

Rastreo de la Información en Páginas Web a través del Significado

(Foraging Information in Webpages through Meaning)

Francisco Torres Guerrero & Ernesto O. López Ramírez*

Resumen. Aunque muchas teorías cognitivas están de acuerdo en que los mecanismos actuales por los cuales el cerebro procesa datos visuales y espaciales son utilizados para el rastreo de información en páginas web, es de interés en la comunidad de investigación de hipermedia dentro de la importancia de la exanimación de las representaciones mentales de los usuarios respecto al rastreo de información en páginas web. Este estudio tiene la intención de alcanzar a entender como el significado que un individuo comparte sobre una temática influye en su forma del rastreo de información en medios de tecnología de información. El grupo estuvo conformado por 40 estudiantes cuyas edades oscilaban entre 17 y 22 años donde aproximadamente la mitad eran mujeres y el resto hombres. Se utilizaron técnicas de representación del conocimiento así como un ambiente web para ejercicios de rastreo de información. De relevancia fue el hecho de que son los mismos contenidos conceptuales que parecen subyacer a ambas organizaciones de información lo cual apoya la idea de un significado único como el elemento central.

Palabras claves. Rastreo de información, representación del conocimiento, significado, interacción humano computadora.

Abstract. Although many cognitive theories agree that the current mechanisms by which the brain processes visual and spatial data are used to track information on websites, is of interest in the hypermedia research community in the importance of exanimation of mental representations of users about the tracking of information on websites. This study intends to reach out to understand as the meaning that an individual shares on a topic influences the way of tracking information means information technology. The group consisted of 40 students aged between 17 and 22 years with approximately half were women and the rest men. It used techniques of knowledge representation and a web environment exercise tracking information. Of relevance was the fact that they are the same conceptual contents that seem to underlie both organizations to information which supports the idea of a unique significance as the central element.

Keywords. Foraging information, knowledge representation, meaning, human computer interaction.

Introducción

En la actualidad existen miles de millones de documentos en internet los cuales son accedidos constantemente por alrededor de 300 millones de usuarios o mas (Davies, Fensel & Van Harmelen, 2003). La interacción humano – páginas web es fundamentalmente una actividad de procesamiento de información. Dentro de esta interacción con un sistema de información dentro de un entorno web el usuario tiene metas y sub metas específicas a lograr en dicha interacción (Sears & Jacko, 2008; Kaluzniacky, 2004). Aquí, la naturaleza cognitiva del usuario que interviene para el

rastreo de información en plataformas de internet demanda procesos de toma de decisiones (Palmquist & Kyung 2000; Hayes & Allinson 1998; Chen & Rada, 1996; Pirolli, 2007), pero por otra parte el uso del significado que un usuario tiene sobre un tema, parece orientar el procesamiento de información para la solución de una actividad de rastreo de información. Esto ha generado líneas de investigación científica para el estudio teórico y experimental dentro de la denominación de ciencias cognitivas, desarrollo de hipertexto, desarrollo multimedia, inteligencia artificial, sistemas educativos, etc. Que pretenden establecer el rol del significado humano dentro del rastreo de información en páginas web (Konar, 2000; Thagard, Gabbay & Woods, 2007).

Desde un enfoque cognitivo la formación de un significado se da por la activación de organizaciones conceptuales que un individuo posee en su memoria a largo plazo (López, 2002; Murphy, 2002; Rogers & McClelland, 2004) y existen ya intentos académicos para tratar de determinar cómo es que dichas organizaciones conceptuales y significado impactan el rastreo de información, como por ejemplo en sistemas de hipertexto (Jonassen, 2000, 2002). De esta línea académica se ha desprendido como uno de los problemas centrales el encontrar herramientas que permitan determinar la organización conceptual del usuario mientras rastrea información. Este problema es de interés al presente documento y se detalla una alternativa al respecto en la sección de método.

En el estado inicial de significado que un usuario posee al rastrear información de un tema en un momento dado inicial (t) en el proceso de un rastreo de información existen dos representaciones conceptuales en juego: La organización conceptual dada la esencia de la información contenida en una página web y la conceptualización del usuario que implica el significado que este impone sobre el tema. Aún y cuando ambas representaciones conceptuales pueden ser idénticas (Isomórficas) es muy probable que sea el caso en que ambas organizaciones conceptuales sean diferentes pero estén hablando de lo mismo (representaciones homomórficas; Holland, Holyoak, Nisbet & Thagard, 1989). Tal y como se ilustra en la Figura 1 es probable que exista una relación (R) entre la representación mental de un usuario y el contenido de una página que inicialmente en el rastreo de información dicha relación imponga estructura sobre la forma de búsqueda de un objeto de información en la web. Esta relación inicial, es una estrategia cognitiva que usa en todo momento el conocimiento existente en el usuario pero que es capaz de usar al mismo momento la esencia de información en la página web. Sin embargo, a diferencia de la web en donde la transición de una organización conceptual a otra (T) está básicamente establecida, la regla de transición en el usuario (T') depende exclusivamente de la forma en cómo el individuo va

significando la información que accesa. El objetivo es la formación de un modelo mental (por ejemplo $t+1$, $t+2\dots$) que le permite comprender el objeto de información que rastrea.

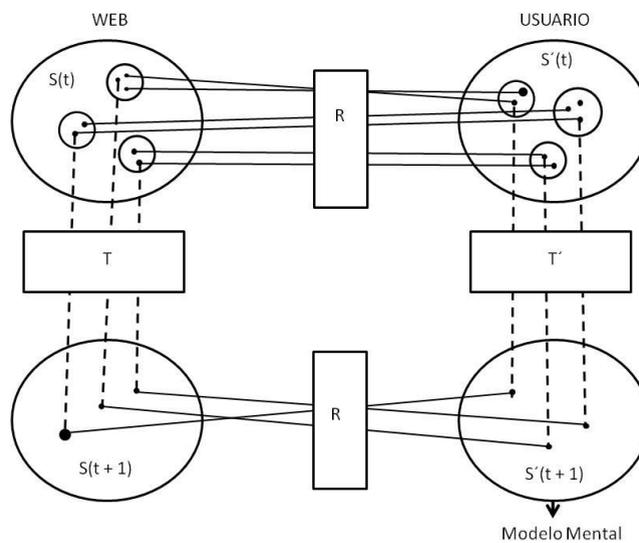


Figura 1 Modelo hipotético en el que el rastreo de información se puede considerar como una estrategia cognitiva (R) que toma en cuenta organizaciones conceptuales homomórficas (representados por puntos unidos) y sus transiciones entre estados (S) de representación conceptual para la formación de un modelo mental del tema que se analiza.

Ahora bien de la Figura 1 se desprende que si $T'R = RT$ entonces existe una equivalencia en representaciones conceptuales aún y cuando estas sean homomórficas. Dicha conmutabilidad es por supuesto al servicio de la creación de un significado final que se constituye como un modelo mental válido del evento que se representa. La no conmutabilidad de esta equivalencia implica la formación equivocada de un significado sobre la temática. De esta forma, el rastreo estratégico de información (implicando R) lleva al usuario a través de estados de representación conceptual S que lo conllevan a un modelo mental que constituye una equivalencia de clases conceptuales entre la información del sitio web y el modelo mental final del usuario. Dicho de forma más formal la ruta de transformación o transición puede ser especificada como:

$$R[T(S(t), O(t))] = T'[R(S(t), O(t))] \quad (1)$$

De tal forma que $O(t)$ se relaciona a la organización conceptual resultante dado una secuencia de rastreo de información. De relevancia es el hecho de que anteriormente no existía una forma específica de determinar $O(t)$ para T' . Una intrigante posibilidad a este respecto se puede visualizar por indicadores de densidad semántica, riqueza de red y/o conectividad de técnicas modernas de análisis de redes semánticas que se basan en información obtenida de individuos y no de forma idiosincrática. Estas pueden ser usadas para determinar R , usando dichos indicadores de la organización semántica en $O(t)$ del usuario.

De esta forma R debe estar supeditada a una regla impuesta por $T'[R(S(t), O(t))]$. Dado que esta regla debe constituirse de una capacidad para representar la dinámica de transformación de estados conceptuales entonces la plausibilidad para usar notación de teoría de autómatas se considera.

Un sistema autómata se compone de una quintupleta de elementos $A = (Q, \Sigma, \delta, q_0, F)$ Donde Q es un conjunto de estados posibles, Σ representa un conjunto estímulos posible al sistema (símbolos), δ es una función que determina la transición de un estado a otro, q_0 es un estado inicial y F un conjunto de estados finales aceptables. De esta forma el caso de un autómata en su estado inicial, q_0 que posee una función de transición δ tal que para pasar de un estado a otro (por ejemplo q_1) entonces $\delta(q_0, a_1) = q_1$ donde dicho autómata procesa su primera cadena estímulo a_1 del conjunto Σ . Si q_n es un miembro de de los estados aceptables F , entonces las entradas a_1, a_2, \dots, a_n son aceptadas sino es rechazado.

Ahora bien supongamos que W es una cadena de caracteres alfanuméricos tal que $W = 1101$ está en realidad compuesto de dos cadenas $x = 110$ y $a = 1$. Entonces

$$\delta'(q, w) = \delta(\delta'(q, x), a)$$

Esto es, para realizar una transición $\delta'(q, w)$, el autómata debe primero procesar $\delta'(q, x)$, que es el estado previo a procesar al entrada "a" (mayor información a este respecto en Hopcroft, Motwani & Ullman, 2001).

Si extrapolamos la notación de autómatas a la conceptualización de representación conceptual de un usuario durante el rastreo de información entonces podemos ver que podemos reconceptualizar la actividad conceptual de un usuario. En particular es posible implementar de 1 la siguiente conmutabilidad:

$$\delta'(q, w) = \delta(\delta'(q, x), a) = T'[R(S(t), O(t))]$$

Nótese de la igualdad que T' entonces implica una transición que condiciona dicha transformación a la consideración de un estado conceptual previo que se une a la siguiente organización, esto es, dicha unión de ambos estados debe estar permitida por el autómata que en el caso del usuario demanda la participación de un significado conceptual que regula la función de transición $\delta'(q, w) = T'$. A este respecto se desglosa a continuación una serie de estudios que pretenden ejemplificar el uso de dicha función de transición desde una perspectiva representacional en el rastreo de la información en páginas web.

Método

A continuación se describen tres partes de un estudio de ciencia cognitiva del rastreo de la información. La primera parte se constituye de un estudio de representación del conocimiento de un usuario de internet que pretende rastrear información sobre ecología en un sitio web. De dicho estudio se deriva la segunda parte que se constituye de un estudio de simulación computacional que permite determinar la proximidad semántica entre conceptos pertenecientes al esquema conceptual de ecología (Eco-Esquema). Dicha organización conceptual permitió establecer los elementos de los conjuntos Σ , Q y F que pueden ser considerados en la función de transición del autómata determinando el rastreo de información del usuario mientras navega en un sitio web. También se describe brevemente la construcción del sitio web.

- Estudio de representación de conocimiento.

En este estudio, a través de una técnica de redes semánticas naturales, se obtuvieron de los participantes definiciones conceptuales de 7 conceptos centrales al esquema de medio ambiente (Eco-esquema).

Participantes

La muestra objetivo que se eligió para el presente estudio fueron estudiantes de ingeniería debido a su experiencia en el uso de sistemas de información en modalidad web para el rastreo de información. De esta forma, la muestra fue por conveniencia. El grupo estuvo conformado por 40 estudiantes cuyas edades oscilaban entre 17 y 22 años donde aproximadamente la mitad eran mujeres y el resto hombres.

Instrumentos

Para la obtención de las redes conceptuales se desarrolló un software denominado SemNet (De la Garza, Sanchez & López, 2010) que es capaz de establecer la relevancia semántica de conceptos obtenidos a través de una técnica de redes semánticas naturales (Figueroa, Gonzales & 1981, López, 2002). En esta técnica se le pide bajo ciertos controles a los participantes generar conceptos definidores (“animales”, “arboles”, etc.) que mejor definan un concepto objetivo (por ejemplo NATURALEZA). SemNet permite obtener valores tradicionales de riqueza conceptual usadas para definir un concepto objetivo (valor J), densidad semántica entre conceptos (valor G), relevancia semántica de un concepto definidor (Valor M) y la elección de los 10 definidores más importantes para cada concepto objetivo (Grupo SAM).

Los estímulos objetivos a definir por los participantes fueron: CIUDAD, CONTAMINACIÓN, ECOLOGÍA, MEDIO AMBIENTE, NATURALEZA, RECICLAJE, REFORESTACION. Estos fueron obtenidos de jueces expertos en la materia.

- Simulaciones computacionales del eco-esquema.

En este estudio se usaron las definiciones conceptuales de los participantes para implementar un análisis de simulación de la eco-representación desde la perspectiva de dos modelos de memoria humana.

En particular, SemNet permite usar los grupos SAM obtenidos de los participantes a través de la técnica de redes semánticas antes mencionada para implementar un modelo conexionista de memoria llamado SASO (López & Theios, 1992). El modelo calcula la probabilidad de co-ocurrencia entre dos conceptos “X” y “Y” a través de los grupos de definiciones conceptuales SAM considerando de la siguiente fórmula:

$$W_{ij} = -\ln \{ [p(X=0 + Y=1) p(X=1 + Y=0)] * [p(X=1 + Y=1) p(X=0 + Y=0)]^{-1} \}$$

La matriz simétrica de pesos de asociación entre los diferentes conceptos resultante permite patrones de activación conceptual del eco esquema de acuerdo a un modelo conexionista de “cerebro en caja” que converge a soluciones usando un algoritmo Boltzman (Rumelhart, Smolensky, McClelland & Hinton, 1986).

La matriz de pesos de asociación entre conceptos dada la co-ocurrencia conceptual, permitió determinar la proximidad semántica entre conceptos usando un

algoritmo denominado análisis Pathfinder (Schvaneveldt, 1990). Pathfinder converge a una solución de mínimo de conexiones semánticas entre dos conceptos (proximidad semántica). Usando una aproximación Minkowski:

$$L(P) = \left(\sum_{i=1}^n l_i^r \right)^{\frac{1}{r}}, 1 \leq r \leq \infty$$

Donde l_i es el peso asociado con la liga i en un camino. El conjunto de todos los pesos en un camino con n ligas estaría dado por $l_i=1,2,\dots,n$. Conforme el valor de r varía sobre un rango permitido el número de ligas en la red resultante varía sistemáticamente.

De esta forma, usando un sistema software denominado KNOT es posible realizar representaciones gráficas de organización conceptual dada la proximidad semántica entre conceptos tal y como se muestra en la sección de resultados. Aquí cabe hacer énfasis que la misma técnica de redes semánticas usada para la obtención de las definiciones conceptuales pudo ser usada para construir definiciones conceptuales basadas en los conceptos activados en las páginas web para así obtener grupos conceptuales SAM resultantes del rastreo de la información. Esto se describe a continuación.

- Construcción de los escenarios de las páginas web usadas para el rastreo de la información sobre conceptos de un eco-portal.

Utilizaremos como base el sitio Ecoportal (<http://www.ecoport.net/>) ya que en este sitio fue diseñando con la meta difundir información con respecto a la ecología donde el alcance va desde cuestiones básicas de información hasta noticias actuales e información de entidades involucradas. Se desarrolló una página que sirviera como marco para monitorear cuestiones de rastreo la cual se accesaba desde la página (<http://www.ciencias-cognitivas.net/>) donde se agregaron dos elementos visuales adicionales a la página de ecoportal, los cuales fueron: El recuadro de la temática, el cual indicaba que tema buscar. Por otra parte existía la opción para pasar al siguiente concepto a rastrear y el evaluador de liga. Esta liga permite dar una calificación a la liga visitada antes de visitar la próxima página seleccionada. El usuario visualizaba la página tal y como aparece en la Figura 2.

The screenshot shows the 'ECO PORTAL.NET' website. At the top, there is a navigation bar with links for 'Cursos', 'Campus', 'Foros', 'Blogs', 'Especialistas', 'ONGs', 'Glosario', 'Anuncios', 'Enlaces', 'Campañas', 'GIFs', and 'Vegetarianos'. Below this is a secondary navigation bar with categories like 'Biodiversidad', 'Cambio Climático', 'Contaminación', etc. The main content area is titled 'Temática: Contaminación' and includes a link 'Pasará al siguiente rastreo'. A modal dialog box is open, asking for a rating of the page visited, with a scale from 1 (lowest) to 10 (highest). The background content includes sections for 'COCINA VEGETARIANA', 'Mapa del Sitio', 'Temas Especiales', 'Servicios', 'ÚLTIMAS NOTICIAS', and 'AMBIENTE Y SOCIEDAD'.

Figura 2. Pagina de navegación.

El procedimiento para el rastreo de la información consistía en que el usuario entrara a una página de inicio en la cual ingresaba sus datos generales. Una vez ingresados los datos presionaba el botón de comenzar. El sistema despliega la información de la página web ecoportal así como también el recuadro con la tarea de rastreo a buscar. El usuario empezaría a rastrear información relevante al tema presionando las ligas que considera relevante. El sistema antes de desplegar la siguiente liga le pregunta al usuario por una calificación en base a la relevancia del sitio revisado. El usuario captura la calificación y presiona ok ocasionando que el sistema guarde esta información en la base de datos. En el caso que el usuario presiona cancel significa que no ha terminado de ver la página y permanece en ella.

Resultados

La Figura 3 muestra las definiciones conceptuales para los 7 conceptos considerados dentro del esquema del medio ambiente que se obtuvieron de los participantes del estudio.

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE			NATURALEZA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
2	CARROS	325	1	HUMO	263	3	RECICLAR	151	4	NATURALEZA	186	3	AGUA	291
1	EDIFICIOS	205	2	BASURA	241	4	AMBIENTE	150	3	AGUA	127	3	ANIMALES	279
2	CIUDAD	195	2	CARROS	177	4	NATURALEZA	128	3	ANIMALES	112	4	ARBOLES	244
3	CONTAMINACION	168	3	CONTAMINACION	152	3	ANIMALES	128	2	ECOSISTEMAS	106	2	PLANTAS	179
1	CALLES	160	2	FABRICAS	113	2	PLANTAS	117	3	CONTAMINACION	105	2	FLORA	140
1	CASAS	148	3	AGUA	112	4	ARBOLES	113	4	ARBOLES	103	3	AIRE	112
1	PERSONAS	139	2	CIUDAD	78	2	ECOSISTEMAS	81	3	AIRE	102	4	AMBIENTE	104
1	GENTE	138	3	AIRE	71	2	VERDE	77	1	TIERRA	81	1	FAUNA	102
2	FABRICAS	121	2	RUIDO	61	4	ECOLOGIA	77	1	CALENTAMIENTO	73	2	BOSQUES	99
2	RUIDO	62	4	AMBIENTE	58	1	RECICLAJE	62	4	ECOLOGIA	57	4	NATURALEZA	94
J = 93 G = 26.3			J = 83 G = 20.5			J = 77 G = 8.9			J = 74 G = 12.9			J = 71 G = 19.7		
RECICLAJE			REFORESTACION											
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	PAPEL	220	4	ARBOLES	333									
2	BASURA	171	2	BOSQUES	167									
1	PLASTICO	152	1	ITALA	109									
1	ALUMINIO	145	3	RECICLAR	89									
3	RECICLAR	133	1	VIDA	84									
1	LATAS	95	4	NATURALEZA	80									
1	CARTON	70	2	FLORA	74									
1	VIDRIO	69	2	VERDE	73									
4	AMBIENTE	65	1	PLANTAR	73									
4	ECOLOGIA	48	4	ECOLOGIA	67									
J = 77 G = 17.2			J = 73 G = 26.6											
Total Conceptos	7	Def. x Concep.	284	Conceptos Diferentes	38	H-S 1		H-S 2		Compara		Predictor		
Total Alumnos	40.0	Def. x Alum.	49.8	Conceptos Comunes	19(51)	Detalles Sam		Esquemata		NetScan				
Total Definidores	1,990	Def. x Alum. x Concep	7.1	Conceptos Libres	19									

Figura 3. Se describe las organizaciones conceptuales de los participantes del estudio para los 7 conceptos del esquema de medio ambiente. Nótese que los conceptos que proveen mayor conectividad (señalados por puntos) están relacionados a naturaleza (arboles, animales y agua) o aspectos de contaminación (basura, contaminación, aire). La columna F de cada grupo SAM señala la cantidad de veces que un concepto fue usado a través de todos los grupos SAM.

De las organizaciones conceptuales de los participantes se puede afirmar que poseen una organización conceptual básicamente orientada hacia el concepto de la naturaleza y aspecto de contaminación de la misma. Esto contrasta con otros estudios en los que se asume que existe mayor relación al medio ambiente (biólogos) en donde las definiciones conceptuales manejan perfiles académicos de ecología (sustentabilidad, energía renovable, etc). Sin embargo queda claro que los participantes tienen al menos un conocimiento base organizado y que permite una significación del tema dado que poseen suficiente conectividad entre los grupos SAM. Ninguno de estos grupos quedo excluido de la conectividad de la red. Todas las definiciones conceptuales presentaron riqueza conceptual (valores J mayores que 100) y es interesante el hecho que el grupo SAM de mayor densidad semántica (valor G) se relaciona a ciudad mientras que el menor densidad semántica se

relación a medio ambiente. Esto último en cierta forma tipifica el tipo de muestra que se consideró en esta investigación.

El orden de conectividad para conceptos de los grupos SAM fue el siguiente Animales (4), Arboles (4), Agua (4), Contaminación (3), Plantas (3), Basura (3), Aire (3), Tierra (3), Verde (3), Carros (2), Ruido (2), Ecosistema (2), Naturaleza (2), Vida (2) y todo los demás solo estuvieron presente en su propio grupo SAM.

Por otra parte se obtuvieron las definiciones conceptuales para el rastreo de información. Recuérdese que en este paso las definiciones fueron capturadas de los conceptos visitados a través de las diferentes opciones que presentaba la página web. Las definiciones conceptuales obtenidas se ilustran en la Figura 4.

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE			NATURALEZA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	EDIFICIOS	347	1	HUMO	349	4	ARBOLES	327	4	ANIMALES	272	3	PLANTAS	290
1	PERSONAS	317	2	BASURA	349	4	AMBIENTE	280	4	ARBOLES	269	3	AGUA	248
2	CARROS	300	2	CARROS	303	3	PLANTAS	271	4	TIERRA	256	4	ANIMALES	245
1	CASAS	293	2	FABRICAS	269	2	ECOSISTEMAS	270	2	FAUNA	238	4	ARBOLES	243
1	CONTAMINACION	285	2	RUIDO	218	4	ANIMALES	270	3	AGUA	229	2	FAUNA	231
1	GENTE	283	1	CIUDAD	198	3	NATURALEZA	208	2	ECOSISTEMAS	225	2	VIDA	224
1	CALLES	246	3	AGUA	190	2	RECICLAR	180	3	NATURALEZA	201	4	AMBIENTE	222
2	RUIDO	227	3	AIRE	181	2	FLORA	159	1	ECOLOGIA	200	3	AIRE	209
2	FABRICAS	191	1	SUCIEDAD	175	3	BOSQUES	153	3	AIRE	172	4	TIERRA	198
1	ESCUELAS	179	1	CALENTAMIENTO	140	4	TIERRA	142	3	BOSQUES	156	2	FLORA	170
J = 35 G = 16.8			J = 42 G = 20.9			J = 38 G = 18.5			J = 33 G = 11.6			J = 27 G = 12		
RECICLAJE			REFORESTACION											
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	PLASTICO	365	4	ARBOLES	358									
1	PAPEL	355	1	PLANTAR	253									
1	CARTON	315	3	BOSQUES	247									
1	VIDRIO	296	3	PLANTAS	217									
1	LATAS	289	4	TIERRA	196									
1	ALUMINIO	289	2	VIDA	177									
2	BASURA	191	3	NATURALEZA	166									
2	RECICLAR	172	1	TALA	159									
1	REUTILIZAR	152	4	ANIMALES	150									
4	AMBIENTE	105	4	AMBIENTE	125									
J = 32 G = 26			J = 25 G = 23.3											
Total Conceptos	7	Def. x Concep.	385	Conceptos Diferentes	39	H-S 1		H-S 2		Compara		Predictor		
Total Alumnos	40.3	Def. x Alum.	67.0	Conceptos Comunes	18(49)	Detalles Sam		Esquemata		NetScan				
Total Definidores	2,698	Def. x Alum.x Concep	9.6	Conceptos Libres	21									

Figura 4. Se ilustran las definiciones conceptuales para la red semántica usada en el rastreo de la información.

Los definidores comunes más relevantes para las definiciones conceptuales de rastreo fueron: Arboles (4), Ambiente (4), Animales (4), Tierra (4), Agua (3), Aire, Bosques (3), Plantas (3), Naturaleza (3).

Puede notarse de ambas definiciones conceptuales que los contenidos conceptuales en ambos grupos de definición conceptual son muy similares. Por ejemplo más del 90% de los contenidos conceptuales son los mismos. También los conceptos comunes a ambas definiciones conceptuales son muy similares por lo que se puede sugerir que tienen una misma organización conceptual.

Para determinar si ambas organizaciones conceptuales (usuario vs rastreo) poseen una organización conceptual similar dada la proximidad semántica entre sus conceptos se procedió a obtener índices de asociación conceptual a través de co-ocurrencia conceptual y un análisis Pathfinder sobre dichos pesos de asociación tal y como se describió en la sección de método.

La Figura 5 así como la Figura 6 muestra de forma gráfica (a través del sistema KNOT) el resultado para cada organización conceptual.

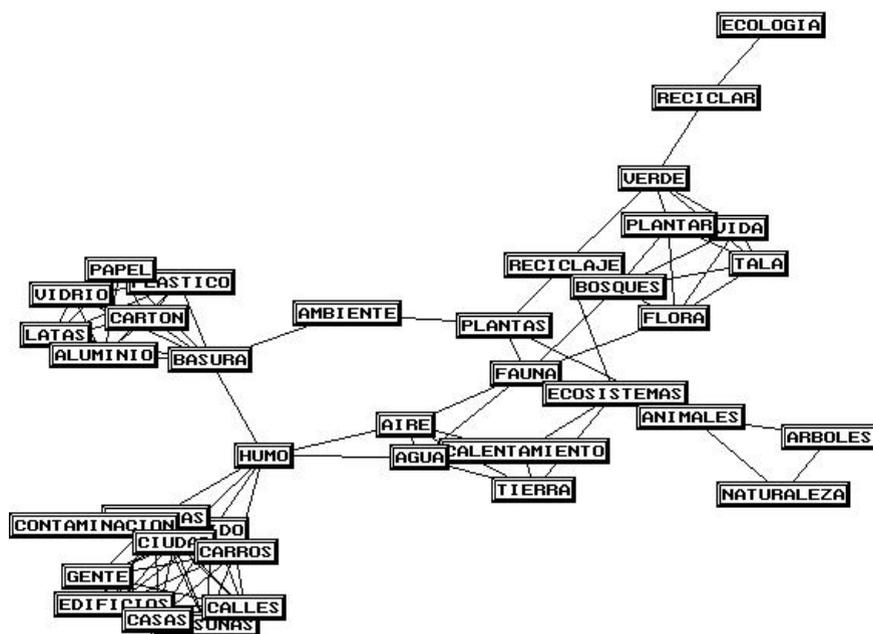


Figura 5. Organización conceptual de proximidad semántica Pathfinder para el grupo de participantes del estudio.

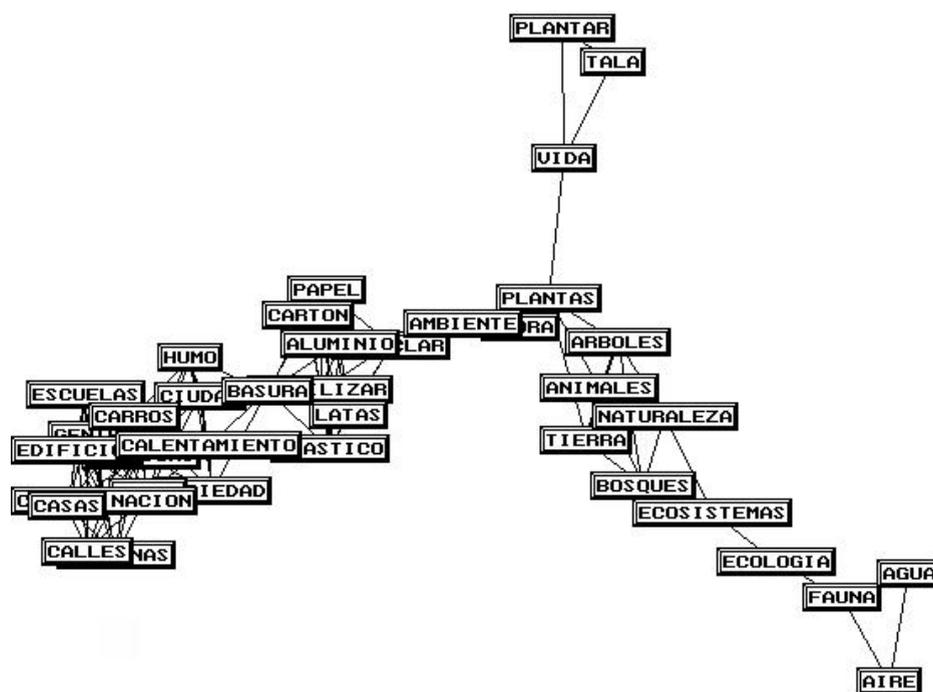


Figura 6. Organización conceptual de proximidad semántica Pathfinder resultante del rastreo de información.

Se puede observar de ambas figuras de proximidad semántica que existen diferencias relevantes en términos de un significado previo a la búsqueda de información. Esto es la representación que emerge durante el rastreo de información es homomórfica a la red semántica del usuario, sugiriendo que aunque ambas redes tienen contenido y conectividad similares son diferentes en términos de proximidad semántica aunque en términos de significado son funcionalmente equivalentes. Por otra parte, índices de similitud de la organización conceptual provenientes del sistema KNOT confirman la poca similitud entre ambas redes.

A pesar de las diferencias de organización conceptual entre ambas redes es posible entender la organización conceptual para cada red como un autómata con estados de transición por ejemplo en el caso de la organización conceptual de los usuarios el hecho que un individuo active información sobre ciudad requiere desde la perspectiva de los grupos SAMS iniciar en un estado q_0 , esto es, en el concepto Agua indicando que este fue uno de los conceptos más relevantes en términos de significado para indagar sobre la temática del medio ambiente de ciudad. Por razones de espacio no se incluye todo el diagrama de transición, pero la Figura 7 señala de forma gráfica que para ir del estado inicial conceptual de AGUA al estado

conceptual final aceptable RUIDO solo existe una secuencia conceptual posible. Para mayor ejemplificación, nótese la Figura 8 que incluye la tabla de transición así como el estado de transición para ir del definidor de ciudad “CARROS” hacia ESCUELAS.

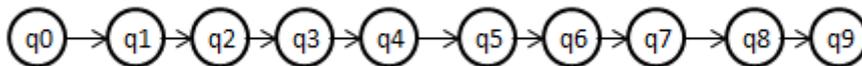


Figura 7. Diagrama de estados de transición conceptual redes semántica partiendo del concepto de carros en la definición conceptual de ciudad.

q ₀	Carros
q ₁	Edificios
q ₂	Ciudad
q ₃	Contaminación
q ₄	Calles
q ₅	Casas
q ₆	Personas
q ₇	Gente
q ₈	Fabricas
q ₉	Ruido
q ₁₀	Escuelas



Figura 8. Diagrama y tabla de estados de transición conceptual rastreo partiendo del concepto de carros en la definición conceptual de ciudad.

Con respecto al diagrama de transición que surge del rastreo de información se observó que no es exactamente la misma secuencia lineal que se da en el caso de organización conceptual del usuario pero es relevante notar que estados aceptables de organización conceptual final pueden llegar por rutas alternativas en el diagrama de transición. Por ejemplo, en el mismo caso de agua hacia ruido la secuencia de estados conceptuales demanda considerar los conceptos (edificios, personas, carros, casas, contaminación, gente, calles, ruido). Esto es, ambos sistemas parecen estar usando el mismo conjunto de símbolos Σ pero con diferente regla de transición. De manera formal:

Redes: $[q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9]$ \cap Rastreo $[q_1, q_6, q_0, q_3, q_5, q_7, q_4, q_9, q_8, q_{10}]$

Es decir que

Redes \cap Rastreo $[q_0, q_1, q_3, q_4, q_5, q_6, q_7, q_8, q_9]$

Así es que 9 de cada 10 conceptos existentes en la organización de redes del usuario se encuentra también en el rastreo de información. Lo cual prácticamente señala que ambas organizaciones conceptuales utilizaron casi los mismos estados de transición pero en diferente orden.

Discusión y Conclusión

En el pasado, el estudio de cómo se relaciona la organización conceptual del usuario a la organización conceptual que se rastrea era inexacta porque no había forma de determinar la organización conceptual “real” del usuario. Como se ilustra a partir de los resultados es posible incluir dicha organización conceptual desde una perspectiva de las redes semánticas siempre y cuando se utilicen técnicas que obtenga la información del individuo y no deforma idiosincráticas. Desde esta perspectiva, los resultados señalan que la información del usuario es homomórfica a la organización conceptual que se da en el rastreo de información. De relevancia fue el hecho de encontrara que son los mismos contenidos conceptuales que parecen subyacer a ambas organizaciones de información. Esto apoya la noción de un significado único como el elemento central, ya que el significado se construye por los contenidos conceptuales sin que este se vea afectado por la forma y entrada de activación de una red conceptual (Lindsay & Norman, 1977; Murphy, 2002). Ahora bien, desde una perspectiva de teoría de autómatas los posibles recorridos en las tablas de transferencia que se puedan dar dado el rastreo de información sugieren un sistema no determinista compuesto de dos subsistema autómatas (Rastreo vs Red de usuario) que implica estados conceptuales simultáneos de organización conceptual en un mismo tiempo pero que no necesitan ser los mismos. De esto resalta entonces que cada subsistema posee su propia regla de transición.

Lo que es posible determinar es que la regla de transición que impacta en las redes del usuario debe estar influenciada por el hecho de que la riqueza conceptual que el usuario posee es mucho mayor que la posee un sitio web. No tanto por la temática que se está rastreando sino por el hecho de que el usuario posee información conceptual que el sitio web no posee. De hecho esto se refleja en la riqueza conceptual generada por la técnica de redes semánticas en donde la organización conceptual de redes del usuario tiene valores J más altas que las de rastreo de información. Si esto funge como un indicador dentro de la regla de transición

todavía no es determinado e invita a plantear diseños experimentales al respecto. Finalmente, obsérvese de los resultados que la densidad semántica que subyace a ambas organizaciones conceptuales es prácticamente la misma. Aquí hay dos opciones la primera es creer que esta es un situación azarosa o que en la conmutabilidad de la ecuación 1 la organización conceptual de redes y rastreo se someten a una relación sistemática de significado.

Dado lo anterior es posible visualizar una nueva forma metodológica empírica para relacionar organización conceptual de un usuario al tipo de organización conceptual que se trata de acceder durante el rastreo de información. Aunque esta propuesta es inicial, nótese que se sugieren técnicas de análisis que permiten tanto el análisis cuantitativo como cualitativo y que son usadas en contextos experimentales de la cognición humana.

Referencias

- Chen, C., & Rada, R. (1996). Interacting with hypertext: A meta analysis of experimental studies *Human-Computer Interaction, 11*, 125-126.
- Davies, J., Fensel, D. & Van Harmelen, F. (2003). Towards the semantic web: Ontology –driven knowledge management.
- De la Garza, G. A., Sánchez, M.M.P. & López, R. E. O (2010). SemNet. <http://www.cognilab.org/RECURSOS.html>.
- Figuroa, J. G., Gonzales, G. E. & Solis, V. M. (1981). Una aproximación al estudio de las redes semánticas. *Revista Latinoamericana de Psicología, 13*, 447-458.
- Hayes, J., & Allinson, C.W. (1998). Cognitive style and the theory and practice of individual and collective learning in organizations *Human Relations, 51*, 847-871.
- Hopcroft, J.E., Motwani, R. & Ullman, J.D. (2001). Introduction to autómatas theory, languages, and computation. New York: Addison-Wesley.
- Hlland, J.H., Holyoak, K.J., Nisbett, R.E. & Thagard, P.R.(1989). Induction: Processes of inference, learning and discovery. Cambridge Massachusetts: MIT Press.
- Jonassen, D.H. (2000). Toward a design theory of problem solving. *Educational Technology: Research & Development, 48* (4), 63-85.
- Jonassen, D.H. (2002). Engaging and supporting problem solving in online learning. *Quarterly Review of Distance Education, 3* (1), 1-13.
- Kaluzniacky, E. (2004). *Managing Psychological Factors in Information Systems Work: An Orientation to Emotional Intelligence* London: Information Science Publishing.
- Konar, A. (2000). *Artificial Intelligence and Soft Computing Behavioral and Cognitive Modeling of the Human Brain* London: CRC Press LLC.
- Lindsay, P. H., & Norman, D. A. (1977). *Human Information Processing: An introduction to Psychology*. New York: Academic Press
- López, R.E.O. (2002). *El enfoque cognitivo de la memoria humana: Técnicas de investigación*. México, D.F.: Trillas
- López, R. E. O. & Theios, J. (1992) Semantic Analyzer of Schemata Organization (SASO). *Behavior Research Methods, Instruments, & Computers, 24* (2), 277-285.

- Murphy, G.L. (2002). *The big book of concepts*. Cambridge, Massachussets: MIT Press.
- Palmquist R.A. & Kyung-Sun K. (2000). Cognitive Style and On-Line Database Search Experience as Predictors of Web Search Performance. *Journal Of The American Society For Information Science.*, 51(6), 558-566.
- Pirolli, P. (2007). *Information Foraging Theory. Adaptive Interaction with Information. Series in Human – Technology Interaction*. New York: Oxford University Press.
- Rogers, T.T., & McClelland, J.J. (2004) *Semantic cognition: A parallel distributed approach*. Cambridge, Massachussets: MIT Press.
- Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G.E.(1986). Schemata and sequential thought processes. In: McClelland, J.L., Rumelhart, D. E. & the PDP research group. *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models*. Massachussets: MIT Press.
- Schvaneveldt, R. W. (1990). Proximities, networks, and schemata. En: R. W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex.
- Sears, A. & Jacko, J.A. (2008). *The Human–Computer Interaction Handbook Fundamentals, Evolving Technologies, and Emerging Applications*. New York: Taylor & Francis Group, LLC.