

CAPÍTULO 5

GRACILEWeb: Traductor Japonés-Inglés en Internet

5.1 Representación del idioma Japonés en el navegador

5.1.1 Charsets

Cuando uno hace una requisición por medio del Protocolo HTTP a algún URL específico la respuesta es enviada (como todo en la red) como una secuencia de bytes. El saber cómo desplegar esa cadena de bytes se da de la siguiente manera.

Por lo general se hace un mapeo de cada byte a un caracter en una tabla en particular . La tabla que se maneja de manera general es la especificada por el estandar ISO-8859-1, también llamada Latin-1 la cual contiene la relación byte-a-caracter para la mayoría de los caracteres usados en los lenguajes de la Europa del Este. De esta manera el browser al recibir una cadena de bytes por lo general tratará de hacer el mapeo de dichos bytes contra esa tabla.

Sin embargo para el manejo de otros idiomas, en nuestro caso del idioma Japonés, el mapear caracter contra byte resulta insuficiente, por lo que para el Japonés, que contiene aproximadamente 2000 caracteres se mapea por medio de charsets, donde un caracter viene dado por un conjunto de bytes. Uno de los charsets comunmente usados en el manejo de el idioma Japonés es el Shift_JIS. Otro charset de uso general pero que también soporta la decodificación a caracteres japoneses a partir de secuencias Unicode es el UTF-8 (UCS Transformation Format, 8-bits Form) el cual maneja una codificación de longitud variable y que trabaja eficientemente con Unicode trasformando dichas secuencias a representaciones de 1, 2 o 3 caracteres.

5.1.2 Caracteres Unicode

En Java, todos los caracteres, cadenas String e indentificadores son manejados internamente como caracteres Unicode que se manejan en 2 bytes. Unicode fue establecido por el Consorcio Unicode y describe el estandar como se muestra en <http://www.unicode.org/unicode/estandar/estandar.html>

Por ejemplo, para el kanji de la palabra "bad":



el código Unicode correspondiente sería: `\u60aa`.

De esta manera Java interpreta cada secuencia Unicode como un caracter único. Un secuencia Unicode viene dada por 4 dígitos hexadecimales de la siguiente manera: \uxxxx. El consorcio Unicode ha desarrollado tablas de equivalencia para secuencias Unicode hacia la mayoría de los lenguajes entre ellos el Japonés. De ahí su importancia para este trabajo. En la tabla 5.1 se presenta como ejemplo la tabla para caracteres Hiragana.

3040		Hiragana					309F
		304	305	306	307	308	309
0		ぐ	だ	ぼ	む	み	
1	あ	け	ち	ぼ	め	ゑ	
2	あ	げ	ち	ひ	も	を	
3	い	こ	っ	び	ゃ	ん	
4	い	ご	っ	び	ゃ	ぐ	
5	う	さ	づ	ふ	ゆ		
6	う	ざ	て	ぶ	ゆ		
7	え	し	で	ふ	よ		
8	え	じ	と	へ	よ		
9	お	す	ど	べ	ら	ゝ	
A	お	ず	な	ぺ	り	ゝ	
B	か	せ	に	ほ	る	ゝ	
C	が	ぜ	ぬ	ぼ	れ	ゝ	
D	き	そ	ね	ぼ	ろ	ゝ	
E	ぎ	ぞ	の	ま	わ	ゝ	
F	く	た	は	み	わ		

570 *The Unicode Standard 3.0, Copyright © 1991-2000, Unicode, Inc. All rights reserved*

Tabla 5.1. tabla Unicode de caracteres Hiragana

5.1.3 Japonés y Prolog

Dentro de SICStus el manejo de cada caracter Japonés se hace mediante secuencias Unicode. Para esto los archivos cargados deben de tener especificado el enunciado.

```
:- prolog_flag(language, _, iso).
```

El cual habilita el uso de la sintaxis `\xHEXCODES\` siendo esta la manera interna en dicho Prolog de secuencias Unicode.

Por lo que una predicado en Prolog haciendo uso de una secuencia Unicode podría quedar representado como:

```
:-predicado('\x#####').
```

Un ejemplo de un archivo con hechos, y reglas que involucren caracteres japoneses podría quedar de la siguiente manera:

```
:- prolog_flag(language, _, iso).  
:- module(japón, [japón/2]).
```

```
significado_Japonés(bad, '\x60aa').
```

Donde se genera el hecho: **significado_Japonés** con dos átomos: un String **bad** y la representación de dicha palabra en Japonés que correspondería al kanji:

悪

5.2 El traductor

El traductor comienza con una página principal en donde se muestra la opción inicial de traducción: Inglés a Japonés o Japonés a Inglés como se muestra en la figura 5.1.



Figura 5.1 Menú principal

Una vez elegido el tipo de traducción es decir Japonés a Inglés (figura 5.2) o Inglés a Japonés (figura 5.3), se presenta un menú en donde se da a elegir por el tipo de texto de entrada a escoger entre un sustantivo(noun), un adjetivo (adjective) o un verbo (verb).



Figura 5.2. Menú Inglés a Japonés



Figura 5.3. Menú Japonés a Inglés

En caso de elegir inicialmente la opción Japonés a Inglés, una vez elegido el tipo de entrada para el texto se presenta una forma de captura de dicho texto para su consecuente procesamiento, como se muestra en la figura 5.4 para sustantivos, en la figura 5.5 para adjetivos y en la figura 5.6 para verbos.

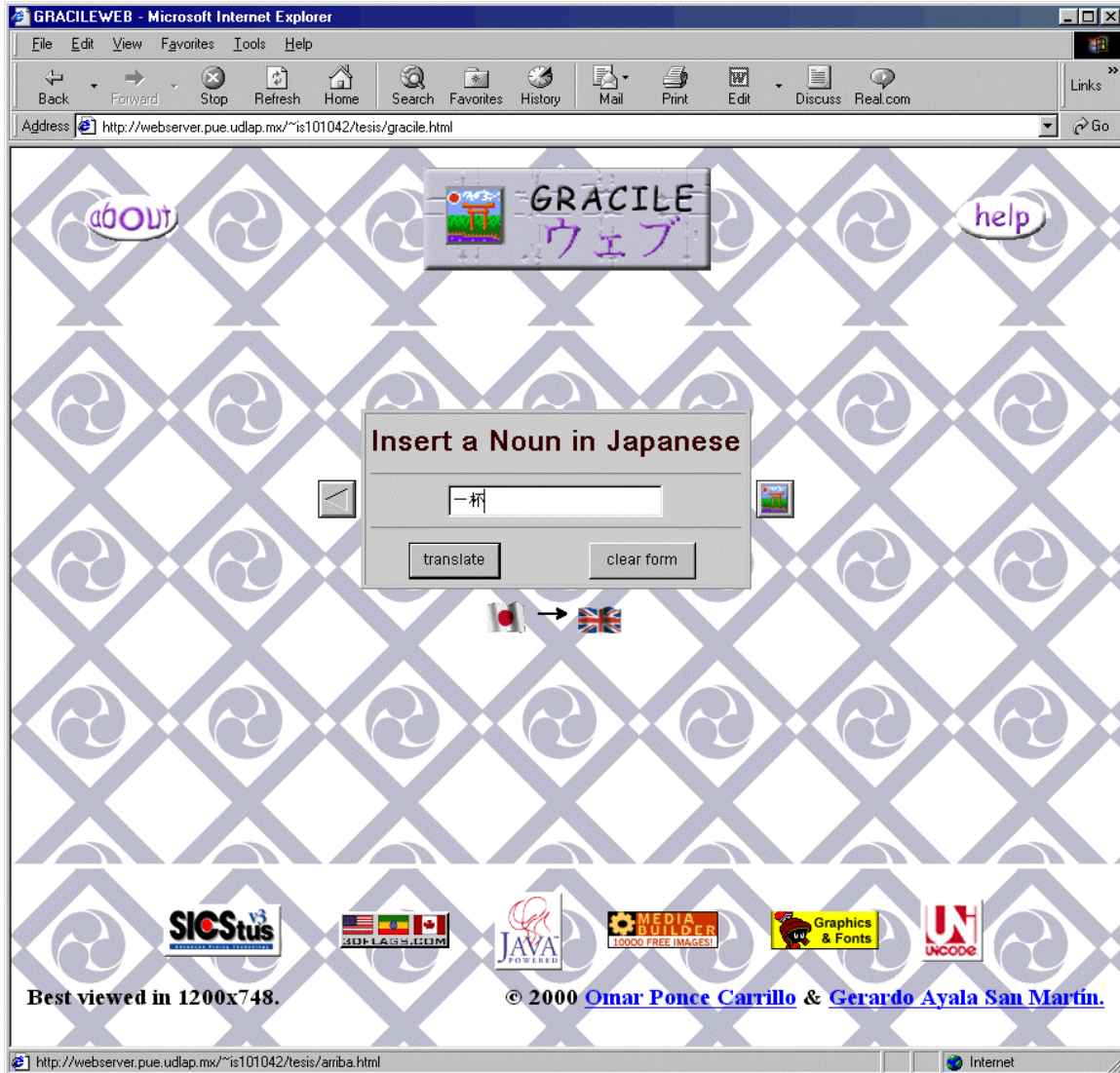


Figura 5.4. Entrada de sustantivo en Japonés



Figura 5.5. Entrada de adjetivo en Japonés



Figura 5.6. Entrada de verbo en Japonés

En caso haber elegido la opción Inglés a Japonés , una vez elegido el tipo de entrada para el texto se presenta una forma con la lista de sustantivos, adjetivos o verbos según sea el caso, con los que se cuentan para traducir. Para sustantivos se muestra en la figura 5.7

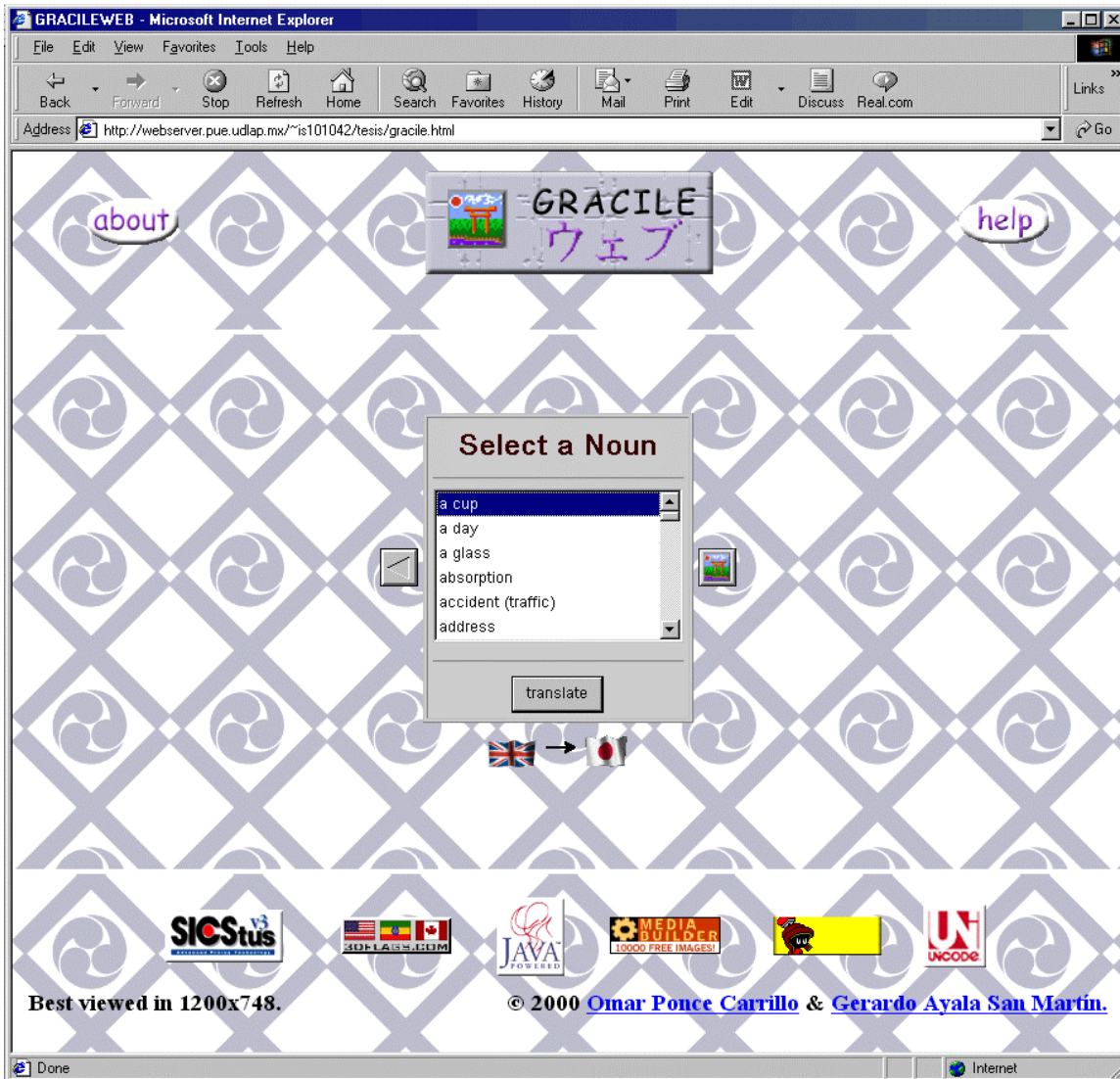


Figura 5.7. Selección de sustantivo en Inglés

Para la opción de adjetivos en inglés, aparte de seleccionar el adjetivo se tiene que seleccionar alguna las opciones de conjugación, donde no todas las combinaciones son posibles, como se muestra en la figura 5.8.



Figura 5.8. Selección de adjetivo en Inglés

Para la opción de Verbos en inglés, aparte de seleccionar el verbo se tiene que seleccionar alguna las opciones de conjugación, donde no todas las combinaciones son posibles, como se muestra en la figura 5.9.



Figura 5.9. Selección de verbo en Inglés

Y finalmente en caso de éxito se presentan los resultados como lo muestra la figura 5.10 para un sustantivo en Japonés, en la figura 5.11 para un adjetivo en Japonés, en la figura 5.12 para un verbo en Japonés; o la figura 5.13 para un sustantivo en Inglés, la figura 5.14 para un adjetivo en Inglés y la figura 5.15 para un verbo en Inglés,

En la figura 5.10 se muestra el resultado de analizar el sustantivo que en Inglés significa tanto *glass* como *cup* con su forma de escritura en Japonés tanto en Hiragana como en Kanji además de mostrar sus significados con cierta información extra que en este caso indica que se comportan como adjetivos *na*.



Figura 5.10. Sustantivo Japonés analizado

En la figura 5.11 se muestra el resultado de analizar el adjetivo que en Inglés significa tanto *bad* como *wrong* mostrando la manera en que el adjetivo fue introducido, su escritura en Hiragana y las formas de conjugación presentadas por dicho adjetivo. Además de mostrar sus significados.



Figura 5.11. Adjetivo Japonés analizado

En la figura 5.12 se muestra el resultado de analizar el verbo que en Inglés significa tanto *abolish* como *do await with* como *repeal* mostrando la manera en que el adjetivo fue introducido, su escritura en Kanji y Hiragana, su escritura en forma diccionario y las formas de conjugación presentadas por dicho adjetivo. Además de mostrar sus significados.



Figura 5.12. Verbo Japonés analizado

En la figura 5.13 se muestra el resultado de analizar el sustantivo en Inglés *cup* con su forma de escritura en Japonés tanto en Hiragana como en Kanji.



Figura 5.13. Sustantivo Inglés analizado

En la figura 5.14 se muestra el resultado de analizar el adjetivo en Inglés *bad* con su forma de escritura en Japonés tanto en Hiragana como en Kanji ambas conjugadas en base a las opciones de conjugación previamente escogidas por el usuario y las opciones de conjugación previamente escogidas.



Figura 5.14. Adjetivo Inglés analizado

En la figura 5.15 se muestra el resultado de analizar el verbo en Inglés *abolish* con su forma de escritura en Japonés tanto en Hiragana como en Kanji ambas conjugadas en base a las opciones de conjugación previamente escogidas por el usuario, su escritura en forma diccionario y las opciones de conjugación previamente escogidas.

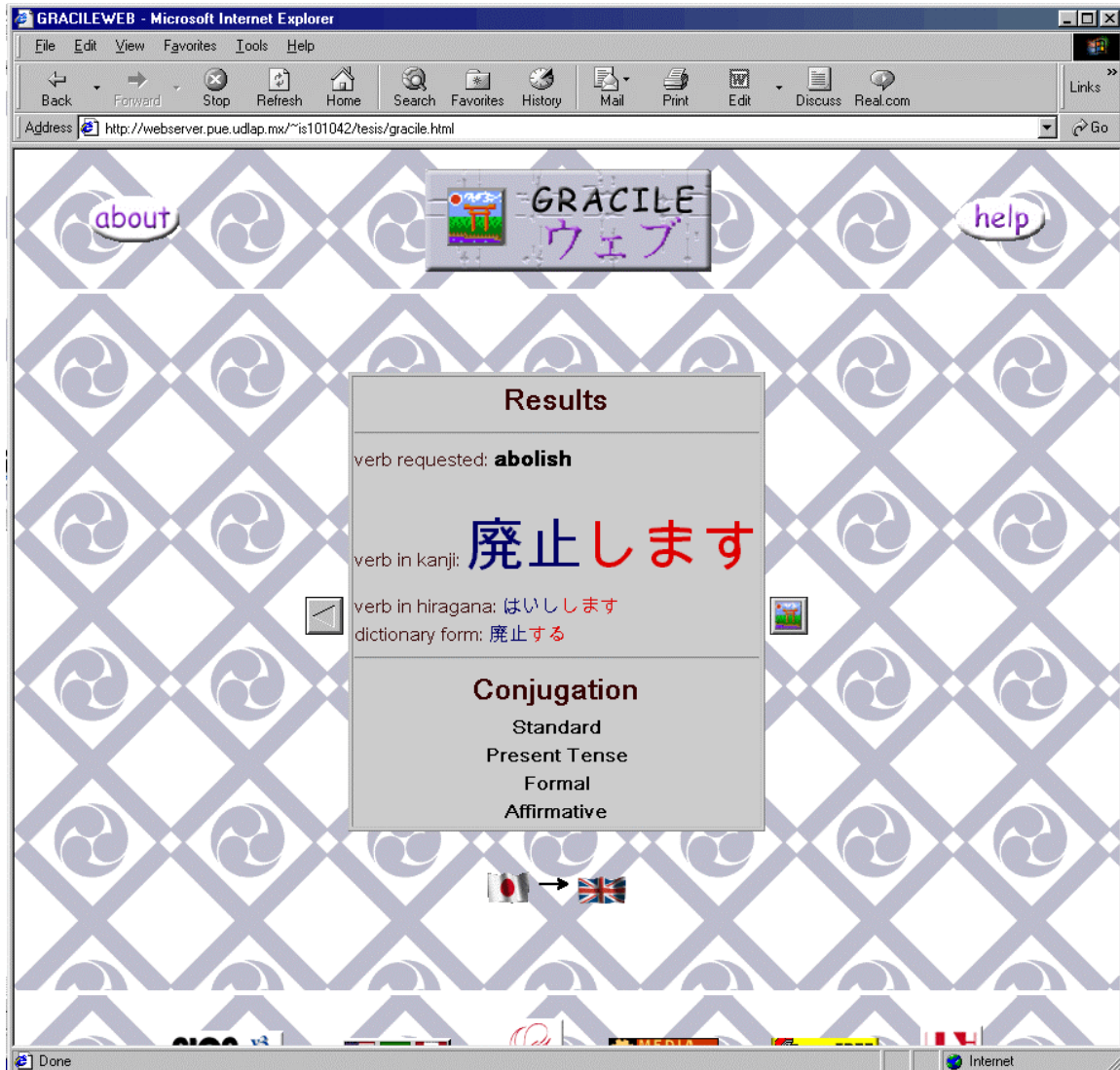


Figura 5.15. Verbo Inglés analizado

El traductor puede ser accedido en la siguiente dirección:
<http://webserver.pue.udlap.mx/~is101042/tesis/gracile.html>

5.3 Modelado en UML

A continuación presentamos los diagramas de UML para el sistema GRACILEWeb. En primer lugar mostramos el modelado de relaciones de las clases de GRACILEWeb en la figura 5.16.

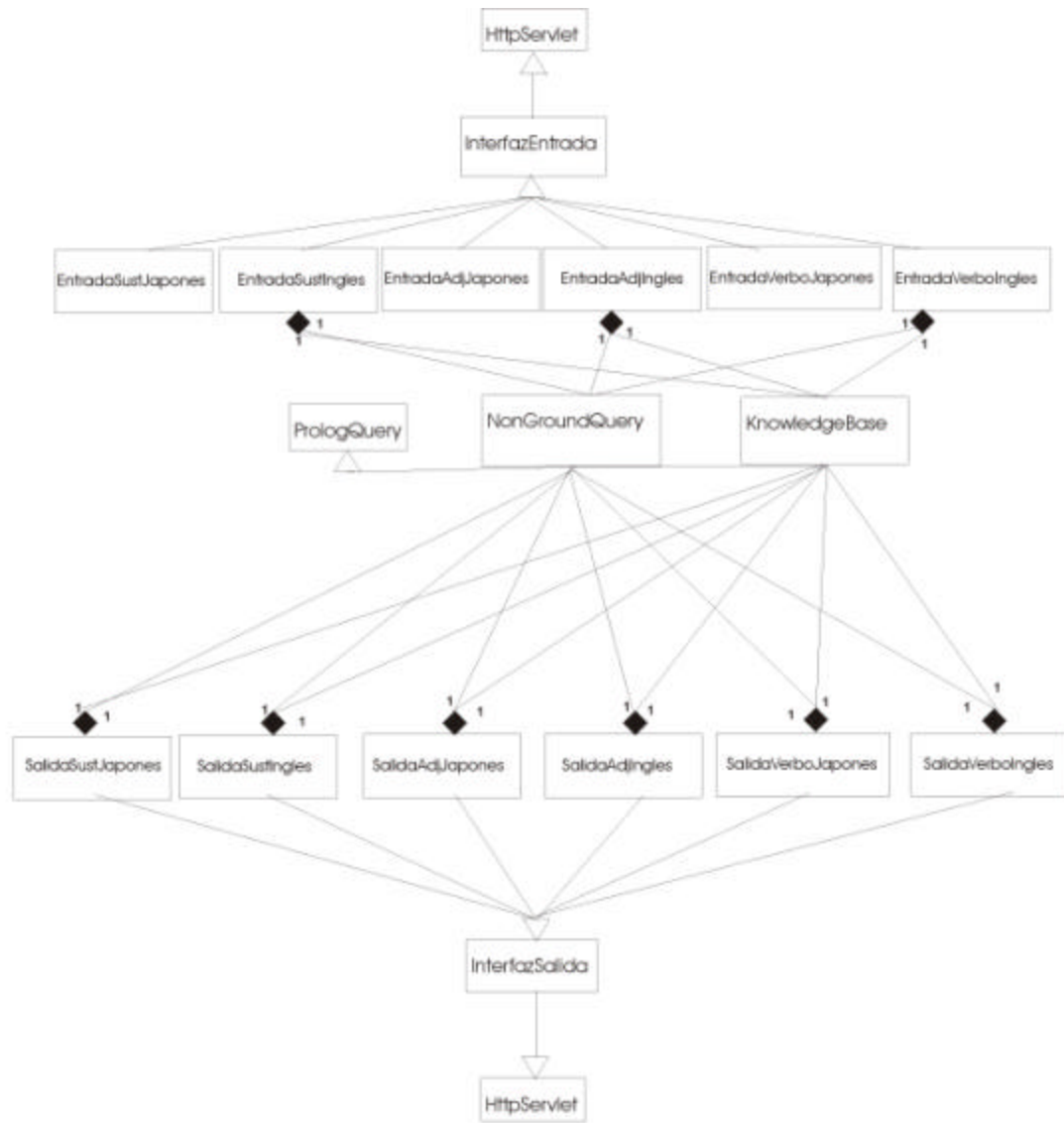


Figura 5.16. Modelado de relaciones de clases de GRACILEWeb.

A continuación presentamos los diagramas de clase de el sistema GRACILEWeb. En la figura 5.17 tenemos la clase InterfazEntrada que es la clase de la cual van a heredar todas las Clases de entrada tanto para el Japonés como para el Inglés.



Figura 5.17 Clase de la interfaz de entrada.

En la figura 5.18 tenemos la clase InterfazSalida que es la clase de la cual van a heredar todas las Clases de salida tanto para el Japonés como para el Inglés



Figura 5.18 Clase de la interfaz de salida.

De la figura 5.19 a la figura 5.24 presentamos las clases de entrada tanto de el Japonés



Figura 5.19 Clase para la entrada de un sustantivo en Japonés.



Figura 5.20 Clase para la entrada de un sustantivo en Inglés.



Figura 5.21 Clase para la entrada de un adjetivo en Japonés.



Figura 5.22 Clase para la entrada de un adjetivo en Inglés.



Figura 5.23 Clase para la entrada de un verbo en Japonés.



Figura 5.24 Clase para la entrada de un Verbo en Inglés.

De la figura 5.25 a la figura 5.30 presentamos las clases de salida tanto de el Japonés como



Figura 5.25 Clase para la Salida de un sustantivo en Japonés.



Figura 5.26 Clase para la Salida de un sustantivo en Inglés.



Figura 5.27 Clase para la Salida de un adjetivo en Japonés.



Figura 5.28 Clase para la Salida de un adjetivo en Inglés.



Figura 5.29 Clase para la Salida de un verbo en Japonés.



Figura 5.30 Clase para la Salida de un verbo en Inglés.

En la figura 5.31 se presenta el Diagrama de Casos de Uso que aunque de facil comprensión muestra la manera en que el usuario interactua con el sistema.

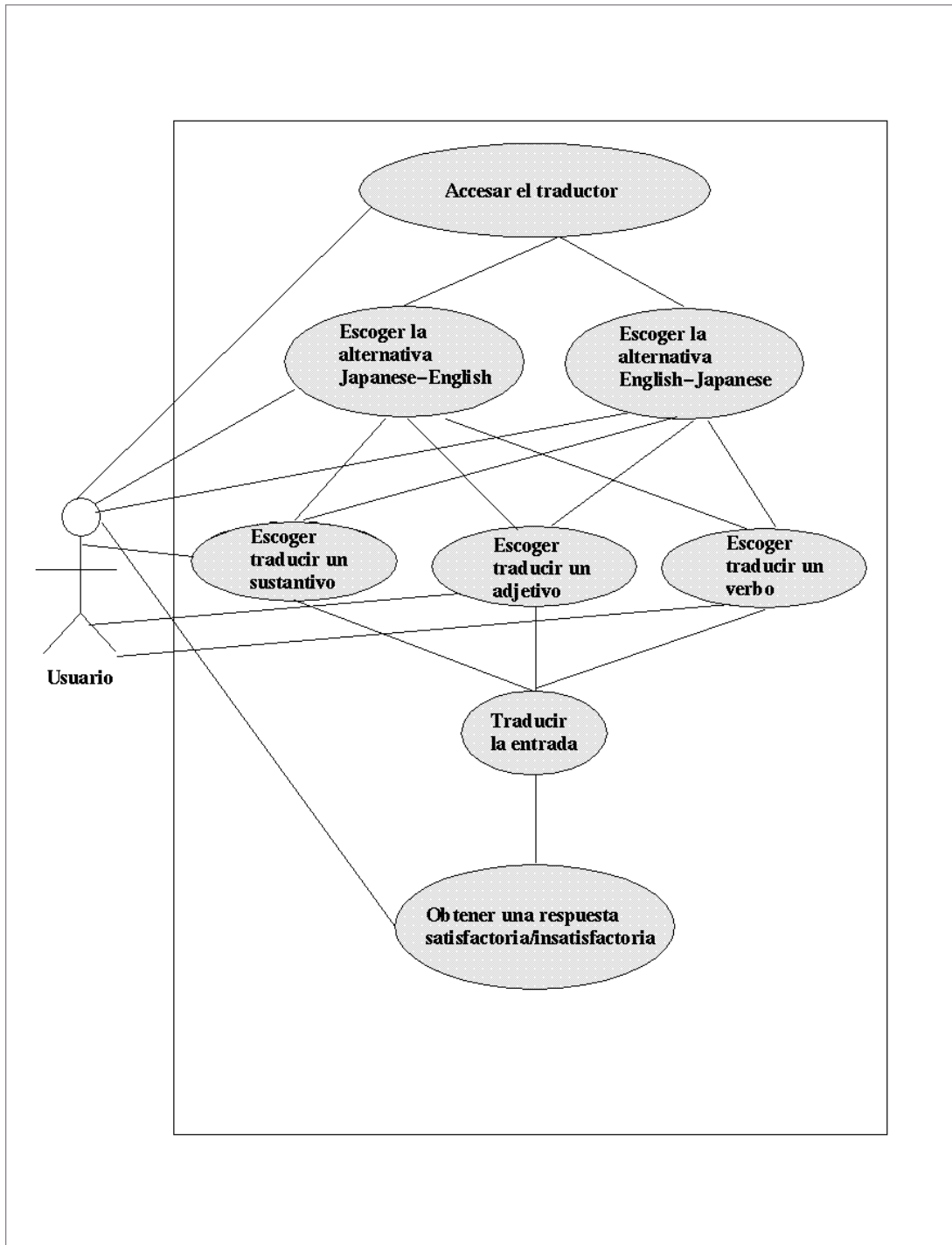


Figura 5.31 Diagrama de Casos de Uso.

En la figura 5.32 y a manera de ejemplo se presenta el Diagrama de Colaboración para la tarea de traducir un sustantivo en inglés como entrada .

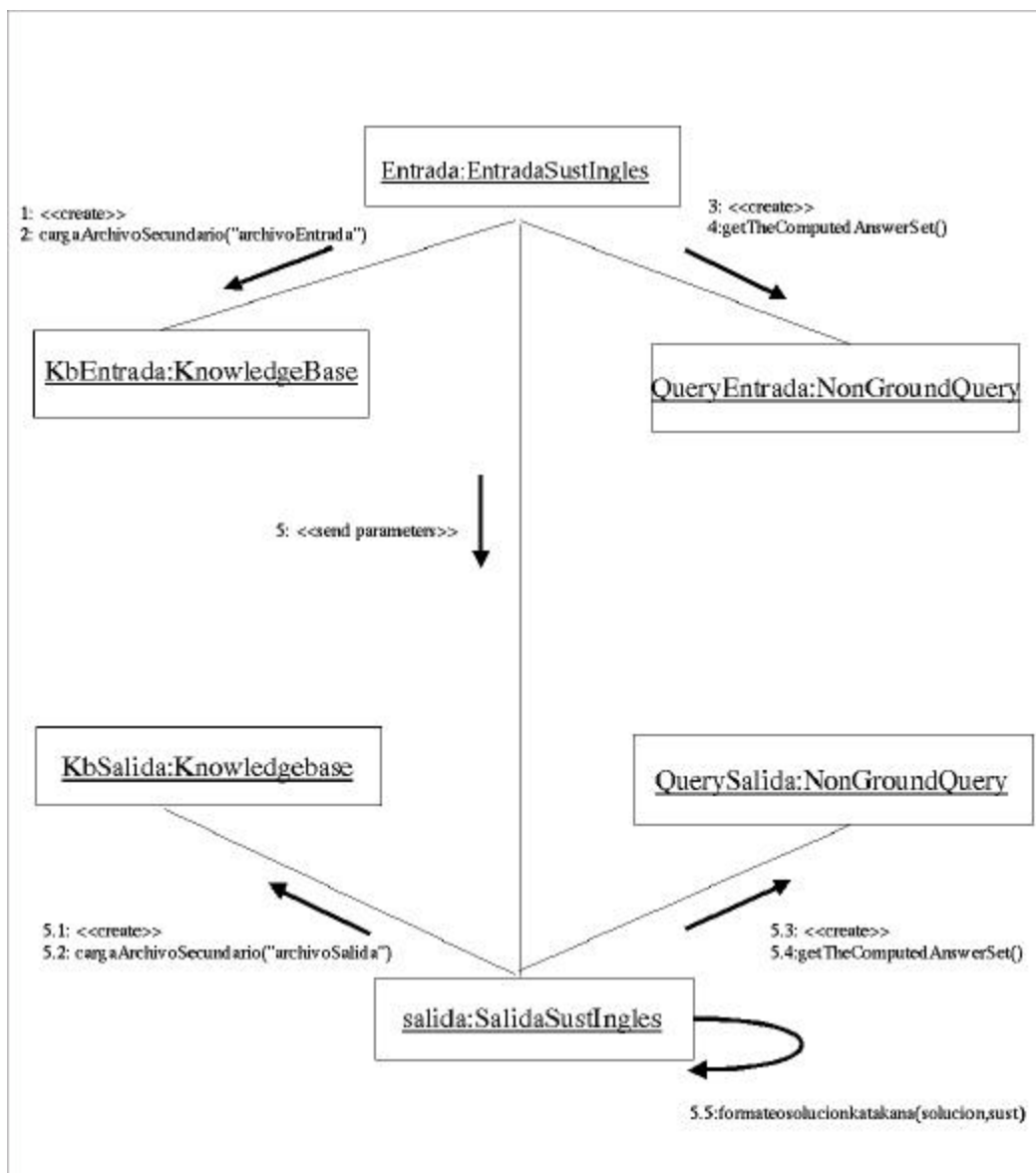


Figura 5.32 Diagrama de Colaboración.

Finalmente en la figura 5.33 presentamos el Diagrama de Secuencia con el ejemplo anterior.

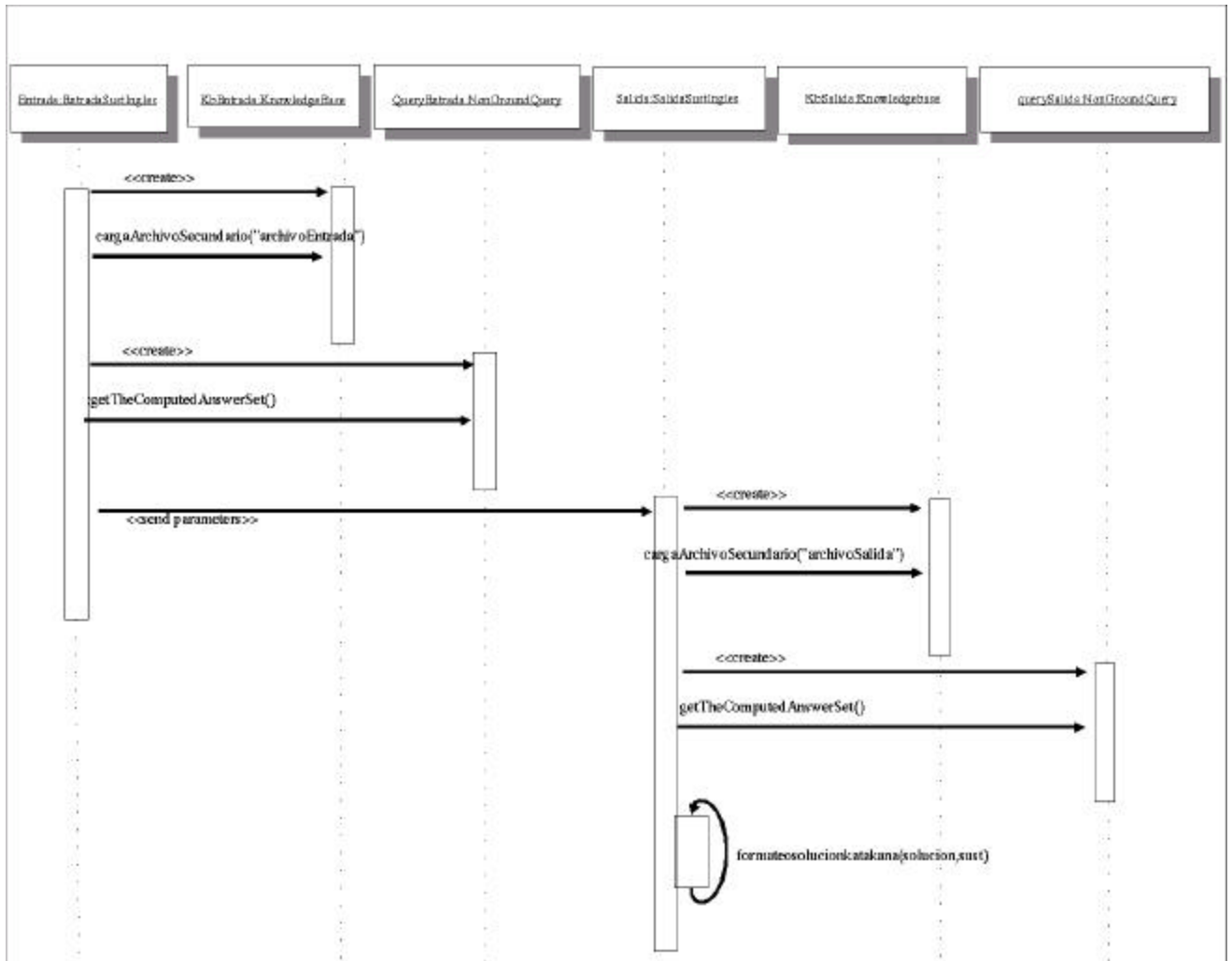


Figura 5.33 Diagrama de Secuencia.