

El sistema de búsqueda de respuestas del que se parte participa en CLEF desde el año 2003, a través del grupo de investigación MIRACLE (Multilingual Information Retrieval for the CLEF campaign) (de Pablo-Sánchez et al., 2007). En este trabajo el sistema monolingüe, desarrollado inicialmente para trabajar en español, se ha evaluado a través de una batería de preguntas extraídas de la colección de CLEF 2008. Otros trabajos previos han abordado también este problema para el tratamiento de otros idiomas, como en (Hartrumpf y Leveling, 2006) o en (Moldovan y Clark, 2005).

El artículo está estructurado como sigue: en la sección 2 se describen las bases para el tratamiento de información temporal. En la sección 3 se presenta el sistema de búsqueda de respuestas sobre el que se ha trabajado, mientras que en la sección 4 se detallan las mejoras llevadas a cabo sobre el mismo. La sección 5 muestra los resultados de la evaluación a la que se ha sometido al sistema. Para finalizar, la sección 6 incluye las conclusiones obtenidas y algunas líneas de trabajo futuro.

2 Reconocimiento y Normalización de Expresiones Temporales

En un SBR es primordial poder resolver referencias que ayuden a responder a cuestiones temporales (“¿En qué mes se celebra en España la Navidad?”) o con restricciones de tiempo (“¿Cuántas películas fueron filmadas en 2005?”). Particularmente en estos sistemas resulta de especial interés la integración de mecanismos de razonamiento sobre el tiempo que doten a la aplicación de una nueva dimensión temporal (Moldovan, Bowden, y Tatu, 2006).

Un tratamiento de información temporal adecuado ha de comenzar con una detección precisa de las expresiones temporales en las colecciones de documentos. Además, ha de ser capaz de manejar cada detección en un formato estándar que capture el valor temporal de la expresión una vez resuelta, permitiendo el razonamiento sin lugar a ambigüedad. De esta manera, en un sistema de recuperación de información que incorpore estas características se puede concretar más el rango de la búsqueda e incrementar la calidad de los resultados.

El sistema implementado para la extracción de información temporal (Vicente-Díez, de

Pablo-Sánchez y Martínez, 2007) procesa los textos de entrada e identifica fechas y expresiones temporales, instantes de referencia, duraciones e intervalos en ellos (reconocimiento). Las expresiones reconocidas son a su vez resueltas cuando por su naturaleza así lo precisan (resolución), devolviéndose en el formato estándar internacional ISO8601 (2004) de representación de fechas y horas (normalización). En la Figura 1 se muestra la arquitectura general del sistema anteriormente descrito.

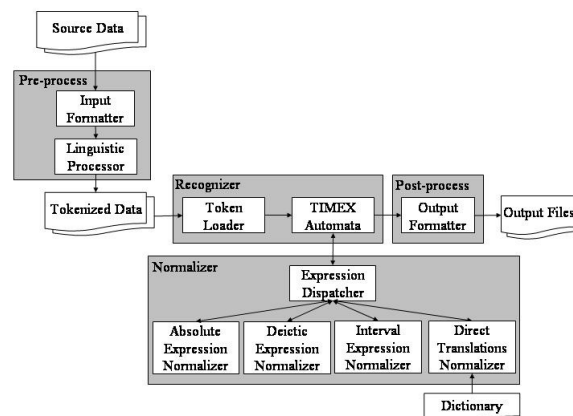


Figura 1: Arquitectura general del sistema de extracción de información temporal

La base del sistema la constituye una gramática de reglas de reconocimiento de expresiones temporales que define el funcionamiento de un autómata de estados finitos. Conjuntamente, se ha desarrollado una propuesta de resolución y normalización de las detecciones llevadas a cabo, que también se realiza de manera automática. Ambas herramientas se han construido a partir de un estudio exhaustivo de los diferentes tipos de expresiones temporales que aparecen en distintos corpus en español. Gracias a este análisis se desarrolló una tipología de las expresiones de tiempo y seguidamente se definió la gramática de reconocimiento y las reglas de resolución y normalización que correspondían con los tipos de aparición más frecuente (Vicente-Díez, Samy y Martínez, 2008).

La Tabla 1 presenta un ejemplo de definición de patrones que constituyen las reglas de la gramática de reconocimiento, así como de las reglas para la resolución de las detecciones y la definición del formato de salida estandarizado.

RECONOCIMIENTO					
PATRON	DESCRIPCIÓN	EJEMPLOS			
COMPLETE_DATE	[ART PREP]? DAY PREP MONTH_NAME PREP YYYY	el_3_de_enero_de_2005			
REL_DEICTIC_UNIT	DEICTIC_UNIT	mañana			
RESOLUCIÓN Y NORMALIZACIÓN					
	FORMATO ENTRADA	REGLA DE RESOLUCIÓN	EJEMPLO		
			ENTRADA	REFERENCIA	SALIDA NORM
ABS_DATE	[ART PREP]? DAY PREP MONTH_NAME PREP YYYY	Day =toDD (DAY) Month=toMM(MONTH_NAME) Year=YYYY	[el] 31 de diciembre de 2005	NA	2005-12-31
REL_DEICTIC_UNIT_FUTURE	mañana	Day=getDD(Creation_Time)+1 Month=getMM(Creation_Time) Year=getYYYY(Creation_Time)	mañana	2008-06-01	2008-06-02

Tabla 1 Ejemplo de reglas de reconocimiento de la gramática temporal y reglas de resolución correspondientes.

3 Sistema de Búsqueda de Respuestas

Los módulos principales que componen la arquitectura general del SBR que ha sido sometido a estudio en este en este trabajo son presentados en la Figura 2.

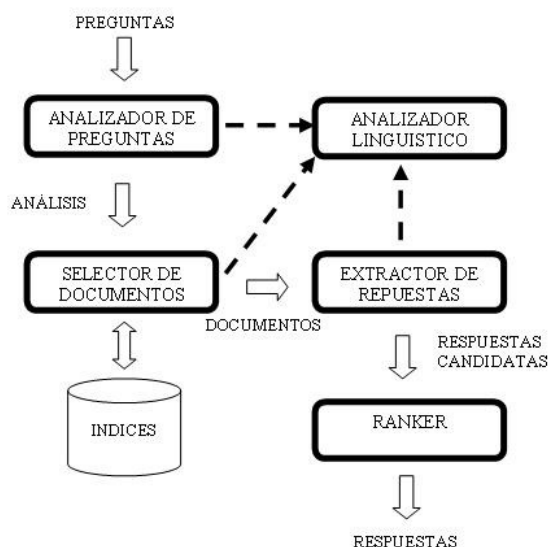


Figura 2: Arquitectura modular del sistema de búsqueda de respuestas

La flecha de bloque marca el recorrido que sigue una pregunta durante su procesamiento. Las flechas discontinuas indican qué uso hacen unos componentes de otros.

A continuación se describe de forma genérica la funcionalidad de cada módulo.

Analizador de Preguntas

Este componente se encarga de clasificar una pregunta de entrada atendiendo a un conjunto de características predefinidas. Para ello emplea un clasificador basado en reglas. Es

el módulo encargado de determinar si una pregunta tiene rasgos de temporalidad, entre otros.

Selector de Documentos

Proporciona acceso a fuentes de información, es decir, a sistemas capaces de proporcionar porciones de texto que pueden contener la respuesta buscada para una pregunta. El sistema da acceso a índices de *Lucene* en los que previamente se han indexado todos los documentos de las colecciones en las que se procederá a buscar las respuestas.

Este módulo precisa de la funcionalidad de análisis lingüístico que ofrece otro de los módulos.

Extractor de Respuestas

Su funcionalidad radica en analizar el contenido de un documento para determinar si aparece o no la respuesta a una pregunta.

Para ello se definen dos tareas: seleccionar las frases que pueden contener una respuesta y determinar qué parte de una frase encierra la respuesta a una pregunta.

Como posibles respuestas se extraen aquellos *tokens* (o grupos de *tokens*) que tienen asignada la etiqueta semántica adecuada. Este componente devuelve una lista de respuestas candidatas a las que asocia un valor de confianza en su corrección.

Este módulo también hace uso de las funciones de análisis lingüístico.

Evaluador de Respuestas (Ranker)

El sistema consta de un componente para puntuar las posibles respuestas. Este sistema de puntuación asigna un peso local y un peso global a cada respuesta. El peso local sólo depende de la frase en la que se encuentra la respuesta. El peso global se calcula teniendo en consideración todas las respuestas encontradas.

Ambos pesos se combinan para proporcionar el peso asignado a la respuesta, siendo la influencia de cada uno de ellos configurable.

Analizador Lingüístico

Este componente incluye toda la funcionalidad relativa al tratamiento lingüístico de los textos con los que se trabaja, sean textos extraídos de documentos o preguntas.

Los textos son divididos y enriquecidos con etiquetado gramatical, morfosintáctico y semántico por las herramientas lingüísticas *StilusTokenizer* y *StilusReader*, desarrolladas por (DAEDALUS, 2008).

4 Integración del tratamiento temporal en la Búsqueda de Respuestas

Algunos autores han denominado búsqueda de respuestas temporal a aquella especialización de la tarea de búsqueda de respuestas en la que las preguntas tienen algún rasgo que denota temporalidad (Saquete, 2005). Este tipo de preguntas pueden ser clasificadas en 3 categorías de acuerdo al papel que juega el dato temporal en su resolución:

Preguntas temporales: aquéllas para las que la respuesta esperada es una fecha o expresión de tiempo (“¿Cuándo se firmo el tratado de Maastricht?”)

Preguntas con restricción temporal: preguntas en cuyo contenido se encuentra una fecha o expresión temporal que circunscribe su respuesta (“¿Quién ganó el Oscar a la mejor actriz en 1995?”).

Preguntas temporales con restricción temporal: combinan las características de los dos tipos anteriormente descritos (“¿En qué temporada anterior a 1994 se enfrentaron Barcelona y Milán?”).

Para que el SBR base sea capaz de resolver este tipo de cuestiones temporales es necesaria la integración de ciertas capacidades de tratamiento de información temporal en algunos de sus módulos. En este trabajo se presenta una evolución del SBR base cuya implementación se ha centrado en la mejora de los módulos de análisis de preguntas y extracción de respuestas.

4.1 Análisis de preguntas

Este módulo es el encargado de la clasificación de las preguntas de entrada. Entre otras, se encarga de clasificar las preguntas con rasgo de temporalidad de acuerdo a la categorización que se describió anteriormente. Su funcionalidad ha sido incrementada dotándole de capacidad para detectar, resolver y

normalizar expresiones temporales dentro de esas preguntas.

Esto permite que en el tratamiento de las preguntas temporales se pueda refinar el tipo de respuesta que se espera del sistema. Por ejemplo, ante una pregunta como “¿En qué año...?” el sistema es capaz de detectar que la respuesta ha de coincidir con un patrón que corresponda a un año exclusivamente.

Por otra parte, en el caso de que la pregunta contenga una expresión temporal, el sistema es capaz de extraerla, clasificando automáticamente dicha pregunta como restringida temporalmente, y utilizando su resolución para acotar la búsqueda.

De manera adicional, el uso de la forma normalizada de las expresiones temporales detectadas en las preguntas permite recuperar aquellos documentos que contienen información temporal que coincide no sólo literal sino también semánticamente con los términos de la búsqueda (“15 de septiembre”, “15/09”, “15 sept.”).

4.2 Extracción de respuestas

Detectar correctamente el tipo de pregunta es fundamental para conseguir una buena respuesta, pero también lo es extraer los fragmentos de texto adecuados para responder a la pregunta, así como asignar una valoración conveniente a las posibles respuestas.

Este módulo extrae como respuesta aquéllos *tokens* que tienen asignada la etiqueta semántica correspondiente al tipo de pregunta formulada. En el caso de las preguntas temporales el sistema con tratamiento de información temporal es capaz de proporcionar respuestas que cumplen la restricción de ser expresiones temporales o fechas, con la granularidad que dicte la pregunta (fecha completa, año, mes,...).

Cuando se trata de responder a preguntas con restricción temporal, el módulo aplica nuevas reglas de extracción basadas en inferencia temporal. En el caso de que una respuesta candidata no cumpla la restricción temporal impuesta en la pregunta, la regla reducirá la valoración final de dicha respuesta. Si por el contrario la cumple, la confianza en que pueda ser una respuesta correcta aumenta con respecto al resto.

Se ha desarrollado un mecanismo de inferencia temporal básico fundamentado en el principio de inclusión: de un instante de tiempo en un intervalo, y de un intervalo en otro. Esta

inferencia es facilitada por la normalización previa de las expresiones de tiempo.

5 Evaluación y resultados

Esta sección muestra una comparativa entre los resultados obtenidos por el SBR base y el ampliado con tratamiento de expresiones temporales. Ambos sistemas trabajan sobre los índices de documentos obtenidos de las colecciones que se muestran en la Tabla 2.

	Documentos	Tamaño	Tipo
EFE 1994	215.738	509 MB	txt
EFE 1995	238.307	577 MB	txt
Wikipedia	≈ 370.000	--	html

Tabla 2 Colecciones de documentos indexadas

Con el fin de realizar esta evaluación ha sido también preciso emplear un corpus de preguntas temporales que permitiera contrastar los resultados de uno y otro sistema. Dicho corpus se ha obtenido de la tarea principal de QA@CLEF2008, y cuenta con un total de 200 preguntas. De éstas, 46 están clasificadas según su cariz temporal dentro de alguno de los 3 tipos definidos en el punto anterior. Este subconjunto de preguntas será sujeto de estudio por su interés particular para este trabajo. La Tabla 3 muestra las proporciones de cada tipo de pregunta con temporalidad en el corpus.

Se ha realizado una evaluación manual de la corrección de las respuestas a las preguntas con rasgos de temporalidad obtenidas por ambos

sistemas, habiéndose llevado a cabo el cómputo en términos cuantitativos y cualitativos. Para esta última medición se ha hecho un análisis en función de si la respuesta correcta ha sido ofrecida como primera, segunda o tercera opción, descartándose el resto de posibilidades.

Preguntas Restricción Temporal	26	13%
Preguntas Temporales	19	10%
Preguntas Temporales con Restricción Temporal	1	0,5%
Total	46	23%

Tabla 3 Preguntas con rasgo de temporalidad en QA@CLEF2008

Los resultados obtenidos por el SBR base, sin mecanismos de inferencia y tratamiento temporal, se muestran en la Tabla 4. Éste responde correctamente a 8 de las 46 preguntas, siendo 5 de esas respuestas ofrecidas como primera opción.

En cuanto al SBR con capacidades temporales, los resultados se muestran en la Tabla 5. En este caso se observa un incremento en el número de respuestas correctas obtenidas, contabilizándose un total de 9. Además, los resultados se ven también mejorados cualitativamente, ya que 7 de esas respuestas se ofrecen como primera opción.

La utilización del SBR temporal supone un incremento en la tasa de acierto del 2,17% con respecto al SBR base, y del 4,35% en el ratio de respuestas correctas en primera instancia.

	Aciertos en					
	1ª respuesta		2ª respuesta		3ª respuesta	
Preguntas con Restricción Temporal	1	3,85%	1	3,85%	1	3,85%
Preguntas Temporales	4	21,05%	0	0%	1	5,26%
Preguntas Temporales con Restricción Temporal	0	0%	0	0%	0	0%
Total	5	10,87%	1	2,17%	2	4,35%

Tabla 4 Resultados del SBR sin tratamiento de información temporal

	Aciertos en					
	1ª respuesta		2ª respuesta		3ª respuesta	
Preguntas con Restricción Temporal	3	11,54%	0	0%	0	0%
Preguntas Temporales	4	21,05%	1	5,26%	1	5,26%
Preguntas Temporales con Restricción Temporal	0	0%	0	0%	0	0%
Total	7	15,22%	1	2,17%	1	2,17%

Tabla 5 Resultados del SBR con tratamiento de información temporal

6 Conclusiones y líneas futuras

A la vista de los resultados obtenidos en este trabajo, se puede concluir que el tratamiento de la información temporal constituye una línea de mejora a tener en cuenta por los SBR actuales.

Bien es cierto que el corpus de preguntas disponibles con rasgos de temporalidad es pequeño y por tanto los resultados no pueden considerarse absolutamente concluyentes. No obstante, mediante la integración de un sistema de reconocimiento y normalización temporal junto con el ajuste de los módulos de análisis de la pregunta y extracción de la respuesta, permitiéndose la realización de una inferencia temporal sencilla, se consigue que el número de respuestas correctamente satisfechas se vea incrementado. Conjuntamente, también la probabilidad de obtener la respuesta correcta en la primera opción aumenta. Ambas características hacen que consideremos prometedora esta línea de trabajo.

Aun habiéndose obtenido resultados positivos en términos cuantitativos y cualitativos, sería interesante ahondar más en las causas que originan dicho incremento. Del mismo modo queda pendiente analizar los motivos por los cuales en determinadas ocasiones el sistema disminuye la valoración de la respuesta correcta.

Se plantea también como línea futura la mejora de otros módulos del SBR, añadiendo nuevas capacidades de tratamiento del tiempo. En este sentido se destaca la posibilidad de combinar el sistema de indexación actual con índices temporales constituidos de expresiones temporales normalizadas. El enriquecimiento del mecanismo actual de inferencia temporal es otra mejora susceptible de llevar a cabo.

Por último, también el sistema de extracción de expresiones temporales podría ser perfeccionado con el fin de incrementar la cantidad de información temporal completa y correctamente manipulada.

Bibliografía

DAEDALUS. 2008. Data, Decisions and Language, S. A. <http://www.daedalus.es>. Visitado: Enero 2009.

Hartrumpf, S. y Leveling, J. 2006. University of Hagen at QA@CLEF 2006: Interpretation and Normalization of Temporal Expressions. En, *Working Notes of the 2006 CLEF Workshop*. Alicante (Spain).

ISO8601:2004(E). 2004. Data elements and interchange formats – Information interchange – Representation of dates and times. Tercera edición 2004-12-01.

Moldovan, D. y Clark, C. 2005. Temporally Relevant Answer Selection. En *Proceedings of the 2005 International Conference on Intelligence Analysis*. Mayo 2005.

Moldovan, D. Bowden, M. y Tatu, M. 2006. A Temporally-Enhanced PowerAnswer in TREC 2006. En *The Fifteenth Text REtrieval Conference (TREC 2006) Proceedings*. Gaithersburg, MD, (USA).

de Pablo-Sánchez, C., Martínez, J.L., González Ledesma, A., Samy, D., Martínez, P., Moreno-Sandoval, A. y Al-Jumaily, H. 2007. MIRACLE Question Answering System for Spanish at CLEF2007. En *Working Notes of the 2007 CLEF Workshop*. Budapest (Hungary). Septiembre 2007.

Saquete, E. Resolución de Información Temporal y su Aplicación a la Búsqueda de Respuestas. 2005. Tesis Doctoral en Informática, Universidad de Alicante.

Vicente-Díez, M.T., de Pablo-Sánchez, C. y Martínez, P. 2007. Evaluación de un Sistema de Reconocimiento y Normalización de Expresiones Temporales en Español. En *Actas del XXIII Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural (SEPLN 2007)*, páginas 113-120. Sevilla, (Spain). Septiembre 2007.

Vicente-Díez, M.T., Samy, D. y Martínez, P. 2008. An Empirical Approach to a Preliminary Successful Identification and Resolution of Temporal Expressions in Spanish News Corpora. En *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*. Marrakech, (Morocco). Mayo 2008.