

**INSTITUTO POLITÉCNICO NACIONAL
CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN**

No. 21 Serie: VERDE Fecha: Julio 99

**La Computación en la
Inter - Ciencia**

Adolfo Guzmán Arenas ¹

RESUMEN

En este trabajo se trata el papel de la computación como instructor cibernético que permite enseñar mezclando porciones de varias disciplinas (enseñar Inter - Ciencia), como medio para obtener información en bibliotecas de otros especialistas y como explotador y usuario (es decir, como científico cibernético que abarca toda la Ciencia).

Palabras Clave: Ciencia, científico cibernético, enciclopedia, inteligencia artificial, inter - ciencia, instructor cibernético, polilibro.

¹Centro de Investigación en Computación del I.P.N.
aguzman@pollux.cenac.ipn.mx

ADVERTENCIA

“Este reporte contiene información desarrollada por el Centro de Investigación en Computación del Instituto Politécnico Nacional a partir de datos y documentos con derechos de propiedad y por lo tanto su uso queda restringido a las aplicaciones que explícitamente se convenga.

La aplicación no convenida exime al Centro de su responsabilidad técnica y da lugar a las consecuencias legales que para tal efecto se determinen.

Información adicional sobre este reporte podrá obtenerse recurriendo a la Unidad de Publicaciones y Reportes Técnicos del centro de Investigación en Computación del I.P.N. Av. Juan de Dios Bátiz s/n, teléfono 729-60-00 ext. 46103, 56608 y 56610”.

La Computación en la Inter-Ciencia

Adolfo Guzmán Arenas
Centro de Investigación en Computación del I. P. N.
aguzman@pollux.cenac.ipn.mx

1. INTRODUCCIÓN Y ALCANCES

1.1 Qué es la Inter-Ciencia

La Inter-Ciencia puede verse como la comunicación e inter-relación efectiva entre varias ciencias –Por ejemplo, Ciencias Sociales, Ciencias Naturales, Humanidades. La Inter-Ciencia surge porque hay problemas que trascienden una especialidad. Estudia áreas que quedan a horcajadas entre dos o más divisiones actuales (disciplinas) de la Ciencia.

La Inter-Ciencia es un fenómeno artificial. No importa cómo se parta o divida la Ciencia, un ser humano solo podrá abarcar una pequeña parte de ella: lo que pueda aprender en digamos 20 años de estudio, de primaria a doctorado (edad de aprendizaje) más lo que pueda aprender en el resto de su vida (etapa productiva).

1.2 Alcance de este documento

El papel de la Computación (§3.1) como lenguaje entre mono-disciplinas ya ocurrió hace unos quince años. El papel de *instructor cibernético* que permite enseñar mezclando porciones de varias disciplinas, sin perder la ilusión o los requisitos (enseñar Inter-Ciencia, §3.2) está ocurriendo ahora. La labor (§3.3) de obtener información en bibliotecas de *otros especialistas* también está ocurriendo ahora. El papel (§3.4) de la Computación como explotador y usuario (es decir, como *científico cibernético* que abarca toda la Ciencia, no solo algunas mono-disciplinas) ocurrirá en unos diez años. Este documento habla sobre todo de los papeles 3.2, 3.3 y un poco del 3.4.

2. QUÉ ES LA CIENCIA

Podemos intentar un par de definiciones de diccionario:

- A. Los conocimientos, métodos, conceptos y artefactos producto del estudio de la naturaleza, del universo y de modelos de ellos.
- B. El reflejo o concepción de la realidad observada y de sus abstracciones.

La ciencia es una, un todo, puesto que “el universo” o “la realidad” no están inherentemente o intrínsecamente divididos.

2.1 La Ciencia es una

Empero, las áreas, disciplinas, especialidades, son productos artificiales. Ejemplos:

- La Reología – estudio de fluidos con viscosidad variable. Líquidos no newtonianos, visco-elásticos.
- La Proctología – estudio de la parte final del tubo digestivo humano.

2.1.1 ¿Por qué dividimos la Ciencia?

Las divisiones son motivadas por consideraciones prácticas (de utilidad: se agrupan los conocimientos que generalmente se utilizan juntos, y así se forma la licenciatura de, digamos, Ingeniero Metalúrgico); de autoconsumo: un ingeniero metalúrgico solo puede enseñar lo que sabe, y si le ha sido útil, tenderá a formar otros como él; y de extensión: ¿cuánto puede aprender una persona en 16 años de estudio?

2.2 El libro único de la Ciencia: La Enciclopedia

Tan es única (un solo todo) la Ciencia que Napoleón concibió que sus súbditos escribiesen *el libro de la ciencia*. Lo ideal es tener un solo “libro” de Ciencia, pensó el emperador. E inventó la *enciclopedia*: el libro que contendría todo el saber humano de esa época. Un gran esfuerzo.

Ahora, una enciclopedia tiene 30 tomos y no cubre todo. Empero, Napoleón estaba en lo cierto. ¿Para qué dividir la Ciencia?

2.3 Volvamos a unificar a la Ciencia

Borremos las fronteras entre disciplinas. Hallemos un lenguaje común.

Esta unificación tendría varias ventajas (sus desventajas se vieron en el §2.1.1), por ejemplo:

- Podríamos tener “una sola Ingeniería”, en vez de Civil, Mecánica, Eléctrica, Aeronáutica, ... Se podrían hacer “menús a la carta” (útiles) de cursos, y así diseñar licenciaturas *ad hoc*, bajo demanda o según los intereses de un estudiante, en vez de que todos los alumnos de Geología lleven *el mismo* plan de estudios.
- Dado un perfil de interés¹ personal, me gustaría que “alguien” (EVA [12-14], §3.2, o mi consejero de estudios) me guiara sobre qué aprender, qué leer, con quiénes platicar, qué simulaciones practicar, qué ejercicios de qué libros realizar, qué conferencias por televisión escuchar ...

2.4 Pasos para unificar la Ciencia

Propongo los siguientes pasos para unificar la Ciencia, y al mismo tiempo permitir que la computadora nos auxilie en esta reunión.

- Hacer una taxonomía del conocimiento (§2.4.1);
- Colocar en ella a documentos y personas (§2.4.2);
- Utilizar la computadora como herramienta
 - ◊ para que intercambien conocimientos y colaboren dos especialistas de disciplinas distintas (§3.1);

¹ ... y mis conocimientos iniciales

- ◊ para que un especialista pueda utilizar (leer, digerir) los artículos que yacen en la biblioteca de otro especialista (§3.3);
- Utilizar la computadora para que enseñe (provea de conocimientos), a partir de un conjunto unificado de conocimientos (§3.2);
- Utilizarla para que procese los conocimientos y llegue a nuevas conclusiones (§3.4).
La unificación resultante es sorprendente (§3.5).

2.4.1 Hacer una taxonomía del Conocimiento: el árbol del conocimiento

La Ciencia se divide en Física, Química, ... La Física en Mecánica, Electricidad, ... La Mecánica en Estática, Cinemática, Dinámica, Resistencia de Materiales, ...

El conocimiento (conceptos, materias, áreas) forma un árbol –un subárbol es el conocimiento común, definido como “el conocimiento de un especialista después de haberle quitado los conocimientos de su especialidad” o “el conocimiento de un niño de 9 años” (u otra persona que no posee especialidad) [9].

El espacio del conocimiento está definido por varios ejes, cada eje es la linearización de (un pedazo) del árbol, normalmente una “rama” reconocida del saber. Por ejemplo, podemos dividir la ciencia de la computación en Sistemas Operativos, Lenguajes de Programación, Bases de Datos, ... (ocho áreas según [1]). Entonces, cada área puede ser representada por un eje. El espacio de computación tendría ocho ejes, ocho dimensiones. Para linearizar un árbol lo recorreremos a profundidad primero («depth first»).

2.4.1.1 Taxonomías existentes

- Existen árboles del conocimiento para algunas disciplinas, por ejemplo:
- A. Taxonomía de la revista *Computer Review*, del A. C. M.
 - B. El libro “Taxonomy of Computer Science”. AFIPS Taxonomy Committee. AFIPS Press. 1980.
 - C. El modelo curricular de la ANIEI, [1] para enseñanza de computación en México, en licenciatura.
 - D. El árbol del sentido común [9].
 - E. La discretización del espacio de conocimientos [12], de los conceptos del posgrado (nivel Maestría) en una especialidad de la computación.

2.4.2 Colocar los conocimientos y poseedores de conocimientos (personas, libros, videos, ...) en el Espacio del Conocimiento

Es posible colocar cualquier conocimiento, libro, pregunta, artículo, estudiante, investigador, sobre el espacio del conocimiento. El conocimiento de Juan Pérez (medido por la computadora mediante un examen –de admisión a la Maestría del C. I. C., digamos–) es un punto o vector en este espacio, con ocho componentes, si seguimos a [1]. Un artículo o un capítulo de un libro es una trayectoria pequeña en este espacio. De Q_0 a Q_1 . Una pregunta (y su respuesta) es una trayectoria. Lo mismo una tarea, un examen (y su respuesta).

Un curso es una trayectoria un poco más larga en este espacio, que incrementa los conocimientos del que lo toma en ciertas cantidades sobre ciertos ejes. También son trayectorias un video, una grabación, *esta* presentación en PowerPoint, el diario (periódico) de hoy. Son recorridos que aumentan el conocimiento del que los recorre o viaja.

Para colocar estos recorridos en el espacio, se indexa o cataloga su posición inicial y su posición final.

2.5 La educación bajo la ciencia unificada

Muchos conceptos tradicionales se verán modificados cuando, con el auxilio de la computadora, se puedan transmitir materias y enseñanzas tomadas de un acervo global de conocimientos (es decir, de la Ciencia Unificada) y no solo del acervo de materias o material de enseñanza de la Licenciatura de Ingeniero Químico, digamos. Presento algunos cuantos.

2.5.1 Carreras y planes de estudio globales

Lo común es que todos los estudiantes de la carrera “Ingeniero en Comunicaciones y Electrónica” estudien lo mismo. Es decir, las carreras convencionales son trayectorias pre-definidas, por comités de profesores, sobre este espacio. Empero, se pueden tener carreras individualizadas, aunque hay que tener cuidado que sean útiles.

2.5.2 Carreras y planes de estudio individuales

El auxilio de las computadoras permitirá individualizar los estudios de cada quien, y armarle y luego transmitirle los conocimientos (artículos, simulaciones, tareas, ...) que él requiere. Armarle un plan de estudio individualizado (y no un plan de estudio “promedio”, como el del §2.5.1).

Para esto, es menester hacer (ya hay mucho hecho, en inglés) el material educativo necesario. Son los *polilibros* [14], artículos, noticias, comunicados de prensa, ..., que ahora se indexarán (§2.4.2) y colocarán en el espacio (o árbol) del conocimiento.

2.5.2.1 Administración automática del progreso del estudiante

La computadora podrá tomar nota del progreso (viaje) del estudiante por este espacio de conocimientos. Es decir, la máquina conoce qué conocimientos posee cada estudiante, y lo ubica en el espacio.

2.5.3 Conceptos antiguos vistos bajo nueva luz

Preguntas. Una pregunta (y su respuesta) puede indexarse como una pequeña trayectoria en este espacio

Tutoriales. Un tutorial inteligente o interactivo es un material educativo + simulador que se representa como una trayectoria en este espacio.

Asesor. Un asesor mío es una persona (punto en el espacio) que está “corriente arriba” en mi línea de aprendizaje (trayectoria).

Colega. Un colega es un punto en el espacio que está cerca² del mío.

Competencia Una “competencia” es un punto en el espacio. Ejemplo: Soldador “C”, curtidor de pieles “A”.

Habilidad de aprendizaje. La habilidad de aprendizaje de un estudiante se mide por la velocidad de viaje sobre la trayectoria de conocimientos. Se expresa en conceptos/mes. Con-

² La distancia entre dos puntos es la raíz cuadrada de la suma de los cuadrados de las distancias (diferencias) sobre cada eje. En computación, según [1], son ocho ejes.

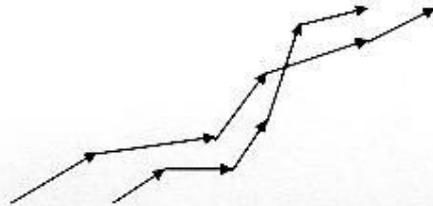
forme haya muchos viajeros (controlados u observados por la computadora) en el espacio, es posible tener datos de velocidades promedio de aprendizaje de tal o cual tema.

Empeño o dedicación. El “empeño” o dedicación de un estudiante se mide por el número de horas dedicadas a tratar de viajar sobre su trayectoria de aprendizaje. Se mide en horas dedicadas/horas totales, o sea, en un porcentaje de su tiempo (es decir, de su vida).

Eficiencia. Número de conceptos adquiridos/horas dedicadas a aprenderlos.

Olvido. Retroceso que un estudiante tiene sobre los conocimientos de un cierto tema.

Afinidad de trayectorias. La afinidad de dos trayectorias de aprendizaje (planes de estudio) es la distancia promedio entre las líneas, dividida por la longitud de ellas.



Entender un concepto. Es asimilarlo (colocarlo) sobre el árbol de conocimientos del lector.

3. EL PAPEL DE LA COMPUTACIÓN EN LA INTER-CIENCIA

¿Qué papeles juega la computación en la Inter-Ciencia, la interdisciplina, la transdisciplina y la multidisciplina? Cuatro distintos, que ocurren conforme aumenta la complejidad del software que interacciona con los especialistas de las mono-disciplinas:

- A. Como lenguaje común, para comunicar a dos especialistas mono-disciplinarios.
- B. Como *instructor cibernético* capaz de proveer conocimiento *bajo demanda* y a la medida, extraído de un banco de información unificado (una enciclopedia del saber humano, que no necesariamente estará centralizada en un solo lugar).
- C. Como administrador del conocimiento de (B).
- D. Como explotador y utilizador del conocimiento de (B).

3.1 Como lenguaje común o unificador

La computación (programas y algoritmos expresados en un lenguaje de programación, interfaces de aplicación que permiten inter-conectar paquetes, y otros artefactos) puede servir como lenguaje de comunicación entre mono-disciplinas. Así como antes se usaba el latín, y ahora se usa el inglés (u otro lenguaje natural) y las matemáticas, la computación sirve de puente para comunicar mono-disciplinas distintas. Este uso de la Computación ocurrió ya hace unos quince años, y continúa. Detalles en §4.1.

3.2 Como instructor capaz de proveer conocimiento (Ciencia) unificado: *el instructor cibernético*

Este es el proyecto EVA [12-14]. Para que la computadora pueda saber qué conocimiento administrar o proveer a cierto estudiante, es necesario:

- A. Saber lo que el estudiante *sabe* actualmente. Esto se logra haciéndole un examen (por computadora) que comprenda varios temas, el que arroja un vector con varios componentes (escalares), uno por cada tema preguntado. Es conveniente *discretizar y normalizar* el espacio de conocimientos,
- A.1. para poder indexar (con llaves exactas³) los conocimientos de cada módulo de la base del punto (B), y para que un módulo de un polilibro pueda embonar con algún otro módulo (si deseamos manufacturar polilibros personalizados);
 - A.1. a fin de que las mediciones en cada eje tengan la misma *granularidad de conceptos* (ver xxx + L. Zadeh), es decir, que una unidad de conocimiento en cualquier eje valga lo mismo. Ya habíamos visto que el conocimiento se mide en número de conceptos, de manera que en cada eje una unidad podría ser el número de conceptos que se aprenden en un día.⁴ Normalizar significa usar esta unidad para la discretización de cada eje.
- B. Lo que el estudiante *desea* saber, su meta final, a dónde quiere llegar. Lo mejor es mostrarle una gráfica del espacio de conocimientos, y que muestre con el ratón a qué punto desea llegar. Con esto, la computadora puede calcular el *tiempo estimado* (para un estudiante con IQ de 100, que le dedica 8 horas diarias) para llegar de su estado inicial a su estado final. También se pueden señalar en el espacio de conocimientos dónde se encuentran personas famosas o conocidas (Marvin Minsky, Emiliano Zapata). También se pueden tener en él “puntos promedio”, por ejemplo, “el promedio de conocimientos de los aspirantes a la Maestría del C. I. C., provenientes de UPIICSA”, y obtenido porque se ha aplicado el examen de admisión a 69 egresados de UPIICSA. O se puede mostrar un punto que represente a Jesús Olivares, o a algún otro aspirante que haya presentado el examen de selección.
- C. Tener acceso a (tener almacenada) una gran cantidad de información (cursos o capítulos de ellos) que el estudiante puede digerir o aprender para aumentar su estado de conocimientos. Junto con cada capítulo o módulo, es menester que la computadora conozca también *el efecto del módulo sobre el espacio de conocimientos*. Cada módulo origina que quien lo lee (y entienda) haga un pequeño viaje, aumentando, digamos, sus conocimientos de Sistemas Operativos.
- Con estas tres premisas, es posible que la computadora escoja de entre toda su *enciclopedia unificada*⁵ de conocimientos (del inciso C) aquél (minúsculo) subconjunto que Juan Pérez requiere, le arme el material (utilizando la relación de precedencia dada por cada eje) y se lo envíe (en un archivo llamado polilibro personalizado) por la red, digamos. Otros servicios proporcionados por *el instructor cibernético* (algunos de los cuales estarán presentes en EVA) son:
- D. Hallarle a un estudiante “compañeros de estudio” (personas que viven cerca y están aprendiendo lo mismo que él), “asesores” (personas que viven cerca y que ya saben lo que él ignora).
- E. Administrar (recoger) sus preguntas y proporcionarle respuestas. Esto puede lograrse:

³ Se conoce asimismo la técnica para utilizar llaves aproximadas o inexactas.

⁴ Si tomamos un eje, por ejemplo “Sistemas Operativos”, entonces los conceptos de esta área se colocan a lo largo de este eje según una relación de precedencia, denotada por “<”. Si A y B yacen en el eje, $A < B$ significa que el concepto A se aprende primero que el concepto B, y que para aprender B se requiere saber A. En principio, cada eje es ortogonal a los otros, de manera que mi capacidad de aprender un concepto de “Sistemas Operativos” solo depende de otros conceptos de “Sistemas Operativos”, y no de conceptos de otro eje, por ejemplo “Matemáticas”.

⁵ Nótese que estamos volviendo al concepto napoleónico, original, de “enciclopedia”.

- E.1. Síncronamente, cuando hay un instructor de carne y hueso respondiendo las preguntas por, digamos, correo electrónico o un foro de discusión.
- E.2. Asíncronamente, cuando los estudiantes envían mensajes al profesor, y este responde después por el mismo medio.
- E.3. En línea, mediante un sistema experto que es capaz de entender y responder a la pregunta. Por ejemplo, existen actualmente programas que ofrecen explicaciones [11], y esta tecnología parece ser la que se requiere.
- E.4. En línea, mediante un acceso con llaves inexactas [6] a bancos de preguntas y respuestas hechas previamente. La computadora analizará la pregunta del estudiante, la descifrará y verá con qué pregunta en la base de datos casa. Y dará la respuesta asociada. Hay programas que casan descripciones de cambios a software ya existente, con descripciones de lo que el código fuente de tales programas significa o contiene [8, 15]. Esta tecnología parece ser la que se requiere. También conviene notar que el problema es más sencillo de lo que parece, pues la computadora *sabe* qué sabe el estudiante, y qué está estudiando en ese momento, de modo que el careo de su pregunta a otra ya almacenada está restringido a un área delimitada del conocimiento.
- F. Proveer de estadísticas de evaluación: qué tan bien se aprende un cierto tema o módulo (“autómatas de Moore”) mediante el *instructor cibernético*, comparado con aprenderlo usando métodos tradicionales; cómo va Juan Pérez, con qué velocidad promedio se aprende el tema “árboles k-d”, ...
- G. Proveer *educación a lo largo de la vida* al estudiante, haciéndole llegar temas que le interesen o sirvan, conforme vayan apareciendo en la enciclopedia (o en la red).
El instructor cibernético funciona como un bibliotecario raro que proporciona información de manera autónoma, según un plan. Este bibliotecario se desempeña como instructor, tratando de llevar a cada usuario de su *estado inicial* de conocimientos a un *estado final* deseado o declarado. Puede medir progreso sobre su trayectoria de aprendizaje.

3.3 Como administrador del conocimiento: almacenamiento y recuperación inteligente de artículos

Como bibliotecario inteligente, permitiendo que cada especialista

- a) guarde sus documentos de trabajo e indexe los artículos que ha leído, en forma electrónica [“guarda este documento en podadoras de pasto; en Austin, Texas; en ‘The Austin Stateman’; en 28 de octubre de 1997”];
- b) los recupere, pese al olvido gradual o deterioro de la información en su mente conforme transcurren los meses [“¿dónde quedó un artículo sobre herramientas de jardín que leí en un periódico de EE. UU. a finales de 1997 o principios de 1998?”];
- c) busque y encuentre en la biblioteca de *otro* especialista temas que le interesan. Es decir, pone a disposición de un especialista las bibliotecas de *todos* los especialistas, no únicamente la suya (como en b).

En este sentido, la Computación funciona como un bibliotecario universal, que sabe lo que cada documento contiene (conoce los temas de los que el documento habla).

Podemos pensar que un administrador del conocimiento es un bibliotecario, un almacenador y recuperador inteligente de conocimientos, de artículos, de libros. Aunque este administrador no entienda mucho de lo que guarda y administra, entiende lo suficiente para

saber los *temas principales* que contiene cada artículo. Esto será pronto posible, debido a CLASITEX [2, 3].

3.3.1 Administrador de mi propia biblioteca, *resistente a mi olvido*

El problema a resolver: dónde guardar lo que leo, de manera que seis meses después, lo encuentre. Esto no es trivial debido a que, conforme pasa el tiempo, los humanos caen en el olvido selectivo, generalizando a medida que van perdiendo bits de información. "Se me olvida parcialmente en qué coordenadas o clasificación lo guardé." Por ejemplo:

- Lo guardo en Austin y lo quiero recuperar en Texas.
- Lo guardo en podadoras de pasto y lo quiero recuperar en herramientas de jardín.

Ya sabemos que la computadora no olvida lo que guardó. Se trata de hacer un administrador que sea resistente a *mi* olvido. Este administrador puede construirse y funcionar si se siguen los siguientes pasos:

A. Indexación de artículos. Los artículos que deseo guardar en mi biblioteca "resistente a mi olvido" ya existen en archivo. Yo (o el usuario) los puedo indexar. Los leo y decido (veo) qué temas abordan. O puedo usar a Clasitex para que él los lea y determine qué temas abordan. De esta manera se pueden clasificar.

B. Preparación de ejes. Yo preparo ejes (pueden ser subdirectorios en Unix) en los cuales guardar los artículos que he leído u hojeado.⁶ (Nótese que esta es una versión simplificada de la *discretización del espacio de conocimientos*, y puedo salir adelante con ella porque, con seguridad, el espacio de conocimientos de *una persona* es minúsculo y no merece mucha atención). Esto es equivalente a: yo etiqueto los anaqueles y subdivisiones de mi biblioteca o bodega de datos.⁷ [Nótese la similitud con las "bodegas de datos" («data warehouses», [7]) del área de Sistemas de Información].

C. Almacenamiento de cada artículo.

C.1. Yo decido dónde guardar el artículo que acabo de leer.

C.2. O, CLASITEX decide (o cuando menos, sugiere) dónde guardarlo.

El procedimiento es similar en ambos casos. Guardamos el artículo (por C.1. ó C.2.):

fecha: 4 de octubre de 1997

tema: la Inter-Ciencia

tema: la computación como lenguaje unificador

expositor: Adolfo Guzmán

lugar: Auditorio Ignacio Chávez, UNAM

motivo: homenaje a Marco Murray-Lasso

D. Recuperación. Seis meses después, deseo recobrar el artículo. Pero tengo de él una idea más vaga:

fecha: fines de 1997

tema: inter-disciplina

expositor: un oaxaqueño

lugar: una universidad en México, D. F., creo que era la UNAM o el I. P. N.

⁶ Cada usuario tiene una forma definida para organizar (guardar, clasificar) los artículos que ha leído, debido a sus inclinaciones, intereses, deformaciones profesionales, etc. Esta organización le permitirá agrupar artículos similares (para él) en áreas cercanas. Cada biblioteca es particular, única. La organización que Juan Pérez usa para su biblioteca no me sirve a mí (me sirve quizá parcialmente, pero prefiero la mía) y viceversa.

⁷ En cambio, la clasificación de la Biblioteca del Congreso (EE. UU.) hace lo propio *para los conocimientos de todos*.

La búsqueda se hace (por el administrador) indexando los temas de los que me acuerdo, yendo al cubito donde se intersectan, y leyendo todos los artículos que ahí existen. También puedo usar CLASITEX sobre los artículos que están en el cubito. Idea: se puede usar la tecnología de llaves inexactas [6].

3.3.1.1 El problema de los predicados incorrectos

Ocurre a veces que al olvidar parcialmente (y generalizar) los temas de un artículo que leímos hace seis meses, nos equivocamos. Cometemos un error, y pensamos que el autor era veracruzano, aunque en realidad era oaxaqueño. Al buscar con el predicado incorrecto,

(fines de 1997) and (tema=inter-disciplina) and (veracruzano), (1)

o no encontramos lo que buscábamos, o no encontramos nada. Nos damos cuenta de que el predicado está incorrecto. Si estamos seguros de que el artículo fue almacenado, que “debe andar por ahí”, entonces podemos pedirle al buscador que *generalice el predicado*, quitándole cláusulas conjuntivas (o cambiándolas por disyuntivas) de derecha a izquierda. En vez del predicado (1) tendremos sucesivamente las búsquedas siguientes :

(fines de 1997) and (tema=inter-disciplina) (2)

(fines de 1997) and (veracruzano) (3)

(tema=inter-disciplina) and (veracruzano) (4)

(fines de 1997) (5)

(tema=inter-disciplina) (6)

(veracruzano) (7)

Esta generalización se basa en que típicamente las primeras cláusulas de un predicado conjuntivo son de las que estamos más seguros. Las más dudosas vienen al final.

3.3.2 Administrador a mi servicio, que usa la biblioteca *organizada de mi colega*

Si algún colega me permite usar su biblioteca (*organizada por él* de acuerdo con §3.3.1), puedo usarla para buscar artículos que me interesen. El problema a resolver es: buscar en su biblioteca artículos descritos con *mis* índices. Él los guardó e indexó a su modo. Yo quiero recuperar uno de ellos, pero lo describo a *mi modo*.

El procedimiento es el siguiente:

A. Utilizo *su* árbol (taxonomía) o configuración de *su biblioteca*.

B. Pero utilizo *mis* índices. Para esto, es necesario mapear mis índices a los suyos (a sus ejes). Se hace este mapeo usando el árbol del conocimiento común [9].⁸ Ver también [8, 15]. Hay tres casos posibles:

Caso 1: mi tema “podadoras de jardín” no lo encuentro tan detallado en *su* biblioteca. Él solo llega a “herramientas”. Está bien. Ahí busco artículos con *mi* tema (usando mis ojos, o CLASITEX). También leo algo de cada artículo, para ver si de hecho habla de podadoras de jardín, porque lo que estoy hurgando es el nodo “herramientas”.

Caso 2: mi tema “podadoras de jardín” está enormemente desarrollado en la biblioteca de mi colega. Él tiene bordeadoras, orilladoras, podadoras industriales, auto-podadoras, podadoras importadas, podadoras a gasolina, etc.

⁸ En ciertas disciplinas los índices ya son estándar, por ejemplo, los códigos asignados a discos de música, o los asignados a estrellas y cuerpos sidéreos, o las coordenadas (*longitud, latitud*) para lugares geográficos, o las fracciones arancelarias.

Eso significa que él tiene como un nodo (padre, abuelo, ...) el nodo "podadoras de jardín". De suerte que yo busco en ese nodo y sus descendientes. Y encuentro (con gran lujo de detalle) lo que yo busco.

Caso 3: el tema "podadoras de jardín" no está.

No busco. No lo tiene. No voy a encontrar algo.

Nótese la similitud con el Proyecto # 38 "Lenguaje para intercambiar información entre agentes" de [4], que mapea el árbol del agente oferente O al árbol del agente C, mediante el árbol común del área.

3.3.3 Administrador a mi servicio, que usa la biblioteca *no organizada* de mi colega

El problema es similar al punto anterior §3.3.2, pero la búsqueda es exhaustiva. Cuando menos, si los artículos están clasificados sobre el árbol de conocimientos, hay un lenguaje común, y podemos utilizar índices (hacer búsqueda sobre los índices). En vez de leer todo el artículo en español.

3.3.1 Una posible mejora

Se puede obtener quizá un algoritmo un poco más sistemático. La idea general es usar un espacio de conocimientos común, discretizado (ver mis comentarios a esta discretización en §3.3.1.B). Sobre él, pueden existir especializaciones (a mí me interesan mucho las bases de datos, por lo que tengo subdivisiones para Oracle, Informix, Progress, etc.).

La computadora puede ayudarnos a explorar un "espacio algo extraño" utilizando *nuestros índices*. Empero, si el espacio es demasiado extraño, nuestros índices serán muy vagos o inapropiados, "no encajarán" con los del colega. Por ejemplo, yo puedo tener ideas raras, no ortodoxas, sobre la clasificación de la música, y por ende no podré usar efectivamente la biblioteca de musicólogos como Armando Zayas o Luis Herrera de la Fuente (QEPD).

3.4 Como explotador y utilizador del conocimiento

Como usuario y generalizador del conocimiento guardado en el §3.3, cuando la Computadora, además de poder acceder la información guardada en muy diversas bibliotecas privadas, o en la enciclopedia unificada,⁹ pueda *entender, utilizar y hacer deducciones* de los textos o documentos que almacena (y estudia). Es decir, como "científico universal". Hablaremos de esto en el §4.4.

3.5 ¿En qué sentido la Computación puede ser un medio *unificador* de varias disciplinas o ciencias?

Sería ideal si la Computación pudiese utilizarse para desvanecer las fronteras (artificiales) que separan una ciencia de otra. ¿En qué sentido esto puede lograrse?

A. En un sentido muy limitado, creo yo, si utilizamos a seres humanos como procesadores (explotadores, investigadores) de esta información. La razón es que para que una persona monodisciplinada pueda abarcar (explotar, usar) los conocimientos "cercaños" de otra disciplina, en realidad debe *entenderlos* (asimilarlos a su árbol de conceptos,

⁹ Pudiese ser que la biblioteca universal (unificada) esté constituida de la unión de bibliotecas particulares.

§2.5.3).¹⁰ Esta asimilación no recibe apoyo especial de la Computación, es una asimilación que la persona debe hacer “a mano” (“a mente”, sería una expresión más apropiada). La Computación ayudará poco a este proceso, excepto quizá a organizar los temas en una base de datos, y en otras tareas auxiliares.

Nótese que ya postulamos la utilidad (§4.1) como lenguaje común o *lingua franca* entre monoespecialistas humanos. Esto no unifica, sino que simplemente permite la discusión. De la misma manera que el español permite que las opiniones a favor y en contra del aborto puedan conocerse “por ambos bandos”, pero no unifica criterios o voluntades.

- B. En un sentido muy amplio, creo yo, si utilizamos a las computadoras como procesadores (explotadores, investigadores) de esta información (§3.4), capaces de generar nuevos conocimientos.

3.6 Como traductor de una jerga especializada a otra

La computación puede traducir de una jerga (o dialecto) especializado a otro, *siempre y cuando se refieran al mismo campo o área de conocimientos* (es decir, que los nodos de los conceptos pertenezcan a la misma región del árbol). Por ejemplo, SoftwarePro International ha desarrollado un programa [5] para poder acceder a y explotar bases de datos poco familiares, concretamente aquéllas conteniendo datos sobre estudiantes. Las bases de datos sobre estudiantes básicamente contienen el mismo tipo de información aunque pertenezcan a Conacyt, al Politécnico, o al C. I. C. Cambian los detalles, las denominaciones, los sinónimos específicos que se utilizan (aspirante, estudiante, estudiante activo). Mapeando una jerga especializada (implementación concreta, es decir las definiciones de tablas, campos y llaves de la base) a un “idioma puro” (que es el expresado en el árbol del conocimiento común), y de ahí a la otra jerga (otra implementación concreta), se obtiene la traducción. Ver también [8, 15]. Lo que no se puede es traducir de un grupo de conceptos a otro, como se aclara a continuación.

3.7 Traducción de un grupo de conceptos a otro

El uso de la Computación como traductor de conceptos de una disciplina (“ácido acético”) a conceptos de otra (“juicio civil”) no tiene sentido y no se persigue.

3.8 Facilitador de interacciones asíncronas y no presenciales entre diversos técnicos

El proyecto EVA del C. I. C. desarrolla parte de estas interacciones. Esta línea de investigación se cubre en la referencia [12, 14], y en breves espacios de este documento.

¹⁰ Ejemplo: no soy químico, pero deseo saber ¿qué son los nitritos? Un profesor (hombre o máquina) me dirá: son las sales del ácido nítrico que aún tienen una valencia sin reducir. Ahh... ¿qué es sal? ¿qué es valencia? ... El diálogo continúa hasta que me topo con conceptos que *ya entiendo*, es decir, que ya están colocados en mi árbol de conocimientos. El diálogo tuvo el efecto de colocar sobre mi árbol los conceptos aprendidos.

4. CONCLUSIONES Y RECOMENDACIONES

4.1 El lenguaje unificador

¿Cuál es el lenguaje común, el unificador? La unificación de la Ciencia se hará sobre el Espacio de Conocimientos. Por ende, el lenguaje unificador está compuesto de *datos* y *procedimientos*, algo así como un lenguaje de objetos. El lenguaje unificador consiste de:

A. Una retícula sobre (la discretización de) este Espacio;

B. Un traductor de "dialectos" o jergas científicas.

B.1. Traductor de una jerga particular a otra al mismo nivel. Cf. §3.6. Ejemplo: Acceso a bases de datos poco familiares [5].

B. 2. Traductor generalizador. Así como hay traductores de BASIC a FORTRAN, también los hay de conceptos especializados a conceptos más generales. Ejemplo: CLASITEX [2, 3].

Otra posibilidad que puede ocurrir si los trabajos de investigación de textos escritos en español avanzan lo suficiente (por ejemplo, los que se efectúan en el Laboratorio de Lenguaje Natural y Procesamiento de Texto del C. I. C.) es usar una superversión de Clasitex para poder usar como lenguaje unificador el español (u otro lenguaje natural). Sería:

C. Un lenguaje natural —el español, digamos— salpicado de notaciones especiales (Matemáticas, Química, ...). Aquí el superClasitex sería el traductor entre las diferentes notaciones especiales, jugando el papel de B.1 y B.2.

4.2 ¿Cómo podrían las personas usar ese lenguaje unificador?

Por un cierto tiempo el material educativo seguirá siendo escrito en lenguaje natural (o en la variante del §4.1.C, a lo sumo), ya sea en libros o en polilibros. El material educativo (libros, capítulos, artículos, nuevos conocimientos) seguirá siendo generado (por un tiempo) por las personas. Pero la computadora lo indexará, lo interrelacionará, y lo enviará a los consumidores finales. La computadora administrará el avance del conocimiento adquirido por tales consumidores finales. Porque la computadora sabe **qué cosas sabe cada quien**, dónde está trabajando, cuánto gana, ...

4.2.1 El futuro está aquí ya

Algunos de estos avances ya se están efectuando:

A. El Proyecto CYC en MCC [9], que buscaba fabricar y explotar el árbol del conocimiento común.

B. El Proyecto EVA del Centro de Investigación en Computación del I. P. N., que es una versión del *instructor cibernético* del §3.2.

C. Clasitex [2,3] y el método de acceso a bases de datos poco familiares [5], que representan engranitos que hacen un poco más factible, quizá, el lenguaje unificador.

D. Quizá la *enciclopedia universal unificada* simplemente sea, con la ayuda de §3.3, la unión de varias bibliotecas particulares.

4.3 Conclusiones

La máquina ya puede ser el *instructor cibernético* del §3.2, capaz de poner a disposición de una persona los conocimientos que ésta requiere. Entonces, concluyo que tiene sentido desarrollar proyectos como EVA que, como primer instructor cibernético, facilitará el aprendizaje a personas que no pueden desplazarse a los grandes centros de enseñanza a obtener educación presencial.

La máquina puede usarse ya para almacenar conocimientos a fin de recobrarlos aun contra nuestro olvido (§3.3.1). Falta desarrollar el software, está claro cómo proceder.

También puede usarse ya (§3.3.2) para utilizar los conocimientos almacenados en la biblioteca de colegas, y, más generalmente, en la red (§3.3.3). La tesis de doctorado de Jesús Olivares en el C. I. C. desarrolla esta línea de investigación.

Por lo anterior, concluimos que la máquina puede ser ya *la enciclopedia universal*, contenedora de todo el conocimiento. Tiene capacidad de indexarlo y accederlo, aunque no lo entiende muy bien.

4.4 Más allá de la Inter-Ciencia: la computadora como Científico Universal

Un siguiente logro importantísimo sería que la computadora *entendiera* lo que almacena, y no tan solo como ahora, que entiende *los temas* de lo que almacena (como CLASITEX). Para esto,

Primero entendería el conocimiento común (sentido común). Este tema lo abordó el Proyecto CYC. Es decir, la computadora tendría *sentido común*.

Luego, usando un *programa inteligente* al que llamaremos CYC++, podría entender textos en español. Por consiguiente, podría aprender más, mediante el acto sencillo de leer libros especializados (de Física, de Reología, Proctología...). Con esto se llegaría a un *conocedor universal*, unificado, que sabe (leyó y entiende) **todos** los conocimientos de todas las disciplinas (según los documentos que logre acceder). Nótese que en este momento la computadora tendría digerido todo el saber humano. Carece de las limitaciones humanas en cuanto a capacidad de aprendizaje.

Agreguemos al *conocedor universal* anterior:

- (a) métodos de inferencia, deducción, generalización (que se conocen a cierto detalle ya en Inteligencia Artificial), reificación, heurísticos;
- (b) acceso a datos [7], mediciones, observaciones; conectándolo también a dispositivos de lectura; dotándolo de agentes que buscan información caracterizada por filtros;
- (c) una agenda o meta a llegar, o un método para generalizar, generar nuevo conocimiento, nuevos teoremas [10].

Entonces habremos construido un sistema capaz de generar y descubrir nuevo conocimiento. Unificado. En cada una de las ciencias. Sin las limitantes de "solo puedo aprender durante 20 años". Este ente no solo sería el conocedor universal (sabría toda la Ciencia, la Ciencia Unificada), sino sería el *generador de nuevos conocimientos*. Naturalmente, sabría **todo** el saber humano. Y todo el saber electrónico.

Es decir, la computadora sería la Ciencia. Y el científico (universal).¹¹ El comienzo de la Ciencia Unificada.

4.5 Agradecimiento

Al Dr. Marco Murray Lasso, por haber celebrado el 4 de octubre de 1997, sus 40 años de vida profesional.

A la Dra. Guillermina Yankelevich, por haberme invitado a esa celebración, y pedirme que hablara precisamente del tema "La computación en la Inter-Ciencia."

Nota: Clasitex es propiedad de SoftwarePro International. Anasin y el programa de Acceso a base de datos poco familiares son propiedad de Idasa.

4.6 Referencias

1. ANIEI, Asociación Nacional de Instituciones de Educación en Informática, A. C. *Modelos curriculares nivel licenciatura, informática, computación*. Publicado por el INEGI, Aguascalientes, México, 1997.
2. Adolfo Guzman. Finding the main themes in a Spanish document. *Journal Expert Systems with Applications*, Vol. 14, No.1/2, Jan/Feb 1998, pages 139-148.
3. Adolfo Guzmán. Hallando los temas principales en un artículo en español. *Soluciones Avanzadas*. Vol. 5, núm. 45, pág. 58. I parte, 15 de Julio de 1997, II parte vol. 5, núm. 49, pág. 66, 15 de septiembre de 1997. También en: *Symposium Internacional de Computación*. Centro de Investigación en Computación. Instituto Politécnico Nacional. Noviembre 12-14, 1997. México, D.F., páginas 36-51.
4. Adolfo Guzmán. *Proyectos y Temas de Tesis en Computación*. Reporte técnico en preparación, Centro de Investigación en Computación, Instituto Politécnico Nacional, México, D. F., 1998.
5. Adolfo Guzmán. Acceso a bases de datos poco familiares. Reporte Final, IDASA, 1994. México, D. F.
6. Adolfo Guzmán. Búsqueda en grandes bases de datos con llaves inexactas. Reporte técnico en preparación, Centro de Investigación en Computación, Instituto Politécnico Nacional, México, D. F., 1998.
7. Adolfo Guzmán. Anasin, Analizador y Sintetizador Inteligente de Información. Informe técnico, manual del usuario. IDASA, 1994. México, D. F.
8. Jae K. Lee, Byoung Y. Lee, Wooju Kim. A knowledge based maintenance of large scale legacy system: Metasoft. *Proc. of the 4th World Congress on Expert Systems, Applications of advanced information technologies*, pp. 65-72. F. Cantú, R. Soto, J. Liebowitz, E. Sucar (eds). ITESM Mexico City Campus, March 16-20, 1988. México City.
9. Douglas B. Lenat, R. V. Guha. *Building large knowledge-based systems*. Addison Wesley. 1990.
10. Douglas Lenat and J. S. Brown. Why AM and Eurisko appear to work. *Artificial Intelligence* 23, 1984, pp. 269-294.
11. Andrew Lymer, Amelia Baldwin, Jan Scott. Explanation generation in accounting expert systems. *Proc. of the 4th World Congress on Expert Systems, Applications of ad-*

¹¹ Por cierto, podríamos preguntarle algunas cosas a este ente, aprender de él. Pero siempre con las limitaciones que actualmente tenemos los seres humanos.

- vanced information technologies, pp. 25-32. F. Cantú, R. Soto, J. Liebowitz, E. Sucar (eds). ITESM Mexico City Campus, March 16-20, 1988. México City.
12. Gustavo Núñez, Leonid Sheremetov, Jesús Martínez, Adolfo Guzmán, Álvaro Albornoz. The EVA Teleteaching Project - the concept and the first experience in the development of Virtual Learning Spaces. *15th IFIP World Computer Congress 'The Global Information Society on the way to the next Millenium', Vienna and Budapest, 31 Aug - Sept 4, 1988*, organized by the Austrian Computer Society and John von Neumann Computer Society on behalf of the International Federation for Information Processing.
 13. Gustavo Núñez, Leonid Sheremetov. Design of a collaborative learning environment, based on process engineering technology. *15th IFIP World Computer Congress 'The Global Information Society on the way to the next Millenium', Vienna and Budapest, 31 Aug - Sept 4, 1988*, organized by the Austrian Computer Society and John von Neumann Computer Society on behalf of the International Federation for Information Processing.
 14. Gustavo Núñez. [Eva] referencia al trabajo que salió en la revista Archipiélago.
 15. Jungsoon Yoo, Sung Yoo. Conceptual clustering on partitioned data: tree weaver. *Proc. of the 4th World Congress on Expert Systems, Applications of advanced information technologies*, pp. 529-536. F. Cantú, R. Soto, J. Liebowitz, E. Sucar (eds). ITESM Mexico City Campus, March 16-20, 1988. México City.