. A continuación se presentan los casos básicos de interacción, aunque el mismo mecanismo se puede extender a casos de negociación más complejos entre más de dos agentes.

A) Un agente es afectado por otro: Ejemplo, el agente A_1 , necesita que el agente A_2 ajuste su salida en algún sentido para que no le afecte o para que le ayude. Ya que la VC_2 de A_2 tiene efectos sobre la VC_1 del agente A_1 , el modelo de control (9) de A_1 considerada a VC_2 como una de las entradas a su RNA:

$$VM_1 = v_{PA_1}(r) + v_2(VC_2) + \sum_{i=1}^n v_i(l_i) \quad (12)$$

Donde: VM_1 es la salida de la RNA de A_1 para ejercer acciones de control sobre VC_1 ; $v_2(VC_2)$ es la perturbación causada por las acciones de A_2 . Para el agente A_2 su modelo de control es:

$$VM_2 = v_{PA_2}(r) + \sum_{i=1}^n v_i(l_i) \quad (13)$$

Donde: $v_i(l_i)$ son las variables causantes de disturbios a la VC de A_2 .

De (12) y (13) se observa que A_1 resiente los efectos de A_2 , pero no al contrario. De modo que si los efectos de A_2 causan que la VM de A_1 opere en uno de sus límites reduciéndole el margen de acción, A_1 puede pedir a A_2 negociar, por ejemplo, que A_2 lleve su VC a un valor que le permita a A_1 operar con holgura. El agente solicitante y el que concede siguen estos pasos:

Agente solicitante: Para iniciar la negociación A_1 debe saber cuánto necesita que varíe VC_2 . Para lo cual se simula internamente el modelo de control (9) aplicando su salida a (10):

$$VC_1 = [v_{PA_1}(PA_1) + v_{VC_2}(VC_2) + \sum_{i \neq 2}^n v_i(l_i)]K_m + VC_2K_2 + \sum_{i \neq 2}^n l_i K_i \quad (14)$$

Donde: $i \neq 2$ suponiendo que l_2 es VC_2 , K_i es la ganancia que tiene cada entrada i respecto a VC_1 .

Lo que se espera de simular el efecto de la VC_2 en (14), es que A_1 opere su VM (9) dentro de un rango útil y no sobre los límites; es decir:

$$VMD = \begin{cases} 90\% & \text{si } 95\% < VM_{actual} < 100\% \\ VM_{actual} & \text{si } 5\% < VM_{actual} < 95\% \\ 10\% & \text{si } 0\% < VM_{actual} < 5\% \end{cases}$$

El valor de variable manipulada deseada VM_D , sustituye a VM_1 , lo que está entre corchetes, en (14) para simular los cambios en VC_2 buscando acercarlo lo más posible su VC al PA.

Para no solicitar una desviación de VC_2 muy grande y también para acotar la búsqueda, se limita el margen de variación de la VC_2 simulada a un rango de $\pm 5\%$ del valor actual de la VC_2 , con incrementos

del 0.5% en cada prueba hasta hallar un valor con el que siendo $VM_i \approx VM_D$ se obtenga $VC_i \approx PA_i$. Entonces A_i solicita a A_2 un cambio en PA_2 .

No obstante, si dentro del margen de búsqueda no se encontrara un valor con el que al menos A_i lograra su MDA se usará el mejor valor para hacer la solicitud.

Agente otorgante: En su momento, A_2 recibe la solicitud de un ajuste en su VC y determina hacia dónde se le pide el ajuste; es decir, $\Delta PA_{s_1} = PA_s - PA_{2,n}$. Donde n es el número de transacciones, iniciando en uno, dado que el proceso de negociación se vuelve iterativo si no hay acuerdos:

$$PA_{2,n} = PA_{2,n-1} + \Delta PA_{s_1} \eta^n \quad (15)$$

Donde: PA_{s_1} es el punto de ajuste solicitado por el agente A_1 , ΔPA_{s_1} el incremento en el punto de ajuste solicitado. $PA_{2,n-1}$ el punto de ajuste de la negociación anterior del agente A_2 , en la primera negociación es el PA original. $PA_{2,n}$ el punto de ajuste por ofrecer. η es el factor de ponderación.

El factor $\eta = p_1/p_2$ pondera la prioridad del agente que solicita el ajuste, A_1 , respecto al otorgante, A_2 , implica que:

$$\eta = \frac{p_o}{p_p} = \begin{cases} > 1 & \text{ceder más} \\ < 1 & \text{ceder menos} \end{cases}$$

Donde: p_o es la prioridad del otro agente. p_p la prioridad propia.

Obsérvese también que al aumentar el número de transacciones, el factor de ponderación η^n aumenta si A_2 es menos importante que A_1 ; y viceversa, se reduce cuando A_2 tiene más importancia que A_1 . No obstante, para evitar que un agente de menor prioridad conceda hasta más de lo que se le pide, lo cual ocurriría si $p_o \gg p_p$, el factor de ponderación se limita de la siguiente forma:

$$\eta = \begin{cases} 1 & \text{si } p_o > p_p \\ \frac{p_o}{p_p} & \text{si } p_o < p_p \end{cases} \quad (16)$$

De (16), se ve que un agente de menor prioridad le cederá a uno de mayor importancia todo lo que le pide y la negociación será más rápida porque desde el inicio le concede. Por el contrario, la negociación será más lenta y posiblemente el agente de menor prioridad tenga que conformarse con "algo" de lo que pide al de mayor importancia.

Siguiendo con la negociación del ejemplo, en respuesta A_2 envía el PA al que puede ajustarse en esta primera transacción. Al recibir la respuesta, A_1 la interpreta como la VC que percibirá y la simula en su modelo (12), si con ese valor puede manejar mejor su VM acepta el trato; en caso contrario, repite la petición tal como la hizo originalmente. Lo segundo se determinó así, porque un agente de mayor prioridad concede menos. En este ejemplo si A_2 ofreció un incremento pequeño, obliga a A_1 a volver a pedir. Entonces se debe considerar que A_2 acumula incrementos cada vez más pequeños a su oferta (15), por ello A_1 debe pedir siempre la misma cantidad a modo obtener lo más posible de PA_{s_1} . No obstante, la magnitud de las peticiones no se incrementa porque ello desvirtuaría el sentido de las negociaciones y de las jerarquías entre las variables del proceso.

Sucesivamente, A_1 recibirá una nueva oferta de A_2 y a su vez enviará una nueva solicitud hasta que A_2 le envíe una oferta final. Esta oferta se produce cuando la desviación que tendrá que aceptar A_2 es mayor al MDA.

Cuando ambos agentes concuerdan, asumen como PA el valor acordado, conservando el PA del operario en espera. El módulo de coordinación verifica que las condiciones de operación sean adecuadas para terminar el acuerdo, en cuyo caso avisa al agente correspondiente. Por su parte el agente que aceptó el acuerdo lo conserva hasta que recibe un aviso de terminación del acuerdo.

B) Ambos agentes se afectan mutuamente: Ambos agentes A_1 y A_2 se afectan mutuamente; una acción que toma el agente A_1 afecta a A_2 y al corregir A_2 la desviación que causó A_1 , disturba a la variable de A_1 cerrando un círculo vicioso que produce oscilaciones si no se impide la interacción. En este caso las VC de ambos agentes están consideradas en sus respectivos modelos de control:

$$VM_1 = v_{pA1}(PA_1) + v_{VC2}(VC_2) + \sum_n^i v_i(I_i) \quad (17)$$

$$VM_2 = v_{pA2}(PA_2) + v_{VC1}(VC_1) + \sum_n^i v_i(I_i) \quad (18)$$

Cada uno, por su cuenta, simula internamente su modelo del proceso y el valor final del acuerdo depende de la prioridad de cada agente y del punto de balance alcanzado por (17) y (18). La distribución de las variables del proceso entre diferentes grupos de agentes, debe hacerse identificando la importancia de cada sección del proceso y asignar series de números a los diferentes grupos en función de su importancia así como prioridades consecutivas a los agentes de un grupo.

C) Dos agentes son afectados por un tercero: Es el caso en que un agente controla una variable que afecta a otros, tal vez más de dos. Por ejemplo A_1 necesita que A_2 ajuste su VM en una dirección y A_2 en la dirección opuesta. A_3 recibirá peticiones de los afectados para cambiar su VC. No obstante, dado que no se puede ceder en ambas direcciones, A_3 hace un balance con (19) y responde a cada agente con la última oferta desde la primera solicitud:

$$PA_{p,n} = PA_{p,n-1} + \sum_m^i \Delta PA_{s_i} \eta_i \quad (19)$$

Donde: PA_{s_i} es el incremento en el punto de ajuste del agente p solicitado por el agente i . η_i es el factor de ponderación entre el agente p y el agente i . m es el número de solicitantes.

Lo contrario sería que un agente sea afectado por otros. En ese caso está en ventaja respecto al que tiene que ceder a diferentes peticiones, ya que tiene distintas opciones para pedir ayuda. Para lograr que $VC \approx PA$ puede: a) negociar con uno de ellos, tal vez al que tuviera que pedirle menos, y b) que calcule con (12) la mejor combinación de VC. La segunda opción es más benéfica ya que combinando las acciones de los agentes que le afectan cada uno tiene que ajustarse menos que si se le pide a uno solo.

En este caso, el modelo de control (9) considera en sus entradas los efectos de las respectivas VC, además de las VD. Suponiendo que A_1 está en esta condición:

$$VM_1 = v_{PA_1}(PA_1) + \sum_{i=1}^m v_{VC_i}(VC_i) + \sum_{i=1}^n v_{I_i}(I_i) \quad (20)$$

Por lo que en el modelo (12), considera también los efectos de las demás VC, VD y su propia acción de control:

$$VC_{s_1} = [v_{PA_1}(PA_1) + \sum_{j=1}^k v_{VC_j}(VC_j) + \sum_{i=1}^n v_{I_i}(I_i)]K_m + \sum_{j=1}^k VC_j K_j + \sum_{i=1}^n I_i K_i \quad (21)$$

Donde: VC_{s_1} es la VC que resultará de simular el modelo (21) bajo el efecto de las k VC supuestas y las n VD actuales, que eventualmente se convertirá en el PA_{s_1} por solicitar en (19).

La combinación de VC es representada por un número en el que cada dígito es una VC. Cada combinación se obtiene mediante un conteo con base fraccionaria; es decir, $i = 1/k$; $k = 4, 6, \dots, 10$. Así, la resolución de la búsqueda es ajustada por el agente tomando en cuenta que al aumentar k , aumenta el tiempo que le toma probar cada combinación. Además los valores máximo y mínimo de cada dígito se limitan a $\pm MDA\%$ del valor actual de cada VC. Cada combinación se prueba en (21) y buscando reducir $|PA_1 - VC_{s_1}|$ en cada simulación.

5. Entorno de simulación y prueba

Se utiliza el control en una torre de destilación binaria, Fig. 5, como interfaz entre los cuatro agentes y el proceso. Los cuatro agentes representados en la figura son: agente de temperatura de domo, ATD; agente de composición de domo, ACD; agente de composición de fondo, ACF; agente de nivel de fondo, ANF.

El proceso de destilación consiste en separar los componentes que vienen mezclados en una corriente de carga (alimentación), basado en las diferentes temperaturas a las que cada componente hierve a presión constante.

Se hace referencia a la sección superior como el domo o sección de enriquecimiento o rectificado; a la sección inferior como fondo o sección de agotamiento o separación. El rehervidor situado en el fondo de la torre, utiliza vapor de agua para proveer el calor que realiza la evaporación del producto ligero. El condensador situado en el domo, transfiere el calor del vapor que ha subido desde el fondo a un fluido de menor temperatura, a veces agua, con lo cual reduce drásticamente la temperatura del vapor de modo que se condensa. En procesos con mayor integración energética el calor del vapor se transfiere a líquido de fondo, impidiendo así el desperdicio de energía a cambio de incrementar la complejidad de las interacciones. Después del condensador el producto ligero es almacenado en el tanque de condensado, desde donde puede regresar a la torre como reflujo o salir como producto de domo. Asimismo, una parte del producto pesado que se quedó en la parte inferior, sale como producto de fondo.

El modelo de la torre de destilación que integra el simulador de pruebas consta de 233 platos, cada uno con tres variables de estado. En la tabla 2, se describen las variables visibles en la interfaz del simulador como lo serían a través de la instrumentación del SCD.

6. Experimentos comparativos

En esta fase experimental se emplean dos torres de destilación con las mismas características, la "Torre 1" (T_1) bajo el esquema de

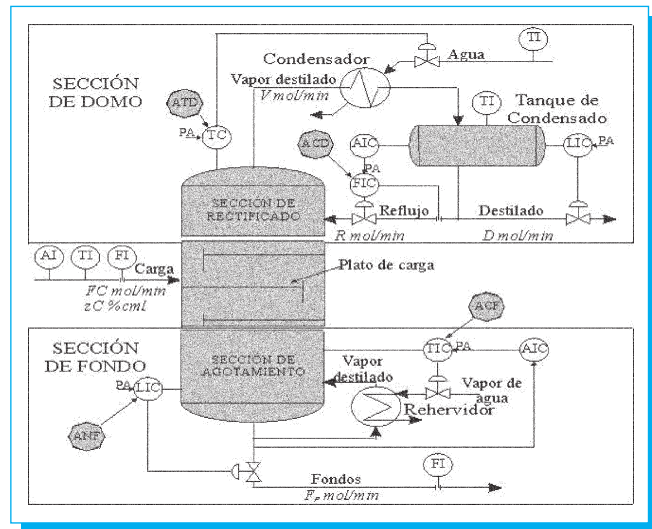


Fig. 5. Diagrama de tubería e instrumentos de la torre de destilación.

control desacoplado y PID y la "Torre 2" (T_2) controlada por los agentes.

El entorno de simulación permite realizar diversos tipos de experimentos como:

- Cambios en el punto de ajuste.
- Inserción de perturbaciones con características como períodos y amplitudes aleatorios o prefijados, etc.
- Simulación de fallas, como atascamiento de válvulas, caída en la presión de los flujos que alimentan la torre, etc.

En los experimentos presentados, se emplean los índices de integral del error absoluto (IEA), integral del error al cuadrado (IEC) para comparar cuantitativamente la eficiencia de los controladores. Asimismo, se utiliza la varianza (σ^2) de la VC y de la VM en cada una de las cuatro VC. Estos dos últimos índices, respectivamente reflejan

Tabla 2. Variables de entrada y salida de la torre de destilación.

SECN.	ENTRADAS	SALIDAS
DOMO	D , Flujo de domo mol/min	ND , Nivel %
	R , Reflujo mol/min	Z_D , Composición %
	F_A , Flujo de agua gal/min	TD , Temperatura de domo °C
CARGA	FC , Flujo de carga mol/min	
	zC , Composición	
FONDO	V , Vapor a rehervidor mol/min	TF , Temperatura °C
	F_F , Flujo de fondo mol/min	Z_F , Composición %
		NF , Nivel de fondo %

la medida en que la VC oscila durante los estados transitorios, mientras que la VM indica la cantidad de movimientos correctivos que el controlador hace con su salida. A mayor cantidad de movimientos es mayor el consumo de energía empleado en las acciones de control. Por razones de espacio se presentan resultados de dos experimentos.

6.1. Cambio de PA a composición de domo (zD).

Descripción del experimento: La siguiente prueba busca comparar la respuesta del ACD con la del controlador PID cuando se cambia el punto de ajuste escalonadamente desde un valor inicial del 75% al 85% y al 95%. Variable afectada: Composición de domo a través de la manipulación de reflujo.

Efecto esperado: Que el ACD aumente el reflujo para enriquecer de ligero del domo de la torre, consecuentemente que la variable controlada cambie siguiendo al nuevo punto de ajuste. Como efecto secundario se espera que el cambio en el reflujo afecte al nivel y la composición de fondo.

Resultados: Los índices de desempeño del experimento se presentan en la Tabla 3 muestran que el error en el control de la T_2 es menor que en T_1 . La varianza de la VC y de la VM muestra que en T_1 la VC tuvo más variación que en T_2 y que el control PID hizo más movimientos correctivos que el agente. Estos resultados se confirman en las gráficas de la Fig. 6.

Las gráficas de las acciones de control correspondientes a cada esquema pueden verse en las figuras 5 y 6, en las que al compararlas se nota que los agentes en conjunto responden simultáneamente al cambio indicado al ACD mientras que en contraste los controladores actúan erráticamente.

En la Fig. 7, se observa que al aumentar el reflujo los demás agentes aumentan su salida en alguna medida, el ANF aumenta el flujo para compensar el incremento de caudal y en respaldo lo hace el ACF al aumentar la temperatura. El ATD aumenta a su vez el flujo de agua para compensar el incremento en la temperatura de produce el ACF. Se nota que las respuestas de los agentes son preventivas a los efectos que tiene el nuevo ajuste.

Tabla 3. Resultados del cambio al PA en composición de domo.

Desempeño del control de composición de domo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	339.9267	1568.1855	1568.1855
T_2	134.4212	1014.6430	1014.6430
Control de temperatura de domo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	37.7509	7.7834	0.0276/0.0099
T_2	7.6650	0.2144	0.0007/0.0001
Control de composición de fondo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	8.6187	0.53186247	0.0018/1.1840
T_2	8.6155	0.46568193	0.0016/1.1840
Control de composición de fondo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	35.017	6.8845	0.0244/0.0866
T_2	6.833	0.2545	0.0009/0.0931

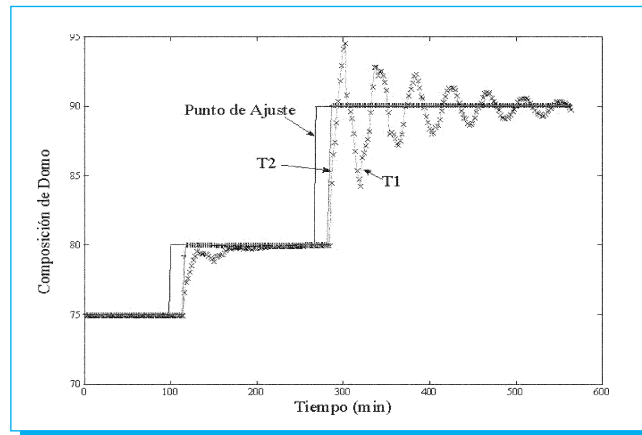


Fig. 6. Efecto de cambios en el punto de ajuste de composición de domo en T_1 y T_2 .

En contraste, como puede verse en la Fig. 8, los controladores PID actúan en forma correctiva al percibir los cambios generados en sus variables controladas. Debido a la baja variabilidad que tienen las acciones de control de los agentes, consumen menos energía al hacer menos movimientos con su VM en comparación con las oscilaciones que produjeron los controladores PID.

6.2. Perturbación en rampa a temperatura de carga

Descripción del experimento: Con este experimento se mide el efecto de una perturbación en la temperatura de la carga sobre las cuatro principales VCs y la correspondiente respuesta de los agentes en comparación con la respuesta de los controladores retroalimentados. La variación en la temperatura de carga es una función rampa con 50% de amplitud pico, el periodo de la oscilación se ajustó a 2900 minutos y se le dejó oscilar por 1.25 ciclos; es decir 3625 minutos (60 horas: 2.5 días). La simulación tardó una hora con treinta minutos.

Variable afectada: Principalmente ambas composiciones.

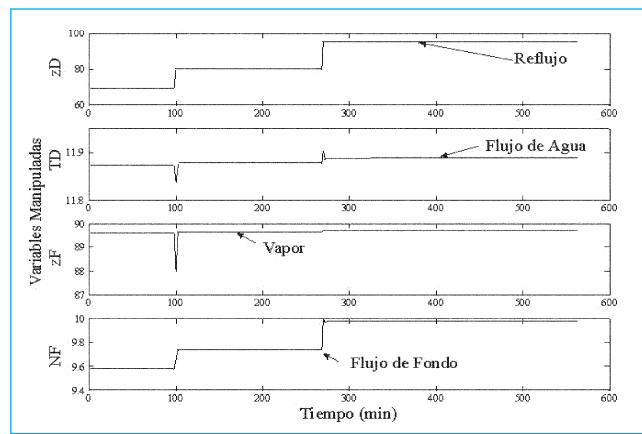


Fig. 7. Acciones de los agentes ante el cambio de punto de ajuste en la composición de domo en T_2 .

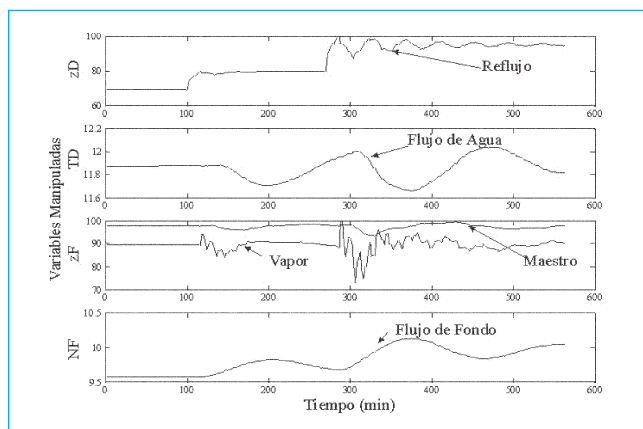


Fig. 8. Reacción de los controladores a los cambios en el punto de ajuste de composición de domo en T_1 .

Efecto esperado: La temperatura de la carga es una variable que afecta directamente a la temperatura de domo y fondo, por lo que se espera que la perturbación sea resuelta principalmente por el ACF mediante ajustes en la temperatura de acuerdo a la variación de entrada, asimismo que el ATD amortigüe las variaciones ajustando el flujo de agua.

Resultados: Los índices de desempeño de controles y agentes durante este experimento se presentan en la Tabla 4. Nuevamente se observa que en T_1 las desviaciones son mayores que en T_2 , además de que los controladores PID producen un número considerable de movimientos con sus VM.

En la Fig. 9 se observan los efectos que la variación de temperatura tiene sobre la composición de domo en cada torre. El efecto causado por la perturbación sobre la composición de fondo se observa en la Fig. 10. En ambas figuras se confirman los resultados vistos en la Tabla 4, que las composiciones en domo y fondo de T_1 se ven más afectadas que en T_2 .

Tabla 4. Condiciones de operación y costos de los generadores.

Desempeño del control de composición de domo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	9.5566	0.61299626	$2.75e-05/36.82$
T_2	0.6275	0.00413361	$0.004/5.58e-05$
Control de temperatura de domo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	2.254	0.06000402	$0.0004/1.20e-05$
T_2	0.246	0.00107254	$7.15e-06/2.76e-06$
Control de composición de fondo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	0.8226	0.01286898	$8.57e-05/0.0371$
T_2	0.2804	0.00073358	$4.89e-06/0.0371$
Control de composición de fondo			
Índices	IAE	ICE	$\sigma^2(VC/VM)$
T_1	3.4276	0.0903384	$0.0006/0.0113$
T_2	0.0397	$4.053e-05$	$2.70e-07/5.2e-07$

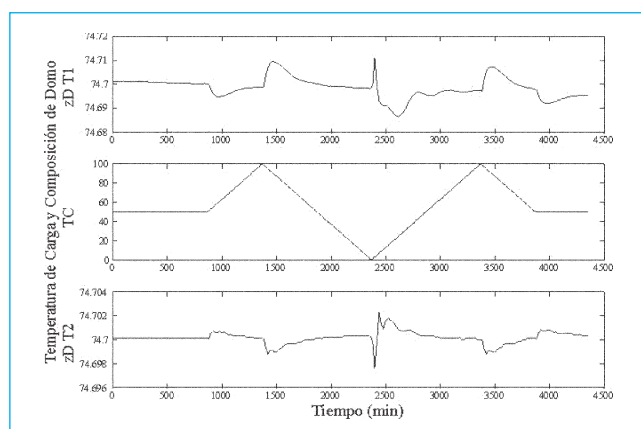


Fig. 9. Efecto de la perturbación en la temperatura de carga sobre la composición de domo de T_1 y T_2 .

Por su parte las acciones de control vistas en la Fig. 11 para los controladores y en la Fig. 12 para los agentes reflejan la considerable cantidad de movimientos que hacen los controladores en comparación con los agentes, obteniendo mejores resultados los segundos.

En la gráfica zF de la Fig. 11 se observa que la señal etiquetada "Maestro" se mantiene sin cambio durante la perturbación, esto se debe a que es el control de temperatura el que resuelve las variaciones antes de que afecten a la composición. No obstante, la razón por la que en las variables de T_1 se nota una variación abrupta antes de los 2500 minutos se debe a que en ese momento los controles que regulan la temperatura en cada sección se contraponen como se ve en las gráficas TD y zF de la misma figura. Al momento en que la temperatura de carga comienza a bajar los dos controladores se ven liberados de la causa que los obliga a mantenerse en ese estado y repentinamente tienen que cambiar la dirección de su acción, el control de TD reduce drásticamente el agua cuando la temperatura de fondo va subiendo lo que produce un incremento en la temperatura de domo. Por el contrario, en T_2 las acciones de los dos agentes que

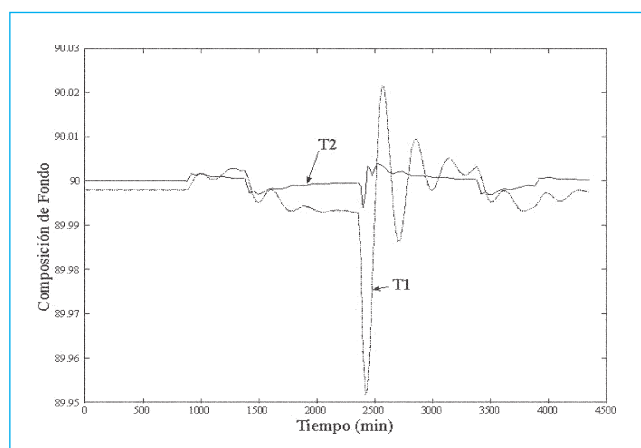


Fig. 10. EComposición de fondo en T_1 y T_2 durante la perturbación.

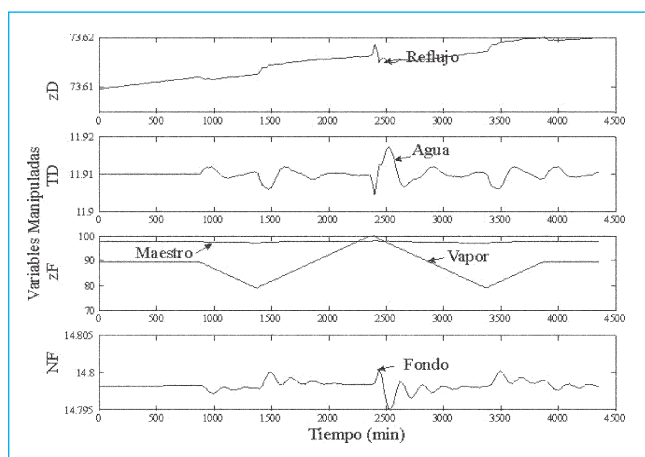


Fig. 11. Acciones de control en T_1 al variar la temperatura de carga.

controlan temperatura van siempre en la misma dirección dado que ambos van balanceando el calor que ingresa en el fondo y el que sale en el domo como se ve en la Fig. 12.

Debido a que en la literatura no se reportan aplicaciones de agentes al control regulatorio del proceso utilizado aquí como banco de prueba, los resultados de esta sección pueden compararse con los obtenidos en [16] y [17], donde se presentan esquemas de control con RNA en columnas de destilación. No obstante, dichos esquemas carecen de la coordinación que permite a los agentes sobreponerse a las interacciones del proceso.

7. Conclusiones

Los experimentos realizados con el esquema de control basado en agentes demuestran que tiene la habilidad de coordinarse para resolver en tiempo real el balance entre las variables de diferentes secciones de la columna de destilación. Individualmente los agentes en este trabajo son controladores reguladores y en conjunto logran mejores resultados que el esquema con lazos independientes. Sus máquinas de inferencia reactivas les permiten tomar decisiones, en este caso bajo las restricciones de tiempo.

Reconocimientos

Esta investigación se realizó en las instalaciones de la Sección de Estudios de Posgrado e Investigación (SEPI) de la Escuela Superior de Ingeniería Mecánica y Eléctrica (ESIME) del Instituto Politécnico Nacional (IPN), y es soportada por el Programa de Investigación en Procesos y Reactores del Instituto Mexicano del Petróleo (IMP). Los autores expresan su gratitud a ambas instituciones.

8. Referencias

[1] S. Skogestad, "Dynamics and Control of Distillation Columns: A Tutorial Introduction," *Trans. IChemE*, 75, Vol. 75, Part A, pp. 539-562, 1997.
 [2] S. Skogestad, I. Postlethwaite, "Multivariable Feedback Control: Analysis and Design", 2nd Ed., Wiley, pp 79-86, 2005.
 [3] S. J. Russell and P. Norvig, "Artificial Intelligence, a Modern Approach", Prentice Hall, New Jersey, USA, pp. 151 & 525, 1995.

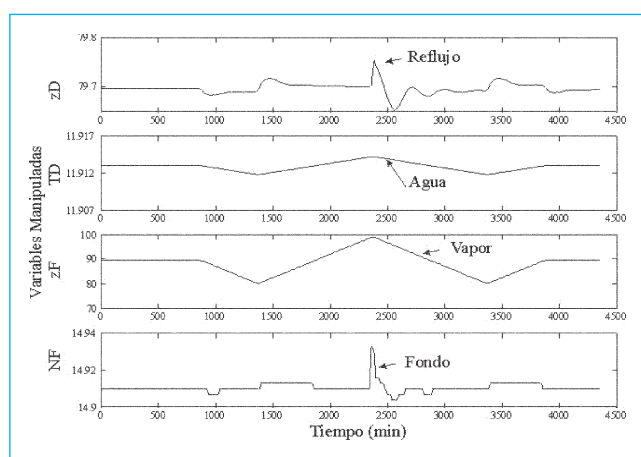


Fig. 12. Acciones de control en T_2 durante la perturbación en la temperatura de carga.

[4] V. Marik & D. McFarlane, "Industrial Adoption of Agent-Based Technologies", *IEEE Intelligent Systems*, Vol. 20, no. 1, pp. 27-35, 2005.
 [5] N. R. Jennings, J. M. Corera, I. Laresgoiti, E. H. Mamdani, F. Perriollat, P. Skarek, L. Z. Varga, "Using ARCHON to Develop Real-World DAI Applications for Electricity Transportation Management and Particle Accelerator Control", *IEEE Expert*, Vol., 11, No., 6, pp., 64-70, 1996.
 [6] H. Ni, H. G. Thomas, "Power System Stability Agents Using Robust Wide Area Control", *IEEE Transactions on Power Systems*, Vol., 17, No. 4, pp. 1123-1131, 2002.
 [7] A. Romanenko & L. O. Santos, P.A.F.N.A. Afonso, "Application of agent technology concepts to the design of a fault-tolerant control system", *Control Engineering Practice*, Vol. 15, No. 4, pp. 459-469, 2007.
 [8] E. Tataru, I. Birol, F. Teymour, A. Cinar, "Agent-Based Control of Autocatalytic Replicators in Networks of Reactors", *Computers and Chemical Engineering*, Vol., 29, No., 4, pp., 807-815, 2007.
 [9] M. Cerrada, J. Cardillo, J. Aguilar, R. Faneite, "Agents-based design for fault management systems in industrial processes", *Computers in Industry*, 58, 4, 313-328, 2007.
 [10] F. Bellifemine, G. Caire, T. Trucco "JADE Programmer's Guide," Disponible en: <http://jade.tilab.com/papers-index.htm>
 [11] R. Batres, S. P. Asprey, Fuchino T., & Naka, Y., "A KQML Multi-Agent Environment for Concurrent Process Engineering", *Computers & Chemical Engineering*, Vol. 23, pp. s653-s656, 1999.
 [12] P. Maes, "Modeling adaptive autonomous agents," *Artificial Life Journal*, no. 1, vol. 1-2, MIT Press, pp. 135-162, 1994.
 [13] C. J. C. H. Watkins and P. Dayan, "Q-Learning", *Machine Learning*, Vol. 8, no. 3-4, pp. 279-292, 1992.
 [14] M. Norgaard, O. Ravn, N. K. Poulsen and L. K. Hansen, "Neural Networks for Modelling and Control of Dynamic Systems", Springer Verlag, pp. 25-33, 2000.
 [15] Hagan M. T., and Menhaj M., "Training Feedforward Networks With The Marquardt Algorithm", *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989-993, 1994.
 [16] Yeh T.-M., Huang M.-C., Huang C.-T., "Estimate of Process Compositions and Plantwide Control From Multiple Secondary Measurements Using Artificial Neural Networks", *Computers and Chemical Engineering*, vol. 27, no. 1, pp. 55-72, 2003.
 [17] L. Fortuna, S. Graziani, M.G. Xibilia, "Soft Sensors for Product Quality Monitoring in Debutanizer Distillation Columns", *Control Engineering Practice*, vol. 13, no. 4, pp. 499-508, 2005.