

Integración de técnicas de Inteligencia Artificial para generar crónicas deportivas

Belén Díaz-Agudo & Pablo Gervás

Dep. Sistemas Informáticos y Programación

Universidad Complutense de Madrid

e-mail: {belend,pgervas}@sip.ucm.es

Resumen

Este artículo describe una aproximación que integra razonamiento basado en casos, razonamiento con reglas, representación de información textual y ontologías para generar automáticamente crónicas escritas, utilizando los datos objetivos que describen un partido de fútbol. La propuesta se enmarca dentro de la asignatura Ingeniería de Sistemas Basados en Conocimiento (ISBC) de la Ingeniería en Informática de la Universidad Complutense de Madrid (UCM).

1. Introducción

El procesamiento del lenguaje natural es una de las áreas de aplicación clásica de la Inteligencia Artificial. En este campo intervienen aspectos tan importantes y cuestionados como el de la creatividad de los programas informáticos. Partiendo de unos datos iniciales, ¿puede un programa transformarlos de forma que se componga un texto original? por ejemplo, una historia literaria interesante, una poesía, una crónica deportiva... [5].

Son muchos los dominios en los que se requiere la generación automática de textos. En este artículo abordaremos la resolución de un problema en un dominio real: generación de crónicas deportivas para partidos de fútbol.

El problema se corresponde con la última práctica que resolvieron los alumnos del curso 2003/2004 de Ingeniería de Sistemas Basados en Conocimiento (ISBC), una asignatura optativa de 9 créditos de la Ingeniería Informática de la Universidad Complutense de Madrid, impartida por la primera autora de este artículo.

Durante los tres años que lleva impartándose esta asignatura, su objetivo ha sido servir como un “laboratorio de Inteligencia Artificial” donde la componente práctica sea el elemento fundamental y los alumnos aprendan a resolver problemas utilizando técnicas de Inteligencia Artificial (IA). Las técnicas

básicas se aprenden en la asignatura troncal de Inteligencia Artificial (Inteligencia Artificial e Ingeniería del Conocimiento, IAIC) y las técnicas más avanzadas en la propia asignatura ISBC¹.

Nuestra propuesta se basa en que las prácticas de ISBC no deben tener como objetivo únicamente el de consolidar el aprendizaje de una técnica de IA, ya que de eso se habrán encargado los ejercicios y prácticas opcionales que se plantean en la asignatura troncal IAIC.

El objetivo principal de la asignatura es que los alumnos sean capaces de integrar y comparar distintas técnicas, así como valorar la adecuación de una u otra técnica, lo que es un aspecto fundamental en una carrera de ingeniería. Para ello los profesores deben elegir aplicaciones adecuadas que permitan la combinación de distintas técnicas de IA para solucionar un problema. Encontramos varios ejemplos de este tipo de aplicaciones relacionados con el dominio del fútbol [4].

En concreto en el planteamiento de las prácticas los alumnos deberán evaluar qué técnica es más adecuada para resolver cada problema, cuáles son sus limitaciones, y cómo una técnica concreta se puede complementar con otra.

La estructura de cada una de las prácticas (menos la última) dentro del mismo curso de la asignatura es la misma: se ofrecen varias técnicas (normalmente 2 ó 3) y los alumnos deben elegir sólo una de ellas, de forma justificada, para resolver un problema dado. Por ejemplo, filtrado de correo spam utilizando el paradigma basado en reglas vs. el paradigma basado en casos (CBR); o, clasificación de páginas Web utilizando técnicas de Recuperación de Información vs. técnicas de CBR Textual; o, diseño de los jugadores de un equipo de fútbol utilizando el simulador SoccerBots².

El objetivo de la última práctica es la integración de diversas técnicas en la resolución de un problema

¹<http://www.fdi.ucm.es/profesor/belend/ISBC/isbc.html>

²<http://www.teambots.org/>

“real”. En concreto, para el problema propuesto –la generación de crónicas deportivas– los alumnos utilizaron de forma integrada el razonamiento basado en reglas, el razonamiento basado en casos, las ontologías y técnicas de tratamiento de lenguaje natural.

Los siguientes apartados describen el problema, los objetivos docentes, los criterios de evaluación y distintas alternativas de resolución del mismo. Terminamos el artículo con una descripción de las conclusiones generales extraídas.

2. Descripción del problema

La práctica consiste en diseñar e implementar un sistema para la generación automática de textos. Son varios los objetivos que se pretenden cubrir:

- Diseñar una ontología para un dominio dado y construir un sistema que aproveche la estructura ontológica creada.
- Comparar los dos paradigmas de razonamiento estudiados: reglas vs. casos.
- Preparar el corpus de textos que se utilizarán como plantillas.
- Diseñar una interfaz adecuada.

La entrada que recibe el sistema son los datos sobre un cierto partido de fútbol junto con ciertas características contextuales. Por ejemplo, nombre de los equipos, distinguiendo el local del visitante, nombre del árbitro y características del arbitraje (¿fue justo? ¿algún dato a reseñar?), nombre del estadio³ y características como su capacidad, el número de espectadores, estado del terreno de juego, el clima, resultado del partido, qué jugadores marcaron, en qué minuto del partido, en qué tipo de jugada –penalti, saque de esquina, contraataque,...– faltas, tarjetas (rojas y amarillas), cambios entre jugadores, ...

Se distinguen distintos tipos de eventos en el juego (cambio, falta, tarjeta, gol) y para cada uno se indica el minuto en el que tuvo lugar y las características específicas que lo describen.

La salida que queremos obtener es una descripción textual detallada de la información de entrada. Como se busca un texto escrito, no son relevantes la pronunciación de las palabras, la entonación del texto, o las pausas.

Para esta práctica queremos llevar a cabo una aproximación CBR en la que la adaptación se hace por

³Se podría inferir del contexto, al saber el nombre del equipo local

sustitución (o instanciación de las plantillas). Estas sustituciones se pueden plantear como un sistema de reglas.

La generación de frases correctas sintácticamente requeriría el uso de una gramática. Como el desarrollo de una gramática exhaustiva queda fuera de los objetivos de esta práctica utilizaremos plantillas, es decir, textos (frases, párrafos o textos completos) preescritos.

El enunciado no está fijado de forma estricta, sino que se permite que los alumnos particularicen ciertos puntos de variación. En concreto, los alumnos deciden qué **casos** (crónicas de fútbol) incluir y su representación, tanto la representación persistente como la representación en memoria, la estructura para indexarlos y recuperarlos, por ejemplo, organizar en párrafos la información de los distintos tipos de eventos (goles, faltas, tarjetas, ..). Además deberán implementar los procesos CBR[1] de recuperación y adaptación que resuelven el problema planteado.

3. Objetivos docentes y Criterios de evaluación

Uno de los mayores problemas asociados al diseño de sistemas basados en conocimiento es el famoso “cuello de botella de la adquisición de conocimiento” [3]. Un problema paralelo se plantea a la hora de diseñar los enunciados de las prácticas de la asignatura ISBC.

En una primera aproximación, llevada a cabo en el curso 2002/2003, los alumnos debían plantear problemas (de tema libre) y desarrollar sistemas basados en conocimiento para resolverlos. El objetivo subyacente era que los alumnos pudieran elegir un dominio en el que les resultara fácil encontrar un experto dispuesto a atender sus necesidades de conocimiento. En opinión de la profesora de la asignatura, el resultado en líneas generales no satisfizo los objetivos docentes de comprensión de la problemática de la adquisición de conocimiento. Casi todos los expertos eran “expertos simulados” o amigos que no planteaban los problemas típicos de los expertos reales. Sin embargo, la experiencia sí resultó positiva en cuanto al aprendizaje sobre la necesidad de representar de una forma adecuada el conocimiento del dominio disponible, a partir de unas u otras fuentes.

En los cursos posteriores, y esperando obtener los mismos resultados, se ha planteado una fase de ad-

quisición de conocimiento basada directamente en fuentes de conocimiento cercanas. Es decir, se han planteado prácticas en dominios familiares para los alumnos, de forma que se utilice como origen el propio conocimiento de los alumnos sobre el tema, así como la consulta y filtrado en la Web.

Por ejemplo, en la práctica que se describe en este artículo, el dominio del fútbol es bien conocido por los alumnos, en general, y les ha permitido el desarrollo de ontologías completas comprendiendo las dificultades asociadas a la representación del conocimiento. Además, la adquisición de crónicas textuales de partidos anteriores, es fácil a través de las páginas Web de periódicos electrónicos, aunque los alumnos deben resolver el problema de cómo representarlas de una forma adecuada para el problema en cuestión, y cómo integrarlas con la ontología que habrán desarrollado de forma previa e independiente a las crónicas concretas.

El grupo de ISBC del curso 2003/2004 contó con 53 alumnos, que trabajaron en grupos de 2 ó 3 personas. La profesora de la asignatura evaluó las 14 prácticas entregadas, según los criterios siguientes:

- Interfaz, facilidad para usar y ampliar el sistema (por ejemplo, añadir más casos)
- Adecuación de la representación interna elegida

- Número y calidad de casos
- Estructura para clasificar los casos
- Persistencia de los casos y de las crónicas generadas
- Conocimiento ontológico del dominio. Nivel de reusabilidad del conocimiento para otras aplicaciones. Se pide distinguir entre el conocimiento común e intrínseco a la aplicación y el conocimiento que variaría si por ejemplo, queremos usar la aplicación en otra liga de fútbol de otro país o entre países (en campeonatos mundiales)
- Documentación (descripción de las dificultades encontradas, puntos fuertes y débiles, conclusiones generales).

4. Descripción de la arquitectura

La Figura 1 esquematiza la arquitectura de sistema CBR que hace uso de los ingredientes que se describen en los apartados siguientes.

La entrada al sistema es una descripción de la consulta, es decir, de los datos del partido. La interfaz

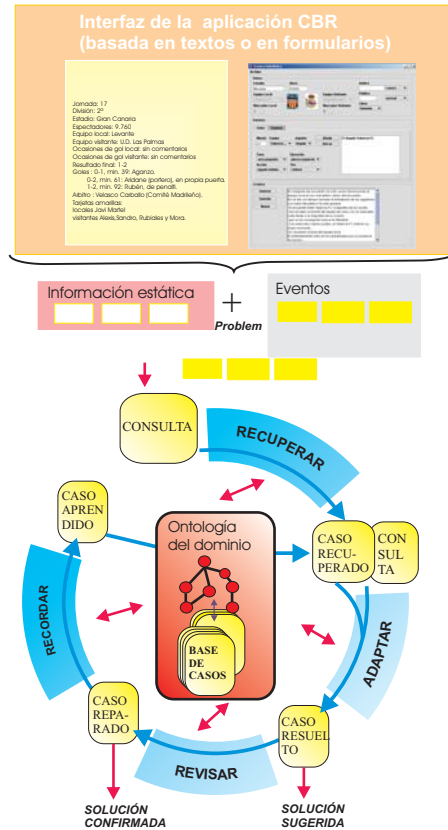


Figura 1: Arquitectura del sistema

puede estar basada en formularios o en textos que seguirán una cierta estructura. Se puede usar texto plano o algún tipo de formato estándar como XML. En la consulta podemos distinguir entre:

- Datos estáticos del partido. Por ejemplo, los datos de los equipos que juegan, el nombre del estadio y del árbitro, clima, hora, número de espectadores, ...
- Eventos. Por ejemplo, los goles, faltas, tarjetas u otros incidentes a resaltar. Cada evento lleva asociado el minuto en el que ocurrió, los actores involucrados y el subtipo concreto de evento. Por ejemplo, los subtipos del evento tarjeta son tarjeta roja y tarjeta amarilla, y los subtipos de goles son: penaltis, gol en propia puerta, gol de cabeza, ..

El sistema CBR dispone de una base de casos previos ya resueltos. Cada caso describe los datos de un

partido de fútbol y una crónica textual que lo resume. Los procesos del ciclo CBR [1] se encargan de recuperar el caso más similar a la consulta y adaptar la crónica para que refleje los datos del partido consulta. La crónica generada es revisada por el usuario.

El plan de trabajo consta de dos fases. Una primera fase de adquisición de conocimiento, en la que se diseñará la ontología del dominio y se recopilarán las crónicas que formarán la base de casos, y una segunda fase en la que se implementarán los procesos CBR, haciendo hincapié en la recuperación y adaptación de casos.

Al terminar la primera fase los alumnos entregarán la documentación correspondiente a la adquisición del conocimiento que debería ser, más o menos, independiente del desarrollo de los procesos que lo manejan. En concreto se entregará una primera versión de la ontología del dominio y un conjunto de textos que describan crónicas de fútbol. De esta forma se quiere evitar que el conocimiento adquirido sea dependiente de la implementación de la aplicación CBR. Los subapartados siguientes describen estas dos fases.

4.1. Fase de adquisición de conocimiento

En esta fase los alumnos deben diseñar una ontología sobre el dominio del fútbol, recopilar una serie de instancias de los conceptos de la ontología, y construir un corpus de crónicas de partidos de fútbol.

La *ontología sobre el dominio del fútbol* incluirá la terminología básica que aparecerá en los textos de las crónicas y en la descripción de la consulta. Por ejemplo, se deben incluir términos como: equipo local, visitante, resultado, balón, jugadores, gol, pelota, entrenador, centrocampista, portero, portería, tarjeta roja, penalti, fuera de juego,

En la aplicación la ontología se utilizará principalmente como fuente de datos para rellenar los campos de los formularios, y durante la generación del texto para encontrar información relacionada, contextual o curiosidades. Por ejemplo, dado el nombre de un equipo, la ontología permite tener acceso a todos sus jugadores; o dado el nombre de un estadio, la ontología tiene información sobre su capacidad y la ciudad en la que está.

En la ontología se separa entre el *conjunto de términos*, clases y relaciones, y el *conjunto de individuos* que son instancias de las clases anteriores y se relacionan con otros individuos.

Para modelar la ontología se propone que los alumnos elijan entre dos herramientas bien conocidas: Protégé⁴ y WebODE⁵. Ambas herramientas proporcionan entornos gráficos para definir clases (en términos de sus atributos y relaciones), jerarquías de clases, atributos y relaciones, restricciones en las relaciones (dominio, rango, cardinalidad, ..)

Se propone dividir la ontología en varios tipos de elementos que representan los elementos del dominio, que incluyen términos que pueden aparecer en los textos, por ejemplo, personas como jugadores (Portero, Delantero, ...), entrenador, árbitro, etc., objetos, como pelota, lugares, como área, centro del campo, y tipos de eventos como gol, falta o tarjeta.

Para su uso en el sistema CBR la ontología será exportada a clases Java. Tanto Protégé como WebODE ofrecen facilidades para exportar la taxonomía conceptual de una ontología a una jerarquía de clases Java.

Después de modelar el conocimiento terminológico del dominio se completará con un conjunto de *instancias*. Por ejemplo, el conjunto de instancias denominado "Liga Española Masculina 2003/2004" incluirá los individuos para representar los equipos, jugadores y árbitros que participaron en la Liga española de fútbol de la temporada 2003/04. El objetivo es que un cambio en un conjunto de instancias debería ser transparente tanto a la parte de conocimiento terminológico como a los procesos de la aplicación CBR.

El *corpus de crónicas de partidos de fútbol* recopilará un conjunto de textos que describan partidos de fútbol de la liga masculina de fútbol de la temporada 2003/04. Estos textos servirán para construir los casos que se reutilizarán en el sistema CBR para generar crónicas de partidos nuevos.

Como fuente de conocimiento se utilizarán páginas Web principalmente de periódicos electrónicos⁶.

4.2. ¿Cómo funciona? El ciclo CBR

Para lograr la generación del texto a partir de un conjunto de datos iniciales se seguirán las fases siguientes:

1. Recuperar un caso similar a la consulta

⁴<http://protege.stanford.edu/>

⁵<http://delicias.dia.fi.upm.es/webODE/>

⁶<http://www.publicacionesdelosur.net/comun/deportes/>, <http://www.noticias.info/>, <http://mensual.prensa.com/>

2. Adaptación de la plantilla rellenando los huecos con los valores de la consulta
3. Revisar el texto generado
4. Escribir los textos en el fichero de salida

Los siguientes apartados describen las fases anteriores, comenzando por un aspecto del que dependen los procesos CBR: la representación e indexación de los casos.

4.2.1. Representación e indexación de Casos

Representación de la base de casos para ser manejada por la aplicación CBR.

Los casos constan de dos partes: descripción y solución. La descripción del caso incluye las características del partido cuya crónica está parafraseada en la solución del caso. La solución del caso es una plantilla, es decir, una representación del texto de la crónica que incluye ciertos "huecos" que se pueden rellenar utilizando la información dada en la descripción del caso.

Por ejemplo, en el fragmento de plantilla “*jugador* recibió una tarjeta *color*”, los huecos *jugador* y *color* se corresponden con clases de la ontología que pueden ser rellenados con valores concretos, instancias concretas de jugador y de color.

Los alumnos deben elegir y valorar si es mejor una representación de casos de granularidad gruesa o fina. Los casos de granularidad gruesa representan la descripción de partidos completos y pueden estar formados de otros casos de granularidad más fina, que describen aspectos específicos de un partido: los goles, las faltas, las tarjetas,...

La elección influye principalmente en dos aspectos: cómo indexar los casos y cómo combinarlos de forma ordenada para componer un discurso coherente.

Si se usan casos de granularidad fina, cada plantilla puede ser indexada según un cierto tema, en este ejemplo tarjetas, y por el tipo que corresponda a ese tema, por ejemplo, partido violento (si hay muchas tarjetas). En el caso de la granularidad gruesa, se indexa por las características del partido: victoria abultada local, partido violento, ...

Casi todos han incluido una clasificación relativa al resultado: Empate (Con muchos o pocos goles), Gana Local (Ventaja Abultada, Poca Ventaja) o Gana Visitante (Ventaja Abultada, Poca Ventaja)

Se pueden añadir niveles adicionales de clasificación de casos dependiendo del número de faltas, cambios, tarjetas, ... En vez de crear clases para organizar los casos se pueden incluir estos criterios en una medida de similitud numérica que se computa entre la consulta y cada uno de los casos.

Respecto a la estructura del texto, en los casos de granularidad gruesa, que describen un partido completo, la estructura del discurso está fijada por el propio caso. Sin embargo, una aproximación más flexible, la elegida por la mayoría de los alumnos, permite combinar distintos casos (párrafos) para formar un texto coherente. En esta situación se debe decidir cómo representar los elementos de la estructura del discurso o secuencia de escenas. Cada escena se describe en un párrafo de texto.

Puede haber distintas secuencias para ordenar las escenas, por ejemplo:

- Introducción, desarrollo del juego durante la primera parte, desarrollo durante la segunda parte, resultado final, comentarios generales y conclusión.
- Introducción, descripción de la actuación del árbitro, descripción de los goles, descripción de tarjetas, comentarios generales y conclusión.

4.2.2. Recuperación

De las diversas posibilidades, la opción más utilizada por los alumnos ha sido la de generar la información de cada tipo de evento en un párrafo aparte: un párrafo sobre goles, párrafo sobre tarjetas, párrafo sobre árbitro y público.

Esto da pie a qué pueda haber varias bases de casos de granularidad fina y no sólo una de casos complejos. El esfuerzo inicial de adquisición de casos es algo mayor, porque hay que disgregar las crónicas completas, pero se consigue mayor flexibilidad de uso de los casos y la necesidad de manejar un número menor de ellos.

La recuperación requiere dos pasos. Primero se selecciona la secuencia de escenas a seguir para desarrollar el argumento y posteriormente se recupera un caso para cada escena del tipo adecuado.

Para la recuperación de casos se utilizarán las estructuras de indexación descritas para hacer un filtrado inicial de los casos que cumplan unas ciertas características generales (por ejemplo, empate). El conjunto de casos resultantes se ordena teniendo en

cuenta el resultado de una función de similitud numérica que se usa para comparar la consulta y los casos.

4.2.3. Adaptación

Para esta práctica queremos llevar a cabo una aproximación CBR en la que la adaptación se hace por sustitución (o instanciación de las plantillas).

Esta instanciación se puede llevar a cabo usando un sistema de reglas de sustitución para lo que algunos de los grupos eligieron una implementación con JESS⁷. Otras aproximaciones más simplistas únicamente sustituyen el hueco de la plantilla (que representa un concepto o clase de la ontología) por una instancia de ese concepto que se construye a partir de los datos dados en la consulta.

La corrección semántica de las frases puede ser "garantizada" por los roles de los distintos elementos involucrados en una escena (que vienen dados por las clases de la ontología del dominio a las que pertenecen los elementos), por ejemplo, lugar, personaje, ...

4.2.4. Aprendizaje

El razonamiento basado en casos es, además de un paradigma de razonamiento, un paradigma de aprendizaje. En un sistema CBR típico (Ver Figura 1)

La pregunta obvia que nos debemos plantear en un sistema CBR como el realizado en esta práctica es si merece la pena aprender las crónicas que generamos por adaptación. Realmente la respuesta debería ser no, porque las crónicas generadas son muy similares a las existentes (son una simple instanciación). El aprendizaje real consiste en añadir nuevas plantillas que sean diferentes de las que ya hay almacenadas.

5. Ejemplos

Mostramos un pequeño ejemplo de una de las plantillas guardadas en archivos de texto (con estructura propia) que se utilizan.

La figura 4 muestra un ejemplo de consulta. Hay muchas posibilidades de parafrasear los mismos datos. Se presentan a continuación dos ejemplos de crónicas para la misma consulta.

Unos 9.760 espectadores acudieron hoy al Estadio de Gran Canaria para ver este encuentro

⁷<http://herzberg.ca.sandia.gov/jess/>

Métodos innovadores aplicados a distintas disciplinas



Figura 2: Interfaz basada en formularios

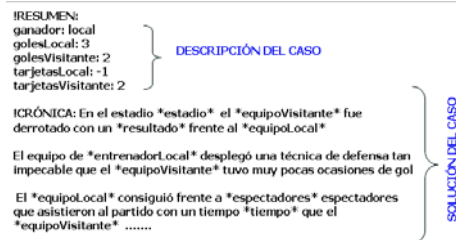


Figura 3: Ejemplo de plantilla

de la decimoséptima jornada de la Liga de fútbol en la Segunda división en el que se enfrentaban los equipos de Levante y Unión Deportiva Las Palmas. Fue un encuentro con bastantes altercados, ya que el colegiado del encuentro Velasco Carballo del comité madrileño, mostró la tarjeta amarilla al jugador local Javi Martel, así como a sus rivales Alexis, Sandro, Rubiales y Mora. En la primera parte del partido se marcó un solo gol por parte del equipo visitante, que aumentó la distancia gracias al portero Aridane que anotó un gol en propia puerta en el minuto 16 de la segunda parte. Los locales pudieron acortar distancias gracias a un penalti que convirtió Ruben cuando ya pasaban 2 minutos del tiempo reglamentario. Resultado final del encuentro, Levante 1, U.D. Las Palmas 2.

El Levante salió al césped del Estadio de Gran Canaria con la lección bien aprendida, querían

Jornada	17
División	2ª
Estadio	Gran Canaria
Espectadores	9.760
Equipo local	Levante
Equipo visitante	U.D. Las Palmas
Ocasiones de gol local	sin comentarios
Ocasiones de gol visitante	sin comentarios
Resultado final	1-2
Goles	0-1, min. 39 Aganzo. 0-2, min. 61 Aridani (pp) 1-2, min. 92 Rubén (p)
Arbitro	Velasco Carballo
Tarjetas amarillas	Locales: Javi Martel Visitantes: Alexis, Sandro, Rubiales y Mora

Figura 4: Ejemplo de consulta

encarrilar rápido el partido y así lo hicieron. Ya en la primera parte (min.40) los locales marcaron el 0-1 gracias a un gol de Aganzo. El juego desplegado por los locales era imprevisto para los isleños que no podían cortar las jugadas de peligro de su rival. La segunda mitad comenzó con el mismo dibujo, el Levante salió a por la victoria y fue Aridani (min.60) en propia puerta el que anotó el 0-2 para los visitantes. Aunque la mala suerte se apoderó de los locales, no dieron el partido por perdido y acertaron distancias en el 92 gracias a un tanto que Rubén consiguió de penalti.

6. Evaluación de la experiencia

En comparación con el funcionamiento de la misma asignatura en años anteriores no se han apreciado diferencias significativas en el rendimiento académico de los alumnos en términos de porcentajes de aprobados o calificaciones medias obtenidas. Sin embargo, cabe destacar que se ha observado a lo largo del curso una motivación mucho mayor de los alumnos a la hora de trabajar en las prácticas propuestas. Esto se ha reflejado especialmente en dos factores fácilmente observables: con respecto a años anteriores se ha observado un incremento en la asistencia a tutorías de los alumnos matriculados, y una notable voluntad de dedicar tiempo a la práctica propuesta más allá de lo estrictamente requerido por el profesor (ha habido grupos que para esta práctica han representado mediante los formalismos propuestos las alineaciones completas de todos los equipos de la liga). Está claro que, impulsados por su notable pasión futbolística, los alumnos participantes han vivido con pasión el esfuerzo de familiarizarse con tecnologías nuevas, de comprender el funcionamiento de las técnicas propuestas, y de aprender a poner al servicio de sus intereses los esquemas teóricos de la asignatura.

El problema propuesto está en la línea de algunos de los

“nifty assignments”⁸ que se han propuesto en las sesiones anuales del congreso SIGCSE. En concreto, cabe destacar los paralelismos entre esta propuesta la similitud con las propuestas *The Random Sentence Generator*, de Julie Zelenski (SIGCSE 1999) – en la que se propone un uso de gramáticas específicas para textos en dominios concretos similar a las plantillas que se utilizan en nuestra propuesta –, y *Random Writer*, de Joe Zachary (SIGCSE 2003) – en la que se aborda una generación aleatoria basada en el estudio probabilístico de textos de muestra.

Asimismo, podrían considerarse ampliaciones posibles de esta propuesta que tuviesen en cuenta las directrices generales para la construcción de sistemas de generación de textos descritas en el trabajo de Bringsjord y Ferrucci [2], que, aunque a una escala muy alejada de una simple práctica universitaria, abordan el desafío de escribir programas dotados de creatividad literaria.

7. Conclusiones

El procesamiento del lenguaje natural es una de las áreas de aplicación clásica de la Inteligencia Artificial. Son muchos los dominios en los que se requiere la generación automática de textos. En este artículo hemos descrito cómo generar crónicas deportivas escritas para partidos de fútbol.

El problema se corresponde con la última práctica que resolvieron los alumnos del curso 2003/2004 de la asignatura Ingeniería de Sistemas Basados en Conocimiento de la UCM. El objetivo docente que se cubre con la práctica es la resolución de un problema complejo integrando varias técnicas de Inteligencia Artificial. En concreto, se ha utilizado razonamiento basado en casos, razonamiento con reglas, representación de información textual y ontologías.

El dominio de aplicación ha resultado muy motivador para los alumnos y ha resuelto un problema típico presente en las prácticas de la asignatura ISBC: el problema de encontrar un “experto” real. La fase de adquisición de conocimiento ha utilizado como fuente el propio conocimiento de los alumnos sobre el tema. También la consulta de algunas páginas Web.

De los resultados obtenidos podemos resaltar la motivación de los alumnos en la resolución del problema, el uso integrado y complementario de distintas técnicas de Inteligencia Artificial en la resolución de un problema real, y la comprensión de las dificultades en la Ingeniería del Conocimiento. Los alumnos han aprendido a separar la parte del conocimiento que es común, reutilizable e intrínseco del dominio y la parte del conocimiento que es específica y dependiente de la aplicación.

⁸<http://nifty.stanford.edu/> (visita 21/04/2005)

Referencias

- [1] A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(i), 1994.
- [2] S. Bringsjord and D. A. Ferruci. *Artificial Intelligence and Literary Creativity. Inside the Mind of BRUTUS, a Storytelling Machine*. Lawrence Erlbaum, 2000.
- [3] J. C. Giarratano and G. Riley. *Expert Systems*. PWS Publishing Co., 1998.
- [4] B. López, M. Montaner, and J. L. de la Rosa. Utilización de un simulador de fútbol para enseñar inteligencia artificial a ingenieros. In *Proc. JENUI, VII Jornadas de Enseñanza Universitaria de la Informática, Palma de Mallorca*, 2001.
- [5] E. Reiter and R. Dale. *Building Natural Language Generation Systems*. Cambridge University Press, 2000.