

Un mecanismo de inferencia aplicado a la búsqueda de respuesta

S. Roger^{1,2}, A. Ferrández², J. Peral², S. Ferrández², P. López² *

¹ Natural Language Processing Group
Department of Computing Sciences University of Comahue, Argentina
sroger@uncoma.edu.ar

² Natural Language Processing and Information Systems Group
Department of Software and Computing Systems University of Alicante, Spain
{sroger, antonio, jperal, sferrandez}@dlsi.ua.es, p.lopez@ua.es

Resumen

Este trabajo describe un mecanismo de inferencia aplicado al Sistema de Búsqueda de Respuestas AliQAn para el idioma español y en dominio abierto. AliQAn está basado fundamentalmente en el uso de patrones sintácticos para identificar las posibles respuestas. Un mecanismo de inferencia es aplicado al conjunto de preguntas de tipo económico. De esta manera, nuestro sistema mejora la precisión de este tipo de preguntas de un 33 % a un 57 %.

This work describe an inference mechanism apply to AliQAn Question Answering System for Spanish in open domain. AliQAn is based in use of syntactic pattern to identify the answers. Our approach apply an inference mechanism to questions set of economic type. Thus, our system enhance the accuracy of this question type of 33 % to 57 %.

Keywords: Keywords: Procesamiento de Lenguaje Natural, Búsqueda de respuestas, Prueba lógica.

1. INTRODUCCIÓN

El Procesamiento del Lenguaje Natural (PLN) es una parte esencial de la Inteligencia Artificial mediante el cual se investiga y formula mecanismos computacionales. Estos mecanismos permiten el desarrollo de sistemas capaces de comprender el conocimiento expresado en textos de un lenguaje dado.

En los últimos años, la investigación en el PLN se ha centrado en el desarrollo de recursos que proveen múltiples niveles de análisis sintáctico y semántico. Dichos recursos son aplicados a diferentes aplicaciones como sistemas de traducción automática, recuperación de información (RI), búsquedas de respuestas (BR), reconocimiento de entidades, clasificación y filtrado de documentos, generación de resúmenes, etc.

Este trabajo se centra en la Búsquedas de Respuestas (BR) o en su denominación inglesa *Question Answering*. Estos sistemas de BR suceden a los sistemas de RI en cuanto necesitamos una respuesta

*This research has been partially funded by the Spanish Government under project CICyT number TIC2003-07158-C04-01 and by the Valencia Government under project numbers GV06/161, and by the University of Comahue under the project 04/E062.

concreta a una pregunta dada. El objetivo de los sistemas clásicos de RI es devolver una lista ordenada de documentos en función de su relevancia sobre la consulta realizada por el usuario. Por otra parte, el objetivo de los sistemas de BR es mucho más ambicioso que el de mejorar la precisión de los resultados devueltos, estos sistemas intentan recuperar el fragmento de texto que contiene la información requerida por el usuario, y no el documento completo. Sin embargo, el estado actual de los sistemas de BR aún está en sus albores. El nivel de conocimiento requerido para el entendimiento de los documentos es cada vez mayor y esto hace la diferencia en la precisión de los sistemas de BR actuales.

La actividad investigadora parte de la implementación del sistema de búsqueda de respuestas AliQAn [25, 5] de dominio abierto para el idioma español. AliQAn ha participado en las competencias del CLEF 2005 monolingüe (QA@CLEF-2005), consiguiendo una precisión del 33 %. Este año, AliQAn también ha participado en la misma tarea, no sólo para el idioma español sino también en el idioma inglés, es decir preguntas en inglés y textos en español.

Si se caracterizan a los sistemas de BR presentados en competencias como el CLEF según el nivel de recursos de PLN utilizado, aquellos que llegan hasta un nivel sintáctico o a lo sumo un nivel semántico superficial (mediante la utilización de sinonimia, hiperonimia entre otras relaciones similares) no superan cierto rango de precisión (50 % de efectividad aproximadamente). Los sistemas que superan en gran medida dicho valor es debido a la utilización de técnicas más complejas mediante la utilización de fuentes de conocimientos. Esta investigación pretende dar comienzo a un trabajo de gran envergadura: desarrollar una herramienta robusta capaz de razonar automáticamente en dominios abiertos para el idioma español que podrá ser integrada a distintas aplicaciones de PLN como BR, implicación (en su denominación inglesa *entailment*) o EI entre otras.

La habilidad de razonar en los sistemas de BR apareció a fines de los 70 en dominios cerrados. Actualmente, los sistemas que integran gradualmente componentes de inferencias aún no han evolucionado como aquellos de dominio cerrado [2].

Pese a que los recursos disponibles para el idioma español son muy escasos, se ha podido demostrar que la mejora de esta técnica puede ayudar significativamente en la efectividad de sistemas tal como la BR.

En la sección 2 se presenta una clasificación de los sistemas que utilizan razonamiento. La sección 3 elabora un procedimiento para la obtención de la representación elegida y desarrolla un conjunto de técnicas de transformación. Finalmente en las dos últimas secciones se discute los resultados de las pruebas elaboradas en la sección anterior y ofrece trabajos futuros para extender el método presentado en este artículo.

2. ANTECEDENTES

En los últimos años, los resultados de los certámenes TREC, CLEF o NTCIR han demostrado la necesidad de métodos con mayor grado de conocimiento semántico. Sin embargo, a pesar de que han surgido técnicas valiosas resta muchos problemas por resolver.

Por ello, es imprescindible tener una representación del conocimiento unificada que provea una codificación jerárquica de las propiedades semántica, relacional y estructural de un texto dado y contar con un mecanismo de inferencia que pueda ser usado sobre tal representación.

El marco generalmente aceptado para el estudio del razonamiento para ser aplicado a diferentes “sistemas inteligentes” son los enfoques basados en conocimiento. La idea es almacenar el conocimiento en algún *lenguaje de representación*. Las sentencias expresadas con este lenguaje son almacenadas en una “Base de Conocimiento”(BC). Finalmente, un mecanismo de razonamiento es usado para determinar qué puede ser inferido desde las sentencias en la BC. Diferentes sistemas de repre-

sentación (conjunto de reglas lógicas, redes probabilísticas, etc.) son asociadas con mecanismos de razonamiento, cada uno con sus propios méritos y rango de aplicación [16, 21]. Dada una base de conocimiento lógica por ejemplo, el razonamiento puede ser abstraído como una tarea de deducción: determinar si una sentencia es lógicamente implicada por la BC.

2.1. Clasificación

Las representaciones de conocimientos de propósito general junto con sus algoritmos de inferencias de propósito general correspondientes no resuelven el tema clave de qué representar y cómo derivar una representación abstracta eficientemente considerando, además, la complejidad del problema. Dentro de este marco podemos distinguir las técnicas o métodos basadas en conocimientos y probabilísticos. Los métodos basados en conocimientos utilizan generalmente como lenguaje de representación subyacente las Fórmulas Lógicas (FL), de las cuales podemos encontrar en la literatura un abanico de variantes.

Para poder realizar una clasificación los sistemas se hará según las tareas involucradas en el proceso de conocimiento:

- **Representación de conocimiento (RC):** El lenguaje para poder representar el conocimiento. Las RC de propósito general pueden ser FOL¹, probabilístico o híbrido.
- **Procedimiento para obtener esta representación:** Es el mecanismo mediante el cual se logra la representación del conocimiento en el lenguaje anteriormente mencionado.
- **Mecanismo de inferencias:** Es un mecanismo efectivo para ser utilizado como medio de deducción para decidir si una consulta es deducible a partir de una base de conocimiento descrita en el lenguaje de presentación elegido.

2.1.1. Representación de conocimiento

Un sistema de razonamiento debería seleccionar una estructura simbólica apropiada para poder representar el conocimiento y un mecanismo de razonamiento apropiado para ser capaz de asimilar nueva información y responder a consultas.

Como una primera introducción en la clasificación podemos ver que la representación puede ser basada en modelos o basada en FL.

- **Representación basada en modelos:** [12, 11, 13] la BC es representada como un conjunto de modelos del mundo en lugar de describirlo como un conjunto de FL. Otra propuesta es la presentada por [4] en la que se utiliza la semántica de modelos estables definida por [7] permitiendo otorgar semántica a los programas lógicos con negación.
- **Representación basada en FL:** Un nivel de representación del significado independiente del contexto son las Fórmulas Lógicas. Hay diversas formas de representación [8, 18, 26, 20, 22]. Podemos considerar que el contenido empírico de la teoría de la FL tiene dos aspectos: uno concierne a la sintaxis y la relación de ésta con la forma de las oraciones; el otro concierne al papel de la FL en la explicación de la noción general de significado. Las teoría-FL y la semántica modelo-teórica son muy parecidas en algunos aspectos. Ambas tienen que ver con el significado estructural, alejándose del significado de la palabra y de la pragmática, y las dos proponen un lenguaje lógico que represente el significado estructural de las oraciones. De este modo, podría decirse que en las dos teorías la FL es una subteoría discreta de la teoría de significado. La semántica modelo-teórica va aún más lejos al afirmar que la FL, más una interpretación teórica establecida de ésta, es una subteoría discreta de la teoría del significado.

¹First Order Logic

2.1.2. *Proceso de obtención de la representación elegida*

El proceso de obtención de la representación elegida es un tema de estudio en sí mismo. Si bien la tarea de elegir una representación adecuada es importante, la forma en que llegamos a ella lo es aún más. Es difícil clasificar a los sistemas en este punto dado que son muy particulares cada uno de ellos y necesitaríamos detallar el mecanismo de cada uno de ellos en forma particular. Los grupos que podemos identificar son los:

- Sistemas basados en relaciones de dependencias [3, 9, 10, 14]
- Sistemas basados en análisis sintáctico [19, 17, 18, 8]
- Sistemas basadas en análisis mixto [23, 24, 1]

2.1.3. *Mecanismos de inferencias*

Inferir es concluir o decidir a partir de algo conocido o asumido. A su vez, razonar es pensar coherente y lógicamente, establecer inferencias o conclusiones a partir de hechos conocidos o asumidos.

El proceso de razonamiento, por lo tanto, involucra la realización de inferencias, a partir de hechos conocidos. Realizar inferencias significa derivar nuevos hechos a partir de un conjunto de hechos conocidos como verdaderos.

Entre los sistemas de inferencias podemos mencionar:

- *Sistemas de inferencia directos*: Aplican un mecanismo de razonamiento “hacia adelante”: partiendo de resultados conocidos (los axiomas, y, en el caso de demostración de razonamientos, las premisas) aplican sucesivamente reglas de inferencia hasta llegar a la fórmula que se pretende probar. Ejemplos: sistemas axiomáticos o sistemas de Hilbert, deducción natural, cálculo de secuentes.
- *Sistemas de inferencias indirectos o por refutación*: Aplican un mecanismo de razonamiento indirecto basado en la técnica de reducción al absurdo: para demostrar que una fórmula es válida se parte de su negación y se realizan inferencias hasta llegar a una contradicción. Ejemplos: método de Resolución, tableaux semánticos.

Desafortunadamente, el PLN utiliza grandes cantidades de texto, que pueden ser transformados en conocimiento. La Lógica Clásica es adecuada para trabajar en ciertos dominios concretos, pero resulta insuficiente para modelar correctamente el razonamiento de este tipo de conocimiento.

Por ello, se han desarrollado otras lógicas, denominadas genéricamente Lógicas no Clásicas, construidas en base a la Lógica Clásica pero con mayor poder expresivo y con múltiples aplicaciones prácticas:

- Lógicas que incorporan operadores aplicables sobre fórmulas (operadores de creencia, posibilidad, necesidad, temporales, ...): Lógicas modales, Lógicas temporales.
- Lógicas que permiten trabajar con información incompleta, incierta o imprecisa: Lógicas multivaluadas, Lógicas probabilistas, Lógicas difusas, Lógicas No Monótonas.

Existen diversos modelos de Razonamiento No Monótono, que pueden clasificarse en dos grandes categorías:

- *Modelos Simbólicos*: Hipótesis del Mundo Cerrado o en su abreviatura inglesa (CWA), Lógicas modales no monótonas, Lógicas defectivas o por defecto, Circunscripción, entre otros.
- *Modelos No Simbólicos*: Modelos probabilísticos, Modelos difusos, Sistemas basados en herencia, Sistemas de mantenimiento de la verdad, entre otros.

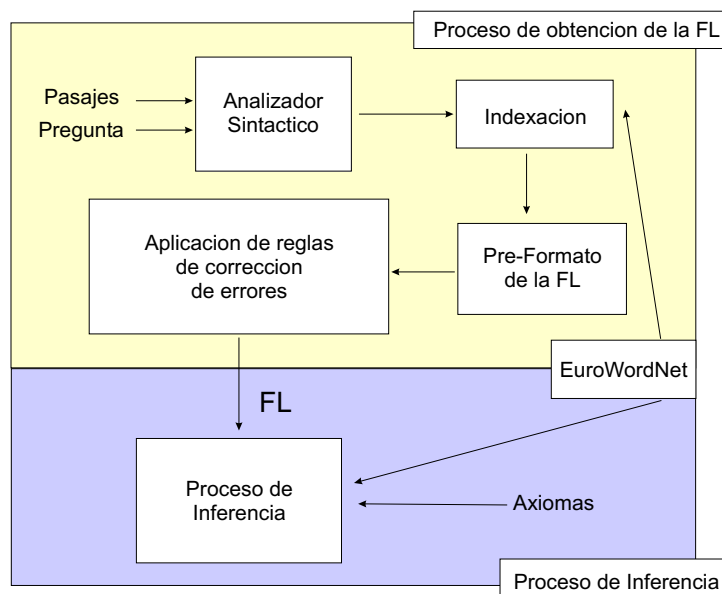


Figura 1: Arquitectura del sistema de razonamiento

3. PROPUESTA

Como se mencionó anteriormente los sistemas que utilizan recursos de PLN hasta un nivel sintáctico no superan un cierto rango de precisión. Una nueva generación de sistemas ha comenzado y dan un paso más allá de estos tipos de sistemas. La nueva tendencia en los sistemas de BR tienden a incorporar más semántica en el proceso de comprensión de los textos mediante la utilización de técnicas más complejas utilizando fuentes de conocimientos externas.

A continuación se detallará la representación elegida, el proceso de traducción de las FL, las correcciones realizadas a partir de la entrada del proceso y, posteriormente, el sistema de inferencia.

3.1. Representación lógica

Para elegir una representación adecuada es necesario tener presente la utilidad que vayamos a hacer de ella, los recursos utilizados, el idioma elegido y futuras extensiones entre otras cosas.

No existen sistemas que utilicen FL para el idioma castellano y por lo tanto no existen recursos factibles de ser usados para ayudar al proceso de inferencia. Es conocido que el idioma inglés, es un idioma en el que la mayoría de recursos se encuentran disponibles. Teniendo en cuenta esto, se ha decidido optar por una representación que sea fácil adaptar a la representación resultante de la traducción al castellano de algunos de estos recursos. Esta representación es la utilizada por el ExtendedWordNet (XWN).

3.2. Proceso de transformación en la representación elegida

La figura 1 muestra la arquitectura del proceso de traducción de la salida del analizador sintáctico a FL y su posterior ingreso al proceso de inferencia.

Para transformar un párrafo en su representación lógica, partimos de la salida del analizador sintáctico parcial SUPAR[6]. Luego de esto, tanto la pregunta como los pasajes son indexados. En el proceso de indexación, se realiza un proceso de desambiguación [25, 5]. El ejemplo 2 muestra un fragmento de un pasaje indexado. Obtenemos una jerarquía del texto dividido en oraciones y a su

- La población mundial aumentará un 50_por_ciento en los próximos 35 años, sobre todo en los países en desarrollo, alcanzando la cifra de 8.500 millones de personas, lo que creará más dificultades, según el Banco Mundial.

```

3|1|C|X;F;F;Fg;#
|O;la;La;NP00000;01547674 01353399 00550797 00366547 01563290 11288204 07485702 12408908
01232923 04513542* poblacion;población;NCF000;00509473 04699283 10100482 12668165
12384195*mundial;mundial;AQ0CS0;#!#!#
|V;aumentar;aumentará;VMIF3S0;00136946 01937583 00754271 02308524,00436817 01227164
01470863 00169989 00153292,00606271 01937583 00754271 02308524,01204939 00136946 01937583
00754271 02308524,01540637 02175245 02007386 02308524,01753713 00866031 01278826,01937583
00754271 02308524,02143984 01705360 00136946 01937583 00754271 02308524,02205361 02520193
00827403
|S;uno;un;DI0MS0;#*50/100;50_por_ciento;Zp;#
!(en;en;el;los;DA0MP0;#*proximo;próximo;AQ0MP0; #*35; 35; Z;#*año; años;NCMP000;05642123
01536186 06022468 04513542,11162972 01536186 06022468 04513542,13991652 01536186 06022468
04513542;#;#)#!#
...

```

Figura 2: Ejemplo de salida de la un pasaje indexado

vez divididas en cláusulas. AliQAn almacena los bloques sintácticos obtenidos a partir del resultado de la indexación. Estos bloques pueden ser de diferentes tipos: Objetos (O), Sujetos (S), Verbo (V), Complementos (C), Relativos (R) y Palabras Sueltas (X).

A partir de la salida de proceso de indexación, se obtendrá un pre-formato de las FL. El proceso de transformación no considera todos los tipos de los elementos. Algunos de ellos son irrelevantes o pueden interferir en un correcto funcionamiento. Otro motivo por el cual se toman en cuenta sólo algunos tipos es la futura extensión del proceso a otras lenguas tal como el inglés. En dicha extensión el proceso de traducción suele ser dificultoso dado que no se cuenta con recursos sustanciales de traducción para grandes volúmenes de palabras.

La representación lógica está dada por el synset de la palabra, de esta manera reducimos el trabajo de considerar la relación de sinonimia de EuroWordNet (EWN) dado que queda implícita. Como primer parámetro tenemos al lema del mismo. Luego el resto de parámetros depende de la etiqueta.

En el caso de los *nombres* que están en EWN, el segundo parámetro representa a los dos primeros caracteres de la etiqueta del PoS tagger, los cuales identifican la categoría y el tipo del mismo, por último parámetro tenemos la variable que será usada para relacionar los predicados entre sí. En el caso de los nombres que no estén en EWN, el nombre del predicado está dado por el lema unido a los caracteres correspondientes al tipo, los parámetros en este caso consistirán del último parámetro del caso anterior correspondiente a la variable.

El caso de los *verbos* es similar a los nombres con la diferencia que agregan dos parámetros más. El segundo parámetro corresponde al evento que representa, el siguiente parámetro liga al sujeto y el tercer argumento representa al objeto que modifica.

Los adjetivos tendrán la misma variable como parámetro correspondiente a la variable del nombre que modifica. Análogamente, los adverbios tendrán la variable de evento que el posee el verbo que está modificando.

Como el analizador sintáctico es parcial, se producen ciertos errores que es necesario identificar y corregir. Como el proceso de corrección es realizada en Prolog, se realiza el reconocimiento anteriormente mencionado para formar un predicado llamado **floo**. Esta estructura es de la forma

- El jugador zimbabuo de 37 años, fue detenido el martes en el transcurso de un amplio operativo policial

```
flooo(3, [[1,c,[
  [o,s03209158(jugador1,nc,x33),zimbabuo1_aq(x33)],
  [c, de_(x35,x36), numero_Z('37',x36), s05642123(ano1,nc,x36)],
  [v,s(ser1,vs,x37), detener1_vm(e1,x38,x39)],
  [s,fecha_W('martes,S0,S1,S2,S3,S4',x39), en_(x39,x40),
    transcurso1_nc(x40), de_(x40,x41),
    amplio1_aq(x41), s01178555(operacion1,nc,x41),
    policial1_aq(x41), [rr, 2]],
  [r,4]
]]).
]]).
```

Figura 3: Ejemplo de salida de la estructura flooo

flooo(OOO, [CCC, Tipo, [[tipo, Predicado1, Predicado2, ...], ...]], [CCC, Tipo, [[tipo, Predicado1, Predicado2, ...], ...]]). Un ejemplo del análisis y transformación a esta estructura está dada por la figura 3. Posteriormente se aplican reglas de reducción y corrección mediante el empleo de heurísticas.

3.3. Proceso de inferencia

El proceso de inferencia es la resolución. La estrategia de resolución elegida es la del conjunto soporte por ser completa y sana, usando multiresolución. Existen diversos razonadores disponibles como Otter, Vampire entre otros. En este trabajo utilizamos el primer razonador mencionado.

En una teoría formal encontramos un conjunto de fórmulas bien formadas, un subconjunto de éstas que son los axiomas y un conjunto de reglas de inferencias. El conjunto de fórmulas bien formadas de nuestra propuesta son las FL descritas anteriormente. El proceso de inferencia utiliza axiomas. Algunos de ellos generados automáticamente y otros creados manualmente. A continuación ejemplificaremos algunas clases de axiomas utilizados. En particular se mostrarán los axiomas que serán utilizados en la siguiente sección:

- Axiomas generados por las relaciones de EWN: por ejemplo, sinonimia, hiperonimia, hiponimia, etc.
- Axiomas Lingüísticos: Estos tipos de axiomas son generados manualmente.
 - Los posesivos que sustituyen $\forall X, X_1 \text{ por}(X, X_1) \rightarrow \text{posesivo}(X, X_1)$
 - Los axiomas que relacionan número con cantidades por ejemplo, esto será útil para la identificación del tipo de respuesta esperado:

$$\forall X, X_1 (\text{numero_Zm}(X, X_1) \rightarrow \text{cantidad}(X_1))$$

- Como último ejemplo, el siguiente axioma lingüístico dice que un verbo tiene como sujeto un nombre o sintagma nominal entonces estos están relacionados por medio de la preposición **de**:

$$\forall X, X_1 (\text{interpol1_np}(X) \ \& \ \text{disponerde1_vm}(E, X, X_1) \ \& \ \text{s02607133}(\text{presupuesto1,nc}, X_1) \rightarrow \text{de_}(X_1, X)).$$

- Axiomas de Unión: Supongamos las siguiente FL:

1. interpol1_np(x98) &
2. disponer1_vm(e1,x98,x99) &
3. de_(x99,x100) &
4. s02607133(presupuesto1,nc,x100) &
5. de_(x100,x101) &
6. numero_Zm(28000000_dolar,x101)

En este caso el verbo **disponer** está seguido por la preposición **de**. Realizando un reconocimiento de multipalabra determinamos que el término **disponerde** está contenido en EWN. De esta manera, podemos reemplazar las FL 2 y 3 en una sola FL y cuyas variables de ligazón serán combinadas de manera tal que la FL resultante será: *disponerde_vm(e1, x98, x100)*

- Axiomas de cambio de variables: estos axiomas corresponde al orden de las variables del verbo **disponerde** y su significado, es decir podemos tener la siguiente implicación:

$$\forall X, X_1, X_2 (\text{interpol1_np}(X_1) \& \text{disponerde_vm}(E_1, X_1, X_2) \& \text{s02607133}(\text{presupuesto1}, \text{nc}, X_2) \rightarrow \text{s02607133}(\text{presupuesto1}, \text{nc}, X_2) \& \text{de_}(X_2, X_1) \& \text{interpol1_np}(X_1))$$

A continuación se muestra un fragmento del párrafo que da solución a la pregunta 112 del CLEF-2006.

1. -cantidad(X2) | -s02607133(Presupuesto1, nc, X2) | -de_(X2,X3) | -interpol1_np(X3) .
2. s06880763(narcotrafico1,nc,x88).
3. representar1_vm(e1,x90,x91).
4. actualmente1_rg(e91).
5. numero_Zm(4000000000000_dolar,x92).
6. de_(x92,x93).
7. s02348211(beneficio1,nc,x93).
8. al_(x93,x94).
9. s05642123(anyo1,nc,x94).
10. mientras_que1_cs(x95).
11. interpol1_np(x98).
12. disponer1_vm(e1,x98,x99).
13. de_(x99,x100).
14. s02607133(presupuesto1,nc,x100).
15. de_(x100,x101).
16. numero_Zm(28000000_dolar,x101).
17. -disponer1_vm(E, X, X1) | -de(X1, X2) | disponerde1_vm(E, X, X2).
18. -interpol1_np(X) | -disponerde1_vm(E, X, X1) | -s02607133(presupuesto1, nc, X1) | de_(X1,X).
19. -interpol1_np(X) | -disponerde1_vm(E, X, X1) .
20. -s02607133(presupuesto1, nc, X1) | -disponerde1_vm(E, X, X1) .
21. -s02607133(presupuesto1, nc, X) | -de(X, X1) | -numero_Zm(C, X1) | numero_Zm(C, X).
22. -de_(X, X1) | -numero_Zm(C, X1) .
23. -numero_Zm(X,X1) | cantidad(X1).
24. disponerde1_vm(E, 98, 100) (12, 13, 17)
25. de_(100,98)(18, 14, 11)
26. numero_Zm(28000000_dolar,x100)(14, 15, 16)
27. cantidad(x100) (23, 26)
28. □ (1, 14, 11, 25, 27)

4. EVALUACIÓN

En esta sección, se detallarán los resultados obtenidos mediante la utilización de la propuesta descrita anteriormente, los cuales demuestran que la estrategia seguida mejora la precisión del sistema de BR AliQAn. Antes de mostrar los resultados realizados, se detallan una serie de aspectos, como son el conjunto de datos utilizado y la medida de evaluación necesarios para obtener conclusiones de los resultados.

La evaluación realizada sobre el sistema AliQAn [25] y la aplicación del módulo de inferencia han sido llevados a cabo usando el conjunto de preguntas propuesto por el CLEF 2006 clasificadas por AliQAn como económicas, y la colección de documentos de la agencia EFE del año 1994 y 1995 proporcionadas por el CLEF para el mismo año.

Cada solución devuelta por el sistema puede catalogarse entre las siguientes categorías:

- **Correcta:** la respuesta devuelta por el sistema es exacta y no contiene elementos o componentes extra.
- **Incorrecta:** pueden ocurrir dos situaciones:
 - las respuestas devueltas por el sistema no responden a la pregunta formulada;
 - las preguntas no han sido contestadas aunque exista solución en los documentos.
- **Inexacta:** la respuesta devuelta por el sistema es correcta pero:
 - incluye algún elemento o componente extra;
 - le falta algún elemento con respecto a lo que debería ser la solución correcta.
- **No soportada:** Son soluciones correctas descubiertas al azar, es decir, el sistema responde una solución correcta que ha sido extraída de documentos de los que no se puede inferir la respuesta.

Para poder realizar una evaluación de la precisión del sistema, se necesita una medida que valora los resultados generales del sistema en cada uno de los experimentos realizados.

Esta tarea se realiza utilizando la medida propuesta en el CLEF 2003 [15]. La ecuación 1 muestra el *Mean Reciprocal Rank* (MRR) en donde sólo las respuestas correctas son puntuadas.

$$MRR = \frac{(\sum_{i=1}^Q \frac{1}{far(i)})}{Q} \quad (1)$$

donde Q es el número totales de preguntas y $far(i)$ indica la posición de la primera respuesta correcta. Para preguntas en las que no se puede encontrar respuestas, el valor de $\frac{1}{far(i)}$ será 0.

Las evaluación del sistema AliQAn es realizada solamente para la tarea monolingüe en castellano. Sólo se ha tenido en cuenta la evaluación de preguntas del tipo económico. En el futuro se ampliará la evaluación a otro tipo de preguntas utilizando nuevos recursos lingüísticos incorporados como conocimiento en el proceso de inferencia. El cuadro 1 muestra el aumento en la precisión de las preguntas de tipo económico del sistema AliQAn mediante la incorporación de técnicas de inferencias.

Cuadro 1: Evaluación de la propuesta de Inferencia sobre AliQAn.

Conjunto de prueba	MRR AliQAn(%)	MRR con Razonamiento (%)
preguntas de tipo económico del CLEF 2006	33	58 (+75 %)

5. CONCLUSIONES

Se puede afirmar que la incorporación de mecanismos de inferencias en sistemas de BR es crucial para el incremento en la precisión de estos sistemas. A pesar de que aún queda mucho trabajo por hacer para que sean realmente eficaces o que logren evolucionar como aquellos en dominios cerrados. Para ello es necesario un estudio pormenorizado y cuidadoso de recursos de PLN, entre otros, para desarrollar una herramienta robusta capaz razonar automáticamente en dominios abiertos.

En este artículo se ha creado una clasificación que caracteriza a los sistemas existentes, se han presentado los mecanismos de razonamientos existentes, aunque quedaron fuera diversas técnicas más elaboradas que serán estudiadas a posteriori como pasos subsiguientes en esta tarea. Dichas técnicas son la aplicación de teorías de razonamiento no monótono, satisfacción de restricciones, entre otras. Además, se ha desarrollado un mecanismo robusto de aplicación de reglas mediante el cual se pretende solucionar algunos tipos de problemas que introducen las fases anteriores. Estas mejoras ayudan a aumentar la calidad de información incorporada al sistema de inferencia. Junto a esto, se ha provisto de técnicas de transformación a partir de los recursos disponibles usados en el sistema AliQAn.

Se ha desarrollado una herramienta de inferencia para el idioma español, idioma en el cual aún no se han desarrollado este tipo de herramientas y del cual existen pocos recursos capaces de ser usados por un método de inferencia.

Finalmente, la representación en FL por sí misma puede ser vista como un cambio en la representación de la salida de un analizador sintáctico y ello no conlleva ninguna mejora en la precisión de los sistemas de BR, sólo el conocimiento proporcionado ya sea explícita o implícitamente lo hace. La aplicación de las reglas mejora la calidad de la salida de las fases anteriores tal como PoS tagger y analizador sintáctico y la incorporación de una mínima cantidad de conocimiento externo ha contribuido a un aumento en la precisión el cual resulta proporcional a la cantidad de conocimiento externo incorporado en un 75 %.

Es importante destacar que esta área de estudio es nueva y es necesaria la creación de recursos para el idioma español, afianzando nuestra lengua en este campo de investigación.

6. TRABAJOS FUTUROS

Como se ha comentado anteriormente, es grande el abanico de técnicas disponibles que puedan posibilitar la creación de un sistema de razonamiento robustos y eficaces. Por ello, es necesario un estudio minucioso de las técnicas utilizadas por los sistemas existentes y también de las técnicas factibles de ser usadas y que aún no lo han sido.

Una tarea que no resulta fácil será el estudio y elaboración de herramientas para ser utilizadas como recurso de conocimientos. Estas herramientas podrán ser resultado de la traducción de herramientas disponibles en otro idioma, de la utilización ontologías a partir de corpus, entre otros medios. Como se ha demostrado en los sistemas que ya están utilizando inferencia en sus sistemas, la adquisición de tales recursos es fundamental en la mejora de la precisión de los sistemas de BR. El idioma español, a diferencia del inglés, por ejemplo, es muy libre en la generación de sentencias váli-

das. Esto produce que una idea sea expresada en un número mayor de formas correctas. Esto juega negativamente en la automatización de un sistema de BR para el español. Sin embargo, es posible compensar esta característica, entre otras cosas mediante la incorporación de más axiomas que representan conocimiento externo válido para ser utilizado en el mecanismo de inferencia y así poder lograr independencia de la representación del texto.

REFERENCIAS

- [1] S. Abreu. Isco: A practical language for heterogeneous information system construction. *In Proceedings of INAP'01. Japan*, October 2001.
- [2] F. Benamara, M. Moens, and P. Saint-Dizier. Knowledge and reasoning for answering questions. *In Workshop associated with IJCAI05, Edinburgh*, July 2005.
- [3] Gosse Bouma, Jori Mur, Gertjan van Noord, Lonneke van der Plas, and Jörg Tiedemann. Question answering for dutch using dependency relations. *In In Workshop of Cross-Language Evaluation Forum (CLEF), 2005*.
- [4] O. Elkhatab, E. Pontelli, and T. Cao Son. Asp-prolog: A system for reasoning about answer set programs in prolog. *In proceedings 10th International Workshop on Non-Monotonic Reasoning (NMR 2004), Whistler, Canada, June 6-8, 2004, 2004*.
- [5] S. Ferrández, S. Roger, A. Ferrández, A. Aguilar, and P. López-Moreno. A new proposal of word sense disambiguation for nouns on a question answering system. *Advances in Natural Language Processing. Research in Computing Science. ISSN: 1665-9899. 7th International Conference, CICling, 18:83–92, February 2006*.
- [6] Antonio Ferrández Rodríguez. *Aproximación computacional al tratamiento de la anáfora pronominal y de tipo adjetivo mediante gramáticas de unificación de huecos*. PhD thesis, Universidad de Alicante, Junio 1998.
- [7] M. Gelfond and V. Lifschitz. The Stable Model Semantics for Logic Programs. *In International Symposium on Logic Programming*, pages 1070–1080, 1988.
- [8] Sanda Harabagiu, Marius A. Pasca, and Steven J. Maiorano. Experiments with open-domain textual question answering. *In Proceedings of the COLING-2000. Association for Computational Linguistics*. Morgan Kaufmann, Aug 2000.
- [9] Valentin Jijkoun, Jori Mur, and Maarten de Rijke. Information extraction in question answering: Improving recall through syntactic patterns. *In Colling 2004*, pages 1284–1290, Geneva, 2003.
- [10] Boris Katz and Jimmy Lin. Selectively using relations to improve precision in question answering. *In Proceedings of the workshop on Natural Language Processing for Question Answering (EACL 2003)*, pages 43–50, Budapest, 2003. EACL.
- [11] H. Kautz, M. Kearns, and B. Selman. Horn approximations of empirical data. *Artificial Intelligence*, 74:129–145, 1995.
- [12] R. Khardon and D. Roth. Reasoning with models. *Artificial Intelligence*, 87:187–213, 1996.

- [13] R. Khardon and D. Roth. Defaults and Relevance in Model Based Reasoning. *Artificial Intelligence*, 1997.
- [14] Kenneth C. Litkowski. Use of metadata for question answering and novelty tasks. In E. M. Voorhees and L. P. Buckland, editors, *Proceeding of the eleventh Text Retrieval Conference (TREC 2003)*, pages 161–170, Gaithersburg, MD, 2004.
- [15] B. Magnini, S. Romagnoli, A. Vallin, J. Herrera, A. Peñas, V. Peinado, F. Verdejo, and M. Rijke. The Multiple Language Question Answering Track at CLEF 2003. *Comparative Evaluation of Multilingual Information Access Systems: 4th Workshop of the Cross-Language Evaluation Forum, CLEF, Trondheim, Norway, August 21-22, 2003, Revised Selected Papers. Lecture Notes in Computer Science*, 2003.
- [16] J. McCarthy and P. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. meltzer and D. Michie, editors, 4, 1969.
- [17] D. Moldovan, S. Harabagiu, R. Girju, P. Morarescu, F. Lacatusu, A.Ñovischi, A. Badulescu, and O. Bolohan. Lcc tools for question answering. In Voorhees and Buckland, editors, *Proceedings of the 11th Text REtrieval Conference (TREC-2002)*, NIST, Gaithersburg, 2002.
- [18] Dan I. Moldovan, Sanda M. Harabagiu, Christine Clark, and Steven J. Maiorano. Cogex: A logic prover for question answering. In *HLT-NAACL*, pages 87–93, 2003.
- [19] Dan I. Moldovan and Vacile Rus. Logic form transformation of wordnet and its applicability to question answering. In *ACL*, pages 394–401, 2001.
- [20] D. Molla and M. Gardiner. Answerfinder: question answering by combining lexical, syntactic and semantic information. In *Australasian Language Technology Workshop(ALTW)*, Sydney, 2004.
- [21] J. Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. In *Workshop of Cross-Language Evaluation Forum (CLEF)*, 1988.
- [22] John M. Prager, Jennifer Chu-Carroll, and Krzysztof Czuba. Question answering using constraint satisfaction: Qa-by-dossier-with-constraints. In *ACL*, pages 574–581, 2004.
- [23] Carlos A. Prolo, Paulo Quaresma, Irene Rodrigues, Pedro Salgueiro, and Renata Vieira. A question-answering system for portuguese. In *Knowledge and Reasoning for Answering Questions. Workshop associated with IJCAI05*, pages 45–48, 2005.
- [24] Paulo Quaresma and Irene Rodrigues. A logic programming based approach to the qa@clef05 track. In *In Cross Language Evaluation Forum: Working Notes for the CLEF 2005 Workshop*, Austria, 2005.
- [25] S. Roger, S. Ferrández, A. Ferrández, J. Peral, F. Llopis, A. Aguilar, and D. Tomás. AliQAn, Spanish QA System at CLEF-2005. In *Workshop of Cross-Language Evaluation Forum (CLEF)*, 2005.
- [26] Vacile Rus. *Logic Forms for WordNet Glosses*. PhD thesis, School of Engineering Southern Methodist University, May 2002.