

Aplicación de técnicas basadas en PLN al tratamiento de preguntas médicas en Búsqueda de Respuestas

Rafael M. Terol, Patricio Martínez-Barco, Manuel Palomar

Universidad de Alicante

Carretera de San Vicente del Raspeig - Alicante

{rafamt, patricio, mpalomar}@dlsi.ua.es

Resumen: Actualmente existe una cierta tendencia investigadora hacia el área de la búsqueda de respuestas en dominios restringidos. En este papel se detalla la parte de nuestro sistema de búsqueda de respuestas en el dominio médico que se encarga del análisis de la pregunta. Este módulo está basado en técnicas sofisticadas de PLN y utiliza el Metatesauro de UMLS como fuente de conocimiento. La principal técnica de PLN utilizada por este módulo de análisis de la pregunta consiste en el tratamiento computacional de la forma lógica de la pregunta y en el emparejado de patrones.

Palabras clave: análisis de la pregunta, formas lógicas, entidades médicas, tratamiento de patrones

Abstract: Nowadays, there is an increasing interest in research on QA over restricted domains. Concretely, in this paper we will show the process of question analysis in our medical QA system. In this system we combine the use of NLP techniques and the UMLS Metathesaurus as knowledge source. The main NLP technique is the use of logic forms and the pattern matching technique in this question analysis performance.

Keywords: question analysis, logic forms, medical named entities, pattern treatment

1. Introducción

Según la definición de la conferencia TREC¹, la tarea de búsqueda de respuestas en dominios abiertos consiste en extraer la respuesta correcta de fragmentos de texto organizados en grandes colecciones documentales a una pregunta dada.

Existen dos enfoques claramente diferenciados para diseñar estos sistemas de búsqueda de respuestas en dominios abiertos: de una parte están los sistemas basados en conocimiento que se diseñan en base a una serie de reglas que definen el comportamiento del sistema; de otra parte, están los sistemas basados en aprendizaje automático que basan su funcionamiento en el entrenamiento de unos coeficientes estocásticos. Ejemplos de ambos tipos de sistemas de búsqueda de respuestas en dominios abiertos los podemos encontrar

en autores como Moldovan (Moldovan et al., 2003), Sasaki (Sasaki, 2005), Vicedo (Vicedo et al., 2004), Zukerman (Zukerman y Raskutti, 2002), etc. Según los resultados oficiales de la conferencia TREC, estos sistemas de búsqueda de respuestas en dominios abiertos obtienen unos resultados de evaluación que rondan entre el 30 y el 40 por cien de precisión. Cuando queremos trabajar en un dominio restringido como puede ser el dominio médico es imprescindible incrementar notablemente esta precisión debido a la información crítica que se maneja en este tipo de dominio ya que una respuesta incorrecta podría poner en grave peligro la salud de las personas (no contestar es mejor que contestar erróneamente).

Nuestro sistema de búsqueda de respuestas en el dominio médico (Terol et al., 2005) tiene una estructura modular basada en cuatro módulos principales: análisis de la pregunta, recuperación documental, filtrado de frases o pasajes, y extracción de la respuesta. El módulo de análisis de la pregunta es el

¹Text REtrieval Conference (TREC) es un conjunto de workshops organizados por the National Institute of Standards and Technology (NIST) que han sido diseñados para avanzar el estado del arte en Recuperación de Información y Búsqueda de Respuestas

núcleo de nuestro sistema de búsqueda de respuestas ya que su correcto o incorrecto funcionamiento condiciona considerablemente al resto de módulos del sistema y, en definitiva, a la respuesta devuelta por el sistema. En este artículo se describen las características de este módulo encargado de clasificar y analizar las preguntas formuladas al sistema. En las siguientes secciones del artículo se muestran las características del mismo así como su evaluación.

2. *Motivación*

En los dominios clínicos interactúan diferentes agentes: médicos, pacientes, enfermeras, laboratorios clínicos, etc. Todos ellos precisan para su actividad de rápidos, efectivos y fáciles métodos de acceso a la información electrónica. El acceso a la última bibliografía médica ayuda a los doctores a seleccionar los mejores diagnósticos, ayuda a los pacientes a conocer todo lo relacionado con su malestar, y permite establecer el tratamiento más efectivo. Esta interacción entre los diferentes agentes produce mucha información que debería ser procesada computacionalmente. Por ejemplo, la gente quiere obtener competentes respuestas a preguntas médicas: cuando tienen determinados síntomas les puede interesar saber a qué causas pueden ser debidos, cuando quieren conocer otra opinión médica acerca de la mejor manera de tratar su enfermedad, cuando quieren conocer la opinión de doctores experimentados en relación a determinados síntomas y su estado, etc. Todas estas características derivan en que el número y tipo de preguntas médicas que un sistema médico de búsqueda de respuestas puede tratar es muy elevado.

Estos hechos nos llevaron a desarrollar nuestro propio sistema de búsqueda de respuestas en el dominio médico (Terol et al., 2005). Este sistema es capaz de responder preguntas médicas basadas en una taxonomía de preguntas concreta. Esta taxonomía de preguntas se obtuvo a partir del estudio desarrollado por Ely *et al* (Ely et al., 2000) cuyo principal cometido fue el desarrollo de una taxonomía de preguntas formuladas por una serie de doctores determinados. En este estudio, los participantes fueron 103 médicos de familia de Iowa y 49 doctores de atención primaria de Oregon. Los autores concluyeron que las preguntas clínicas en atención primaria pueden ser categorizadas en un limitado número

de tipos genéricos. En base a la taxonomía de preguntas derivada de este estudio, las diez preguntas más frecuentes formuladas por estos doctores están ordenadas en la siguiente enumeración:

1. What is the drug of choice for condition x?
2. What is the cause of symptom x?
3. What test is indicated in situation x?
4. What is the dose of drug x?
5. How should I treat condition x (not limited to drug treatment)?
6. How should I manage condition x (not specifying diagnostic or therapeutic)?
7. What is the cause of physical finding x?
8. What is the cause of test finding x?
9. Can drug x cause (adverse) finding y?
10. Could this patient have condition x?

De este modo, nuestro sistema de búsqueda de respuestas en el dominio médico es capaz de responder preguntas en base a esta taxonomía de preguntas genérica, descartando otro tipo de preguntas. Este hecho produce por una parte, una baja cobertura pero, por otra parte, una elevada precisión con el objetivo de que este sistema sea muy funcional en el dominio médico acorde a esta taxonomía de preguntas. La siguiente sección muestra con detalle la parte de nuestro sistema que se encarga de construir los patrones de las preguntas tratadas por el sistema acorde a la taxonomía de preguntas médicas presentada.

3. *Procesamiento de la taxonomía de preguntas*

En esta sección se detalla el modo en el que nuestro sistema construye un patrón para cada pregunta genérica tratada. Los patrones de preguntas se construyen según dos enfoques: generación manual de patrones y generación automática de patrones supervisada. Para poder desarrollar los patrones, nuestro sistema hace uso de los recursos de PLN detallados a continuación.

3.1. Recursos PLN

Los recursos de PLN que utiliza nuestro sistema en el cumplimiento de sus funciones son los siguientes: Formas Lógicas, Reconocimiento de Entidades Médicas y Conocimiento

Semántico. A continuación mostramos como nuestro sistema utiliza estos diferentes recursos:

3.1.1. Formas Lógicas.

La forma lógica de una frase se calcula a partir del análisis de dependencias entre las palabras de la frase. Este recurso está desarrollado a partir de una serie de reglas que infieren distintos atributos en la forma lógica como el predicado, el tipo de predicado, el identificador del predicado y sus relaciones con otros predicados de la forma lógica. Las técnicas empleadas por este recurso son diferentes que las empleadas por el recurso de formas lógicas desarrollado por Moldovan *et al* (Moldovan y Rus, 2001) que infiera la forma lógica a partir del árbol de análisis sintáctico, o el mismo recurso de Diego Mollá *et al* (Mollá *et al.*, 2002) que introduce la forma plana como paso intermedio entre la frase y su correspondiente forma lógica. Al igual que el recurso de Moldovan *et al*, nuestro recurso de formas lógicas está basado en la nomenclatura empleada por el recurso Logic Form Transformation de eXtended WordNet (Harabagiu *et al.*, 1999). Por ejemplo, la forma lógica “*nerve:NN(x3) NNC(x4, x3, x5) cell:NN(x5) consist_of:VB(e1, x4, x10) large:JJ(x1) cell:NN(x2) NNC(x1, x2, x9) body:NN(x9) and:CC(x10, x1, x7) nerve:NN(x6) NNC(x7, x6, x8) fiber:NN(x8)*” es inferida automáticamente a partir del análisis de las dependencias entre las palabras de la frase “*Nerve cells consist of a large cell body and nerve fibers*”. Según el formato de la forma lógica, cada predicado tiene al menos un argumento, pudiendo tener más de uno: el primer argumento se corresponde con el identificador del predicado mientras que el resto de argumentos se se asocian con los identificadores de otros predicados relacionados con el predicado actual.

3.1.2. Reconocimiento de Entidades Médicas.

Este recurso aporta al sistema toda la información referente a las entidades del dominio médico: nombres de medicamentos, síntomas, enfermedades, disfunciones, etc. La base de conocimiento de este recurso es el Metatesauro de UMLS (Lindberg *et al.*, 1993).

3.1.3. Conocimiento Semántico.

Para extraer el conocimiento semántico utilizamos dos fuentes de conocimiento. Por una parte, usamos WordNet (Miller, 1990)

para extraer las relaciones emánticas entre términos de propósito general. Por otra parte, utilizamos el Metatesauro de UMLS para obtener las relaciones semánticas entre los términos médicos.

Una vez presentados estos tres recursos utilizados por nuestro sistema, el resto de la sección se centra en detallar el uso de estos recursos por el sistema para el desarrollo de los patrones de preguntas tratadas. Esta tarea previa al análisis de la pregunta consiste en la obtención de los diferentes patrones asociados a cada pregunta genérica tratada por el sistema. Un patrón está formado por una serie de entidades médicas y uno o más verbos. A continuación describimos los dos métodos de obtención de patrones.

3.2. Generación Manual de Patrones

La generación manual de estos patrones se desarrolla del siguiente modo:

- Identificación de los tipos de entidades médicas que se deben emparejar en la pregunta genérica.
- Identificación de los verbos que se deben emparejar en la pregunta genérica.
- Expansión automática de estos verbos en base a sus relaciones de similitud definidas en WordNet.
- Seleccionar el umbral inferior de entidades médicas (MELT) de cada patrón. MELT se define como el número mínimo de entidades médicas que se deben emparejar entre el patrón de la pregunta genérica y la pregunta del usuario.
- Seleccionar el umbral superior de entidades médicas (MEUT) de cada patrón. MEUT se define como el número máximo de entidades médicas que se pueden emparejar entre el patrón de la pregunta genérica y la pregunta del usuario.
- Identificar los posibles tipos de respuesta esperados.

Como ejemplo de este proceso de generación manual de patrones para la primera pregunta “What is the drug of choice for condition x?”, la Tabla 1 muestra las entidades médicas asociadas a la primera pregunta genérica mientras que la Tabla 2 muestra los patrones resultado de la combinación de estas entidades.

Id. Patrón	Patrón	MELT	MEUT
P_{11}	Sustancia Farmacológica + Enfermedad o Síndrome	2	3
P_{12}	Sustancia Farmacológica + Concepto Cuantitativo	2	3
P_{13}	Sustancia Farmacológica + Signo o Síntoma	2	3
P_{14}	Droga Clínica + Enfermedad o Síndrome	2	3
P_{15}	Droga Clínica + Concepto Cuantitativo	2	3
P_{16}	Droga Clínica + Signo o Síntoma	2	3

Cuadro 2: Patrones asociados a la primera pregunta genérica

Palabra	Entidad Médica
drug	Sustancia Farmacológica Droga Clínica
x	Enfermedad o Síndrome Concepto Cuantitativo Signo o Síntoma

Cuadro 1: Entidades Médicas asociadas a la primera pregunta genérica

Continuando con este ejemplo, el verbo principal de la pregunta de tipo 1 formulada por el usuario se debe corresponder con alguno de los verbos de la lista (*treat, control, take, associate with, help, prevent, manage, indicate, relieve, evaluate, help, fight and solve*). Esta lista de verbos se completa automáticamente con los verbos expandidos de WordNet que tienen alguna relación de similitud con los primeros. Finalmente, los tipos de respuesta esperados para esta primera pregunta genérica se instancian manualmente a los tipos semánticos de UMLS “Sustancia Farmacológica” y “Droga Clínica”.

3.3. Generación Automática de Patrones Supervisada

La generación automática de estos patrones se realiza a través del procesamiento de las preguntas emparejadas a la taxonomía de preguntas del siguiente modo:

- Derivación de la forma lógica asociada a cada pregunta.
- Reconocimiento de las entidades médicas en la forma lógica. Las entidades médicas sólo pueden ser sustantivos (predicado del tipo NN) o nominales complejos (predicado del tipo NNC) incluyendo sus posibles modificadores (predicado del tipo JJ).
- Reconocimiento del verbo principal en la forma lógica.

- Expansión automática de este verbo principal a través de las relaciones de similitud con otros verbos dadas por WordNet.
- Asignación automática del umbral MELT cuyo valor es igual al número de entidades médicas en la forma lógica menos 1.
- Asignación automática del umbral MEUT cuyo valor es igual al número de entidades médicas en la forma lógica.
- Asignación manual de los tipos de respuesta esperados para cada pregunta.

Este proceso está supervisado por un usuario administrador avanzado del sistema pudiendo modificar los resultados obtenidos en cada uno de estos pasos.

La Tabla 3 muestra un ejemplo de este proceso de generación automática de patrones supervisada para la primera pregunta genérica “What is the drug of choice for condition x?”.

Continuando con este ejemplo, el verbo principal de la pregunta (*administer*) se extiende automáticamente con los verbos devueltos por WordNet que tienen relaciones de similitud con el primero. Los tipos de respuesta esperados se asocian también manualmente con al menos un tipo semántico de UMLS.

Una vez que el sistema tiene construidos los patrones asociados a cada una de las preguntas genéricas tratadas, el proceso puro de búsqueda de respuestas ya está listo para poder iniciarse, pudiendo formular ya de este modo preguntas al sistema. La siguiente sección muestra en detalle el módulo del sistema encargado de clasificar y analizar las preguntas en lenguaje natural formuladas al sistema.

4. Clasificación y Análisis de Preguntas

Esta sección muestra las características de diseño del módulo de nuestro sistema médico

Aplicación técnicas basadas en PLN al tratamiento preguntas médicas en BR

Pregunta:	What drugs are administered for the treatment of hypertension?
FL:	drug:NN(x2) administer:VB(e1, x4, x2) for:IN(e1, x3) treatment:NN(x3) of:IN(x3, x1) hypertension:NN(x1)
EM de drug:	Sustancia Farmacológica (SF)
EM de treatment:	Concepto Funcional (CF)
EM de hypertension:	Enfermedad o Síndrome (ES)
Patrón Automático:	SF + CF + ES. MELT=2 and MEUT=3
Patrón Supervisado:	SF + ES. MELT=2 and MEUT=2

Cuadro 3: Generación automática supervisada de un patrón asociado a la primera pregunta genérica

de búsqueda de respuestas encargado de clasificar y analizar las preguntas en lenguaje natural que los usuarios pueden formular al sistema. Este proceso computacional está basado en dos tareas distintas:

- **Clasificación de la Pregunta:** consiste en asignar uno de los patrones genéricos a la pregunta del usuario siempre y cuando dicha pregunta se asocie con una de las diez preguntas genéricas tratadas por el sistema.
- **Análisis de la Pregunta:** consiste en realizar un complejo procesamiento de la pregunta del usuario en base al patrón emparejado y su correspondiente pregunta genérica.
- Reconocer las entidades médicas de aquellos predicados de la forma lógica cuyo tipo se corresponda con nombre (NN) o nominal complejo (NNC) incluyendo sus posibles modificadores (JJ).
- Construcción de la forma de la pregunta del usuario y asignación de su marcador de entidades médicas (MESQ). MESQ se define como el número de entidades médicas en la forma de la pregunta del usuario.
- Obtención de aquellos patrones en cuya lista de verbos esté contenido el verbo principal de la forma lógica de la pregunta del usuario y que además cumplan la restricción $MELT \leq MESQ \leq MEUT$.
- Asignación de la medida de emparejado de entidades (EMM) definida como el número de entidades médicas que se deben emparejar entre la pregunta y el patrón.
- Selección de patrón que minimiza diferencia entre EMM y MELT.

A continuación pasamos a detallar estas dos tareas.

4.1. Clasificación de la Pregunta

Esta tarea de clasificación de la pregunta comienza después de que el usuario lance su pregunta al sistema. Nuestro sistema maneja diez tipos de preguntas en base a las diez preguntas genéricas presentadas anteriormente. De este modo, esta tarea debe decidir si la pregunta del usuario pertenece a alguna clase (se empareja con alguna pregunta genérica) o no pertenece a ninguna clase (no se empareja con ninguna pregunta genérica). Para conseguir este objetivo, esta tarea se enfoca en el tratamiento de patrones derivados de las preguntas del usuario al sistema y se realiza siguiendo los siguientes pasos:

- Formular la pregunta del usuario al sistema.
- Derivar la forma lógica asociada a la pregunta del usuario.
- Extraer el verbo principal de la forma lógica.

La Tabla 4 muestra un ejemplo de la tarea de clasificación de la pregunta utilizando la siguiente pregunta de usuario “What drug treats temperature?”.

4.2. Análisis de la Pregunta

Una vez que la pregunta del usuario ha sido emparejada con un patrón, la tarea de análisis de la pregunta obtiene su semántica. Tal y como hemos mencionado anteriormente, el Metatesauro de UMLS y WordNet son los recursos utilizados para el desempeño de esta función. El siguiente paso consiste en el reconocimiento del tipo de respuesta esperado. Estos tipos de respuestas médicas pueden ser enfermedades, síntomas, dosis de medicinas, ..., de acuerdo con las posibles res-

Pregunta:	What drug manages temperature?
FL:	drug:NN(x2) manage:VB(e1, x2, x1) temperature:NN(x1)
Verbo Principal:	manage
EM of drug:	Sustancia Farmacológica
EM of temperature:	Concepto Cuantitativo
PP:	Sustancia Farmacológica + Concepto Cuantitativo. MESQ=2
Patrones Comparables:	$P_{11}, P_{12}, P_{13}, P_{14}, P_{15}$ and P_{16}
EMM:	$P_{11}^Q = 1, P_{12}^Q = 2, P_{13}^Q = 1, P_{14}^Q = 0, P_{15}^Q = 1$ and $P_{16}^Q = 0$
Patrón Seleccionado:	P_{12}
Clase de Pregunta Genérica:	GE_1 (primera pregunta genérica)

Cuadro 4: Ejemplo de la tarea de clasificación de preguntas

puestas a las diez preguntas médicas tratadas por el sistema. Después de esto, se identifican las palabras clave de la pregunta del usuario. Estas palabras clave se identifican directamente mediante la aplicación de una serie de heurísticas a los predicados y sus relaciones con otros predicados en la forma lógica. Las palabras clave son nominales complejos o nombres reconocidos como expresiones médicas (utilizando el reconocedor de entidades médicas) incluyendo sus posibles adjetivos modificadores, el resto de nominales complejos y nombres no reconocidos como expresiones médicas incluyendo también sus posibles modificadores y el verbo principal en la forma lógica de la pregunta del usuario. Por ejemplo, en la parte de la forma lógica “... high:JJ(x3) blood:NN(x1) NNC(x3, x1, x2) pressure:NN(x2) ...”, el predicado x3 es reconocido como una *Enfermedad o Síndrome* y, en consecuencia, “high blood pressure” tiene el tratamiento de palabra clave. Estas palabras clave de la pregunta del usuario pueden ser expandidas mediante la aplicación de una serie de heurísticas. Por ejemplo, las expresiones médicas pueden ser expandidas utilizando las relaciones de similitud definidas en el Metatesauro de UMLS. De este modo, “high blood pressure” se expande a “hypertension”.

Este conjunto de palabras clave es elegido por prioridad, así si muchas palabras clave han sido extraídas de la pregunta del usuario, solamente un número máximo de palabras clave son consideradas por el siguiente proceso de recuperación documental.

5. Evaluación

A pesar de que los sistemas de búsqueda de respuestas en dominios abiertos pueden ser evaluados en competiciones como el TREC o el CLEF, cuando estamos trabajando en un dominio restringido no existe ningún tipo de tareas de evaluación de esta índole. Esta es

la principal motivación por la que la evaluación de nuestra tarea de clasificación se basa en la tarea de evaluación presentada por Chung *et al* en su trabajo de investigación previo (Chung et al., 2004). De este modo, para evaluar la efectividad de la tarea de clasificación de la pregunta, se ha desarrollado un conjunto de preguntas del siguiente modo:

- GQ_1 : Cinco preguntas asociadas a la primera pregunta genérica; GQ_2 : cinco preguntas asociadas a la segunda pregunta genérica; y así sucesivamente hasta GQ_{10} . Estas preguntas han sido desarrolladas por un grupo de gente ajena al diseño y desarrollo del sistema de búsqueda de respuestas al que se le dio una serie de directrices específicas para el desarrollo de cada tipo de pregunta.
- OQ : El conjunto de las 200 preguntas en inglés de la última tarea de evaluación de los sistemas de búsqueda de respuestas en la competición CLEF. Ejemplos de este tipo de preguntas pueden ser: “What is BMW?”, “What is the FARC?”, “Who is Silvio Berlusconi?”, etc.

La Figura 1 muestra el modo en el que la tarea de clasificación de preguntas es capaz de clasificar cada una de las preguntas del usuario en una de las siguientes clases de preguntas:

- GE : Estas clases de preguntas incluyen cada una de las diez preguntas genéricas. De este modo, la clase GE_1 se corresponde con la primera pregunta genérica “What is the drug of choice for condition x?”; la clase GE_2 se asocia con la segunda pregunta genérica “What is the cause of symptom x?”, ..., y así sucesivamente hasta completar la clase GE_{10} .

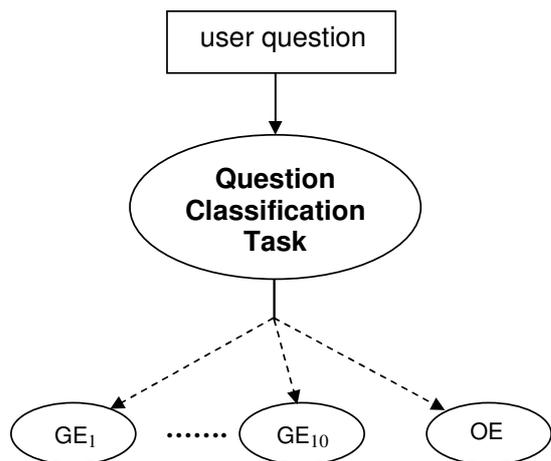


Figura 1: Tarea de Clasificación de Preguntas

- OE: Esta clase incluye el resto de preguntas que no se emparejan con ninguna de las preguntas genéricas tratadas por el sistema.

De este modo, la tarea de evaluación consiste en comparar si cada una de las 250 preguntas de evaluación (GQ_1, \dots, GQ_{10} y OQ) ha sido clasificada correctamente en la apropiada clase de preguntas (GE_1, \dots, GE_{10} u OE). Como medida de evaluación, aplicamos la medida de precisión (P) definida como el cociente entre el número de preguntas clasificadas correctamente y el número de preguntas. La Tabla 5 muestra los resultados obtenidos en esta tarea de clasificación de preguntas. La columna *Clase Clasificada* expresa la clase de preguntas que estamos evaluando. La columna *Clase Relacionada* muestra la clase relacionada correcta asociada a cada clase clasificada. La columna *Preguntas* refleja el número de preguntas clasificadas. La columna *Correctas* indica el número de preguntas que han sido clasificadas correctamente. Finalmente, la columna *Precisión* muestra la precisión acorde a la medida de evaluación definida previamente.

De acuerdo a la fila *Global* de la Tabla 5, la precisión de esta tarea de clasificación de preguntas es del 94,4%. Este buen resultado condiciona positivamente que el resto de módulos del sistema trabajen con preguntas que han sido correctamente clasificadas con una alta probabilidad.

6. Conclusiones y Trabajos Futuros

En este artículo se justifica la necesidad de mejorar la precisión de los sistemas de búsqueda de respuestas en dominios restringidos, especialmente en el dominio médico. En este dominio médico, nuestro esfuerzo investigador se centra en el diseño y desarrollo de un sistema de búsqueda de respuestas capaz de resolver una serie de preguntas basadas en una taxonomía de preguntas médicas formuladas a través de una serie de estudios de doctores sobre la atención a los pacientes. También se detalla el método de procesamiento de las preguntas en el sistema de búsqueda de respuestas en el dominio médico. Este método hace uso del Metatesauro de UMLS y de WordNet. Este proceso computacional consta de tres tareas claramente diferenciadas: (1) generación de patrones, (2) clasificación de preguntas, y (3) análisis de la pregunta.

Actualmente estamos trabajando en el desarrollo de un marco genérico de sistemas de búsqueda de respuestas en cualquier dominio restringido.

Bibliografía

- Hoojung Chung, Young-In Song, Kyoung-Soo Han, Do-Sang Yoon, Joo-Young Lee, Hae-Chang Rim and Soo-Hong Kim. A Practical QA System in Restricted Domains. En *Proceedings of 42nd Annual Meeting of the Association for Computational Linguistics, Workshop on Question Answering in Restricted Domains*, Barcelona, Spain, Julio 2004.
- John W Ely, Jerome A Osherooff, PaulÑ Gorman, Mark H Ebell, M Lee Chambliss, Eric A Pifer and P Zoe Stavri. A taxonomy of generic clinical questions: classification study. *BMJ* 2000, 321:429–432, 2000.
- S. Harabagiu, G.A. Miller, and D.I. Moldovan. WordNet 2 - A Morphologically and Semantically Enhanced Resource. In *Proceedings of ACL-SIGLEX99: Standardizing Lexical Resources*, Maryland, June 1999, pp.1-8.
- Donald A. B. Lindberg, Betsy L. Humphreys, and Alexa. T. McCray. The Unified Medical Language System. In *Methods of Information in Medicine*, 32(4), páginas 281–291, August 1993.

Clase Clasificada	Clase Relacionada	Preguntas	Correctas	Precisión
<i>GQ</i>	<i>GE</i>	50	42	0.84
0Q	0E	200	194	0.97
Global	-	250	231	0.944

Cuadro 5: Evaluación de la tarea de clasificación de preguntas

G.A. Miller WordNet: An on-line lexical database. *International Journal of Lexicography* 3, 4 (Winter 1990), pp.235-312.

Dan Moldovan and Vasile Rus. Logic Form Transformation of WordNet and its Applicability to Question-Answering. En *Proceedings of 39th Annual Meeting of the Association for Computational Linguistics*, Toulouse, France, Julio 2001.

Dan Moldovan, Christine Clark, Sanda Harabagiu y Steve Maiorano. COGEX: A Logic Prover for Question Answering. En *Proceedings of HLT-NAACL 2003. Human Language Technology Conference*, páginas 87-93, Edmonton, Canada, 2003.

Diego Mollá, Rolf Schwitter, Michael Hess and Rachel Fournier. ExtrAns, an answer extraction system. *T.A.L. special issue on Information Retrieval oriented Natural Language Processing*, páginas 495-522, 2002.

Yutaka Sasaki. Question Answering as Question-Biased Term Extraction: A New Approach toward Multilingual QA. En *Proceedings of 43th Annual Meeting of the Association for Computational Linguistics*, Michigan, USA, Junio 2005.

Rafael M. Terol, Patricio Martinez-Barco and Manuel Palomar. Applying Logic Forms to Biomedical Q-A. En *Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications*, Istanbul, Turkey, Junio 2005.

Jose Luis Vicedo, Maximiliano Saiz, Ruben Izquierdo y Fernando Llopis. Does English Help Question Answering in Spanish. En *Proceedings of the 5th Workshop of the Cross-Language Evaluation Forum, CLEF 2004*, Bath, UK, Septiembre 2004.

Ingrid Zukerman y Bhavani Raskutti. Lexical Query Paraphrasing for Document Retrieval. En Hsin-Hsi Chen and Chin-Yew Lin, editors, *Proceedings of the 19th International Conference on Computational Linguistics, COLING 2002*, Taipei, Taiwan, Agosto 2002.