

Metodología para realizar evaluación de respuestas a preguntas escritas en formato de texto

Methodology for evaluation of answers to written questions in text

María Alejandra Paz Menvielle, Mario A. Groppo, Marcelo M. Marciszack, Martín Casatti
Departamento de Ingeniería en Sistemas de Información,
Universidad Tecnológica Nacional, Facultad Regional Córdoba
Córdoba, Argentina
pazmalejandra@gmail.com, proyale@groppo.com.ar, marciszack@gmail.com, mcasatti@gmail.com

Resumo — El presente documento presenta un método para analizar las respuestas escritas por alumnos en forma de texto redactado en lenguaje natural, a preguntas de un examen, con el fin de contrastar su grado de coincidencia con alternativas de respuestas suministrada por un docente. Se muestran la utilización de grafos y las técnicas que permitirán asignar valores a los conceptos y relaciones a fin de poder ponderar la respuesta suministrada por el alumno y compararla con la ponderación de la respuesta base elaborada por un docente. A partir de allí se obtendrán los distintos grados de acierto que pueda tener la respuesta del alumno.

Palabras clave: análisis de textos; grafos; detección de patrones; reconocimiento; detección de rutas.

Abstract — This paper presents a method for analyzing answers written by students as a text in natural language, to questions from an exam, to ascertain their degree of overlap with alternative answers provided by a teacher. Its explain the use of graphs and techniques that allow assigning values to the concepts and relationships in order to ponder the answer provided by the students and compare it with the weight of the base reply prepared by a teacher is. From there you get varying degrees of success you may have the student's response.

Keywords - text analysis; graphs; pattern detection; recognition; detection routes.

I. INTRODUCCIÓN

El presente trabajo se desarrolla en el marco del proyecto de investigación y desarrollo homologado por la Secretaría de Investigación, Desarrollo y Posgrado de la Universidad Tecnológica Nacional.

Debe dar cumplimiento en simultáneo a dos grandes requisitos. Por un lado los contenidos mínimos fijados para la asignatura Paradigmas de Programación de la carrera Ingeniería en Sistemas de Información, los cuales pertenecen al bloque de

tecnologías básicas dentro del área programación, que están principalmente referidos a los paradigmas lógicos, funcional y de orientación a objetos. Y por otro lado cumplir con los descriptores y criterios de intensidad de formación práctica de la Resolución Ministerio de Educación, los que se encuentran definidos en el área de tecnologías básicas, sub-área programación que incluyen a los paradigmas y lenguajes de programación.

El presente proyecto surge a partir de la necesidad de desarrollar un sistema que represente de manera adecuada y simple el dominio de la cátedra Paradigmas de Programación perteneciente a la carrera Ingeniería en sistemas de Información, y también las estructuras del lenguaje natural, con el fin de determinar si la respuesta a una pregunta es correcta, dentro del dominio especificado y aplicando las técnicas y reglas gramaticales relacionada con los lenguajes estructurados [4].

Se pensó en aplicar la Teoría de Grafos según la cual es posible construir un grafo que cumpla con la siguiente ecuación: $G = (V, A)$

En donde V es el conjunto de todos los términos pertenecientes al dominio en cuestión, constituyendo los vértices que se denominarán t_1, t_2, \dots, t_n , y A es el conjunto de todas las relaciones válidas entre dos términos cualquiera.

Dados entonces dos términos, t_1 y t_2 existirá una línea a_{ij} que une dichos términos, si y sólo si, los términos están relacionados.

Si las relaciones no son bidireccionales se utilizan grafos dirigidos para representar el sentido de dicha relación. [2]

Existen dos tipos de enfoque relacionados a la búsqueda de patrones dentro de información almacenada en forma de grafos.

El primer enfoque, es la detección de sub-grafos [5], donde se buscan ciertas estructuras, contenidas en un grafo de mayor tamaño. Generalmente el resultado de dichas búsquedas es

binario y simplemente se determina si la estructura se encuentra, o no, en el grafo destino.

El segundo enfoque es el de búsquedas inexactas, en donde se debe especificar al algoritmo, además de la estructura que se desea encontrar, un cierto umbral que debe ser considerado si la estructura exacta no se encuentra. Teniendo en cuenta este umbral, el algoritmo busca los elementos de la estructura especificados con cierto grado de semejanza, por ejemplo con conceptos faltantes, con relaciones similares pero no idénticas, o con órdenes alterados o invertidos en las relaciones [6].

Debido a la naturaleza inherentemente variable de las respuestas registradas en exámenes al mismo conjunto de preguntas, se utilizará el enfoque de búsqueda inexacta, para la determinación de la validez, total o parcial, de las respuestas obtenidas.

II. MODELADO DEL CONOCIMIENTO

Los conceptos impartidos en la materia se representaron a través de una estructura de grafos, en la forma de un grafo dirigido, con atributos que amplían el contenido semántico de las entidades involucradas.

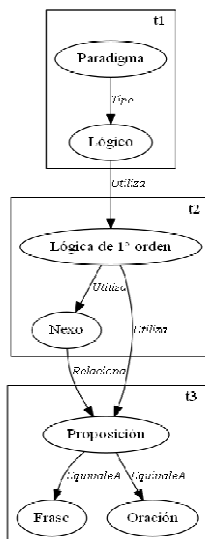


Figura 1: Ejemplo de evolución de encadenamiento de nodos en momentos t1, t2,t3

Se partió del estudio de la modalidad académica de la materia para determinar los ejes fundamentales y las unidades temáticas contenidas en la misma, así se definió el primer conjunto de nodos, de nivel cero, que comprendían los conceptos desde los cuales se derivarían el resto de los datos. El resto de los nodos se fue encadenando de forma natural como parte de la evolución normal del dictado de la materia. Véase ejemplo de la Figura 1.

Un camino similar se siguió con las relaciones que surgieron entre dichos nodos, las mismas se fueron definiendo a medida que las asociaciones entre conceptos se hacían necesarias.

La teoría de grafos en la que se basa este modelo [2] [3], establece que cada nodo contiene un conjunto de etiquetas diseñado para representar correctamente el concepto asociado y, a la vez, facilitar la posterior búsqueda.

En la figura 2 se pueden apreciar algunos de los tipos de relación que actualmente intervienen, en el ejemplo se trabaja sobre conceptos relacionados con el Paradigma Orientado a Objetos (POO).

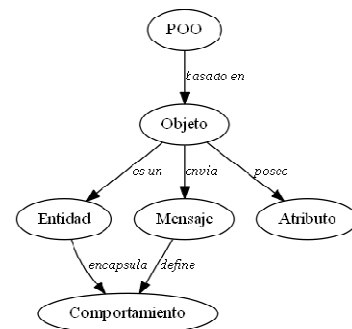


Figura 2: Tipos de relaciones

III. MECANISMO DE BÚSQUDA

A. Corrección ortográfica y gramatical

Se debe comenzar realizando una revisión ortográfica y sintáctica del texto respuesta a fin de evitar la búsqueda de términos con errores o relaciones inválidas. La misma se basa en una revisión de palabras, contra un diccionario de idioma español y se verifican las reglas gramaticales básicas, antes de suministrar el texto al motor de búsqueda.

B. Simplificación del texto

El texto respuesta, previamente validado en su ortografía, se divide en tokens o unidades atómicas, y se determinará si son conceptos o relaciones previamente guardadas en el grafo.

El motor de búsqueda devuelve la misma lista de tokens con marcadores que indican si cada uno de los términos existe, no existe o es una equivalencia de un concepto existente (véase Figura 3). El usuario decidirá qué acción tomar con respecto a los términos inexistentes. Puede descartarlos de la consulta, reformularlos o solicitar que se agreguen como conceptos nuevos en la próxima pasada. De esta forma, al ir realizando consultas sobre el grafo también se pueden ir realizando los pasos necesarios para que la base de conocimientos se mantenga actualizada.

C. Búsqueda de rutas

Una vez que el motor no tiene más sugerencias al respecto de los términos solicitados se procederá a la búsqueda de las rutas válidas que interconectan los conceptos.

»	A	B	C	D	E	F	Vector inicial
«	A	B	C	D	E	F	B, D, E (no existen)
»	A	B[+]	C	H		F	Agregar B. Reemplazar D por H. Descartar E
«	A	B	C	H		F	B Agregado. H existe.
»	A	B	C	H		F	Vector final de búsqueda

Figura 3: Ajuste del vector de búsqueda

Para ellos se utilizarán mecanismos iterativos y recursivos para obtener las listas de relaciones que unen los conceptos o nodos. El objetivo de este paso no es obtener una única ruta, ni la mejor ruta, sino el conjunto de todas las rutas que contienen los conceptos indicados, sin discriminar los tipos de relaciones que los unen.

D. Valoración y ponderación de conceptos y relaciones

Considerando que la búsqueda de relaciones entre conceptos deberá ser empleada para poder evaluar el grado de validez de una respuesta, se planteó la necesidad de comparar las posibles respuestas con una respuesta base, definida como correcta. De esta forma, todas las posibles rutas serán comparadas con esta respuesta base y se computará su grado de validez en base a esa comparación.

Los pasos necesarios para realizar la valoración de respuestas son los siguientes:

Para cada ruta encontrada, la llamaremos ruta candidata, se aplicará un algoritmo cuyo objetivo es contar la cantidad de conceptos y relaciones exactas que la componen. Se considerarán conceptos y relaciones exactas a aquellas que se hallan con el mismo nombre y en la misma posición que en la respuesta base.

Si la respuesta candidata tiene los mismos conceptos y relaciones, en las mismas ubicaciones y con los mismo nombres que en la respuesta base, se la denomina "respuesta perfecta" y tendrá un valor que se calcula como: $V_r = C + R$

Siendo:

- V_r : Valor de la respuesta
- C : Cantidad de conceptos
- R : Cantidad de relaciones

Cualquier otro caso implicará que hay conceptos o relaciones inexactas (con valores ponderados < 1) y se requerirá un análisis diferente para obtener la mejor respuesta.

A continuación se describen los distintos casos que se pueden presentar al momento de hacer las valoraciones de conceptos y de relaciones.

1) Valoración de los conceptos:

Caso 1: Los conceptos son coincidentes pero las ubicaciones dentro de la respuesta difieren de la respuesta base. En este caso se calculará para cada concepto inexacto el desplazamiento con respecto a la respuesta base. Se considera que mientras más alejado esté el concepto de su ubicación correcta, su incidencia en la respuesta disminuye [7]. Es por ello que el valor del concepto se ponderará de la siguiente forma:

$$C_d = 1 - \left| \frac{P_c - P_b}{n} \right|$$

Siendo:

- C_d : Valor del concepto desplazado
- P_c : Posición del concepto candidato
- P_b : Posición del concepto en la respuesta base
- n : Cantidad de conceptos en la respuesta

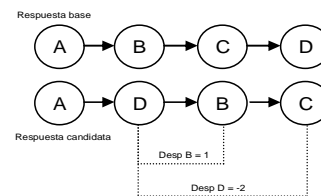


Figura 4: Desplazamiento entre conceptos

Caso 2: Los conceptos no son exactos en el nombre, sino que son equivalentes a los conceptos de la respuesta base.

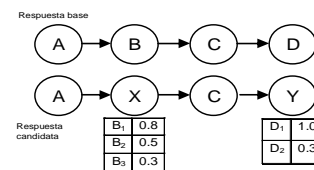


Figura 5. Peso de conceptos no-exactos

En este caso, cada concepto tendrá una tabla de equivalencias, definida por el docente, cuyo fin será ponderar cada una de las equivalencias a un valor numérico entre 0 y 1, donde 1 será una equivalencia completa, que indicará que ambos términos pueden usarse indistintamente. De esta forma el valor de C_e es el valor que se obtiene de las tablas de equivalencia, para el concepto dado tal como se observa en la figura 6.

Los casos 1 y 2 descriptos se podrán combinar cuando existan conceptos que no son exactos en el nombre, y además se encuentren en ubicaciones diferentes a las definidas en la respuesta base, ponderando el concepto candidato según:

$$C_{de} = C_d \cdot C_e$$

Siendo:

- C_{de} : Valor del concepto desplazado equivalente
- C_d : Valor del concepto desplazado

C_e : Valor del concepto equivalente

2) *Valoración de las relaciones*: Las relaciones deben ser valoradas de una forma similar a los conceptos para obtener la valoración total de la respuesta. Se han detectado distintas situaciones o casos que serán descritos a continuación.

Caso 1: La relación es exacta y los nodos de origen y destino son exactos. En este caso se considerará una relación exacta pura y su valor será igual a 1.

Caso 2: La relación es exacta pero un nodo o ambos, nodo origen y nodo destino, no son exactos. En este caso el peso de la relación se reducirá, debido a que no hay forma de asegurar que la relación original siga siendo válida al haber cambiado la exactitud de los nodos asociados. En este caso el valor de la relación se calculará como:

$$V_r = \frac{(1 + C_o + C_d)}{3}$$

Siendo:

V_r : Valor de la relación

C_o : Valor del concepto de origen

C_d : Valor del concepto de destino

Caso 3: La relación no es exacta pero los nodos de origen y destino si lo son. En el caso en que la relación no es exacta se deberá analizar la lista de equivalencias suministrada por el sistema. Esta lista contendrá los términos equivalentes con su ponderación asociada para una relación dada en la respuesta base. Si la relación en la respuesta candidata se encuentra dentro de la tabla de equivalencias, el valor R_e de dicha relación, se obtendrá directamente de la tabla.

Caso 4: La relación no es exacta y los nodos de origen y/o destino tampoco lo son. En este caso se deberá utilizar también la tabla de equivalencias pero los pesos ponderados de los nodos de origen y destino podrán ser alterados de acuerdo a sus propias equivalencias.

Aquí el valor de la relación se definirá como la media matemática entre los valores de los nodos de origen y de destino y el valor ponderado de la relación equivalente.

$$V_r = \frac{(R_e + C_o + C_d)}{3}$$

Siendo:

V_r : Valor de la relación

R_e : Valor de la relación equivalente

C_o : Valor del concepto de origen

C_d : Valor del concepto de destino

En el ejemplo de la figura 6, la relación: $A_1 \rightarrow R_1 \rightarrow D_2$

Tendría un valor ponderado de:

$$V = \frac{(A_1 + R_1 + D_2)}{3}$$

$$V = \frac{(0.7 + 0.9 + 0.3)}{3} = 0.633$$

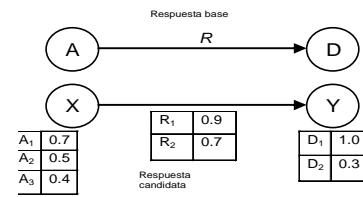


Figura 6. Relaciones y conceptos no-exactos

IV. EVALUACIÓN

Una vez realizado el cálculo del valor de la respuesta, el valor obtenido variará entre 0 y el valor de la respuesta base, si son iguales implicaría una respuesta correcta "perfecta" con un grado de coincidencia conceptual del 100%.

Los valores intermedios indicarán el grado de aproximación de la respuesta suministrada a la respuesta base y podrán ser útiles como indicadores para calcular el porcentaje de acierto de la evaluación en que la respuesta analizada estará incluida. En tal sentido será posible fijar dos umbrales (mínimo y máximo) a partir de los cuales se considerará a las respuestas.

Si el valor obtenido no se encontrara dentro de los extremos mínimo y máximo establecidos, la respuesta debería ser evaluada por un docente a fin de colocar a la misma una calificación correcta en base a una interpretación o un contexto que exceden aquellos representados por la base de conocimientos disponible, y por ende, al mecanismo de evaluación automatizado expuesto.

V. CONCLUSIONES

Se ha presentado un método que permitirá calificar exámenes formados por preguntas que se responderán como ensayo escrito en forma de texto libre, que emplea la teoría de grafos para construir el dominio de trabajo, y se corresponde con el área de conocimiento de la cátedra Paradigmas de Programación perteneciente a la carrera Ingeniería en Sistemas de Información.

REFERENCIAS BIBLIOGRÁFICA

- [1] J Glenn Brookshear, Teoría de la Computación, Addison Wesley, 1989.
- [2] Sowa, John F, "Semantic networks" en Encyclopedia of Cognitive Science, 2006.
- [3] Sowa, John F. "Conceptual graph summary", en Conceptual Structures: Current Research and Practice. Ellis Horwood, New York London Toront, 1992, pp. 3-66.
- [4] Hopcroft, Motwani, Ullman. Teoría de autómatas, lenguajes y computación. Pearson, Addison-Wesley, 2008.
- [5] Pavlidis, Theodosios, Structural pattern recognition. Vol. 2. New York: Springer-verlag, 1977.
- [6] Olmos, Ivan, Jesus A. Gonzalez, and Mauricio Osorio. "Inexact Graph Matching: A Case of Study", en FLAIRS Conference, 2006.
- [7] Buckley, Fred, and Frank Harary. Distance in graphs. Addison-Wesley Longman, 1990.