

Forschungs- und Entwicklungsprojekte

Informatik Forsch. Entw. (2004) 18: 201–206
Digital Object Identifier (DOI) 10.1007/s00450-004-0148-6

In dieser Rubrik erscheinen in unregelmäßiger Folge Kurzdarstellungen geplanter, laufender oder abgeschlossener Projekte. Die Darstellungen werden in der Regel von den Projektbeteiligten geliefert. Die Auswahl erfolgt durch die Herausgeber. Dabei wird die Bedeutung des Projekts für die Fortentwicklung der Informatik das Hauptkriterium sein. Bei geplanten und laufenden Projekten ist ein wichtiges Kriterium der Wunsch, Kontakte zu etablieren und die Zusammenarbeit zwischen verschiedenen Gruppen zu fördern. Bei abgeschlossenen Projekten geht es primär um die Vermittlung von Erfahrungen und Ergebnissen, die sich nicht für die Veröffentlichung in redaktionellen Beiträgen eignen.

Wissen bei Menschen und Maschinen – eine informatikbezogene Betrachtung

Albert Endres

Sindelfingen

Zusammenfassung. Neben Information ist Wissen ein Begriff, der in der Informatik eine große Rolle spielt. Beim Menschen ist Wissen ein Teil seiner Gedächtnisleistung. Erwerben und Nutzen von Wissen ist eng verwandt mit dem Begriff Denken. Das Ziel der Informatik ist es, Leistungen des Menschen im Bereich von Gedächtnis und Denken mit maschinellen Mitteln zu unterstützen. Deshalb ist eine klare begriffliche Abgrenzung zwischen dem, was Menschen tun, und dem, was Maschinen leisten oder zu leisten vermögen, für Informatikerinnen und Informatiker sehr wichtig. Maschinen sind Menschen heute bereits überlegen, wenn es um das Speichern und Wiederfinden von Information geht, oder bei sehr speziellen Aufgaben wie dem Schachspielen. Anders ist es bei Denkprozessen wie Generalisieren, Abstrahieren und dem analogen Schließen, sowie bei der Lösung alltäglicher und wenig strukturierter Probleme.

Schlüsselwörter: Information, Wissen, Semantik, Wissensverarbeitung

Abstract. Besides information, knowledge is a key term used in computing. For humans, knowledge is part of what we call intelligence. Its acquisition and use closely relates to thinking. It is the aim of computer science to support or emulate human achievements in the realm of memory and thinking. It is therefore important that students of computing understand the difference between what humans do and what machines can achieve in this respect. Machines are superior to humans when storing and retrieving information and in the case of very specialized tasks like playing chess. It is different for mental processes like generalizing, abstracting and analog reasoning, and for the solution of broad and ill-structured problems.

Keywords: Information, Knowledge, Semantics, Knowledge processing

CR Subject Classification: E.4, H.1, I.2

1 Einleitung

Die Informatik, gleich ob als Wissenschaft oder als Industrie gesehen, hat das Ziel, Produkte zu erzeugen, welche die kognitiven Fähigkeiten und Leistungen von Menschen unterstützen. Das ist analog zu dem, was andere technische Produkte wie Kräne, Fahrstühle, Autos, Schiffe und Flugzeuge tun. Diese verbessern oder erweitern die physischen Kräfte und Fähigkeiten des Menschen. Zu den kognitiven Fähigkeiten des Menschen gehören Wahrnehmung, Kommunikation, Gedächtnis und Denken. Ein Teil dessen, was wir im Gedächtnis aufbewahren, ist unser Wissen. Die Informatik benutzt den Begriff Wissen im Zusammenhang mit wissensbasierten Systemen oder beim Wissens-Management. Auch wenn dies nicht ausdrücklich dazu gesagt wird, enthalten alle Informatikprodukte eine erhebliche Menge an Wissen. Das gilt insbesondere für Software. In diesem Beitrag wird der sehr breite Begriff Wissen, wie er in verschiedenen Fachgebieten benutzt wird, eingeschränkt auf das, was die Informatik damit verbindet. Von Informatikerinnen und Informatikern wird zu Recht erwartet, dass sie in der Lage sind, diesen Unterschied zu erklären, auch Laien gegenüber.

2 Information und Wissen

Der Begriff Information wurde in einem früheren Beitrag [4] ausführlich behandelt. Deshalb sei hier nur kurz wiederholt: Bei *Information* handelt es sich um interpretierbare, d.h. mit einer Bedeutung verknüpfbare Nachrichten oder Signalfolgen, auch „semantisierte Signale“ genannt [6]. Für die Informatik ist es sinnvoll und ausreichend, Wissen als eine Untermenge von Information aufzufassen, die einem besonderen Zweck dient und besondere Gütekriterien erfüllt. Es ist Information, die auf ein ganz bestimmtes Ziel hin bewertet und selektiert wird. Für die Zwecke digitaler Bibliotheken hatten Fellner und ich [5] uns folgender Definition bedient: „*Wissen ist der Bestand an Modellen über Objekte und andere*

Tabelle 1. Wahre und falsche Aussagen

Wahrheitstyp	Wahre Aussagen	Falsche Aussagen
Umgangssprache	Schwager ist der Mann der Schwester	Enkel ist der Sohn des Bruders
Biologie	Wal ist ein Säugetier	Fische sind Säugetiere
Physik	Sonne ist ein Fixstern	Sonne bewegt sich um die Erde
Chemie	Atomzahl von Eisen ist 26	Wasser ist ein Element
Mathematik	Winkel im Halbkreis ist ein rechter	Winkelsumme im Dreieck ist 128°
Wirtschaft	Verschuldung ist zu hoch	Schulden sind gut

Sachverhalte, über den Individuen oder die Menschheit als Ganzes zu bestimmten Zeitpunkten verfügen, zu dem sie Zugang haben und der für wahr gehalten wird.“ Wissen im engeren (pädagogischen) Sinne sind Modelle, welche die Welt in und um uns erklären (Orientierungswissen) und unser Handeln vernünftig werden lassen (Handlungswissen). Für andere der Informatik weniger nahestehende Zwecke wie Hobbys, Sport und Esoterik oder bestimmte geschäftliche Unternehmungen (Schmuggel, Spionage) können ganz andere Aspekte relevant sein.

Die zentralen Begriffe dieser Definition sind Modell und Objekt. Beide spielen in der Informatik eine große Rolle. Sie sind hier etwas umfassender zu verstehen. Ein Modell ist eine (für wahr gehaltene) Aussage über Objekte. Diese kann man mittels mathematischer oder logischer Formeln, Programmier- oder Entwurfssprachen (z. B. UML), geometrischer Skizzen, natürlichsprachlicher Sätze, Gesten, Mienen oder dergleichen ausdrücken. Das Modell kann sich auf Eigenschaften von Objekten beziehen oder verschiedene Objekte in Beziehung zueinander bringen. Ähnlich wie in der Physik sind Modelle operationelle, d. h. besonders leicht handhabbare Darstellungen einer Theorie¹. Oft werden Modell und Theorie auch gleichgesetzt, so z. B. beim Atommodell von Bohr oder beim DNS-Modell von Watson und Crick. Objekte können Gegenstände in der Welt sein oder aber Konzepte im Denken darstellen. Sie haben eine Identität und eine Menge von Eigenschaften oder Merkmalen. Den jeweiligen Wert dieser Eigenschaften kann man auch als Zustände ansehen. Es gibt einfache Objekte, die sich nicht zerlegen lassen, und zusammengesetzte Objekte.

Mit dem Begriff der Wahrheit des Wissens wird die auf Aristoteles zurückgehende “epistemologische Skala“ angesprochen. Danach gibt es eine Stufenskala ausgehend von falschen Aussagen, plausiblen Aussagen, über Vermutungen und Meinungen bis hin zu wahren Aussagen. Ob eine Aussage objektiv wahr ist oder nur subjektiv für wahr gehalten wird, ist für unsere Zwecke sekundär. Wie in Tabelle 1 angedeutet, gibt es viele Formen von Wahrheit. Sie kann sich aus der umgangssprachlichen Definition eines Begriffes ergeben oder kann durch ein wissenschaftliches Spezialgebiet begründet sein. Eine viel beachtete Definition von (wissenschaftlicher) Wahrheit geht auf Popper [10] zurück. Für ihn gibt es keine absolute Wahrheit, sondern nur wahrheitsähnliche Theorien. Stimmt eine Theorie mit den Tatsachen überein, ist sie bezüglich dieser Tatsachen wahr (verifiziert). Sie erklärt die Tatsachen. Stimmt

¹ Diese Definition ist anders als etwa in der Logik, wo ein Modell ein Objekt ist, das eine Formel erfüllt, oder in der Datenbanktheorie, wo ein Modell die Typinformation der Daten oder das Datenschema ist.

sie nicht überein, ist sie falsch (falsifiziert). Sie wird dann geändert oder verworfen. Daraus folgt, dass wir die Welt um und in uns nur soweit verstehen, wie es dazu eine Theorie gibt, die noch nicht falsifiziert wurde. Mathematische Wahrheiten beruhen auf Theoremen, deren Richtigkeit nicht durch Beobachtungen, sondern durch einen Beweis abgeleitet wurde.

Nicht alles Wissen lässt sich als Information darstellen. Dieser Anteil (das „informatisierte“ Wissen) nimmt aber stetig zu. Information kann falsch oder redundant sein. Wissen kann unvollständig oder irrelevant sein. Können oder Kompetenz basieren meist auf Wissen, sind aber etwas anderes. Tennis- oder Klavierspielen zu lernen aufgrund des Wissens, das man anhand eines Vortrags oder aus einem Buch gewonnen hat, das schafft kaum jemand. Bei Können gehört Übung ganz entscheidend dazu.

3 Kategorien des Wissens

Die Kriterien, nach denen wir Wissen aufteilen, sind nicht sauber getrennt (nicht orthogonal), sondern teilweise überlappend. Einige der wichtigsten Kategorien sollen kurz vorgestellt werden.

- *Nicht-symbolisches und symbolisches Wissen:* Bevor wir Dinge mit Begriffen und Namen bezeichnen, existieren sie sowohl in der Natur als auch in unserem Bewusstsein. Dort steht dann der optische, akustische, haptische, elektrische oder chemische Eindruck, den das betreffende Objekt macht, für sich selbst. Wir sprechen in diesem Fall von nicht-symbolischen Wissen. Auch Menschenaffen erkennen Werkzeuge (z.B. eine Termitenangel) gemäß Funktion und Eigenschaften. Sie können sie nachahmen, aber sie können sie nicht benennen. Geben wir Dingen einen Namen, so benutzen wir diesen, wenn wir auf ein Objekt verweisen. Bei diesem symbolischem Wissen müssen wir zwischen dem Objekt und seiner Bezeichnung unterscheiden. Diese Namen oder Symbole haben nicht nur die Form von Wörtern. Das können auch Standardbildzeichen (Piktogramme) sein wie bei unseren Verkehrszeichen, Ziffern oder Wurzelzeichen wie in der Mathematik, Schaltzeichen wie in der Elektrotechnik oder Strukturformeln wie in der Chemie.
- *Allgemein- und Spezialwissen:* Allgemeinwissen hat den Charakter von Breitenwissen, d.h. man weiß von sehr vielen Dingen etwas. Man kann schnell neue Dinge einordnen, ist aber bei Detailfragen leicht überfordert. Es sind Millionen von Menschen, die über dasselbe Allgemeinwissen verfügen, eben ein Kulturkreis. Spezialwissen ist eher Tiefenwissen. Man weiß über wenige Dinge sehr viel.

Spezialwissen macht seinen Besitzer zum Experten für ein Thema oder ein Fachgebiet. Das Spezialwissen unterliegt stetigen Änderungen bezüglich seiner Menge und seines Inhalts.

- *Aktives und passives Wissen*: Das Wissen, das wir laufend verwenden, ist unser aktives Wissen. Es wird weder aufgebraucht, noch gerät es in Vergessenheit. Allerdings ist der Umfang dessen, was wir als aktives Wissen besitzen, begrenzt. Haben wir mehrere Minuten Zeit, kann auch das Langzeitgedächtnis miteinbezogen werden. Seine Kapazität ist sehr groß und aus Sicht des Einzelnen so gut wie unbegrenzt. Sind Inhalte auch innerhalb von einigen Minuten nicht abzurufen, rechnen wir diesen Teil zum passiven Wissen. Um es wieder aktiv zu nutzen, muss es durch einen äußeren Anstoß aufgefrischt oder reaktiviert werden.
- *Explizites und implizites Wissen*: Nicht alles Wissen, das ein Individuum oder eine Gruppe benutzt, besteht aus bekannten Tatsachen oder Regeln. Wurden diese als Wissen erkannt und formuliert (externalisiert), bezeichnet man es als explizites Wissen. Wir handeln manchmal vernünftig, ohne zu wissen warum. Dann benutzen wir implizites Wissen (engl. tacit knowledge). Implizites Wissen kann man lernen, indem man andere Menschen beobachtet und nachahmt. In ihren ersten Lebensjahren lernen Kinder sehr viel implizites Wissen. Der Prozess, der dies bewirkt, heißt Sozialisation. Implizites Wissen kann sich sowohl auf symbolisches wie nicht-symbolisches Wissen beziehen.

Die Bedeutung des impliziten Wissens wird von vielen Autoren hervorgehoben, so z.B. von Polanyi [9] in wissenschaftlich-philosophischer Hinsicht und von Nonaka/Takeuchi [8] in Bezug auf das Wissens-Management.

4 Information, die kein Wissen ist

Für eine Abgrenzung des Begriffs Wissen ist es hilfreich zu betrachten, welche Arten von Information man *nicht* als Wissen bezeichnen kann. Wenn man Information mit Schall vergleicht, verhalten diese sich zu Wissen etwa so wie Geräusche zu Musik. Die folgenden Beispiele fallen in diese Kategorie:

- *Unterhaltung*: Zur Unterhaltung dienen die meisten der verkauften und gelesenen Bücher und Zeitschriften, fast alle Musikaufnahmen und die meisten Filme und Fernsehsendungen.
- *Werbung*: Obwohl verlässliche Produktinformation sehr wichtig ist, um wirtschaftlich sinnvoll wählen und handeln zu können, geht Werbung leider oft darüber hinaus.
- *Kunst*: Viele Kunstformen bedienen sich der Information als Ausdrucksmedium. Die bekanntesten sind Belletristik (die schöngeistige Literatur), Musik und Malerei.
- *Erbauung*: Für die meisten Menschen ist die Besinnung auf philosophische oder religiöse Themen dazu geeignet, um zu Bescheidenheit, positiver Lebenshaltung und Solidarität mit den Mitmenschen anzuregen.
- *Indoktrination*: Die politische Agitation und Propaganda ist das abschreckende Beispiel der Nutzung von Information für die Irreführung von Menschen.
- *Steuerung und Automation*: Um Werkzeuge und Geräte des täglichen Lebens zu steuern, wird in zunehmendem Maße von Information Gebrauch gemacht.

- *Genetik*: Eines der herausragenden Merkmale lebender Wesen ist die Tatsache, daß von ihnen Information gespeichert und genutzt wird. Information steuert ihren Wachstumsprozess und bestimmt damit die Besonderheit jeder Art und Rasse.

Bei den letzten beiden Beispielen handelt es sich typischerweise um Wissen, das einmal von Menschen oder von der Natur erworben wurde. Maschinen bzw. Organismen benutzen diese Information, ohne in der Lage zu sein, sie (erneut) zu bewerten². Ohne Zweifel sind einige der oben erwähnten Informationsinhalte wirtschaftlich gesehen attraktiver als die Nutzung von Wissen. Es ist deshalb durchaus sinnvoll, Wissensvermittlung mit Unterhaltung zu verbinden.

5 Entstehen und Verändern von Wissen

Wissen ist nicht statisch, weder für den Einzelnen noch für die Gesellschaft. Der Bedarf verändert sich im Laufe der Zeit, ebenso wie das Angebot. Lernen ist Gewinnen von Wissen und Können. Das gilt für Individuen wie für Gruppen. Das individuelle Lernen ist besonders ausgeprägt in den ersten Lebensjahren. Die Gesamtheit des Wissens aller Menschen umfasst nicht nur das Wissen aller Menschen, die heute leben, sondern auch das Wissen früherer Generationen. Jeder Versuch, dieses Wissen insgesamt zu definieren, ist zwecklos. Es ist nur in Teilaspekten möglich. Bibliothekare haben jedoch sehr umfassende Klassifikationssysteme bereitgestellt [5].

Neues Wissen kann erworben werden durch Experimente, Beobachtungen oder durch Nachdenken. Seit einigen Jahrzehnten beteiligen sich immer mehr Menschen an dieser Aufgabe, außerdem verfügen sie über immer bessere Methoden und Werkzeuge, um neues Wissen zu erzeugen und zu verbreiten. Als Folge davon vermehrt sich der Umfang von Wissen immer schneller. Das Gedächtnis und die Lebensdauer eines Einzelnen reichen nur aus, um einen kleinen Ausschnitt des Wissens der Menschheit zu erfassen. Lernen auf Vorrat wird immer problematischer. Wird Wissen nicht benutzt, so kann es in Vergessenheit geraten. Es wird von anderem Wissen verdrängt. Es ist unklar, ob wir Wissen, das einmal im Langzeitgedächtnis enthalten war, jemals ganz verlieren.

6 Wissensspeicherung in Maschinen

Wegen der im vorigen Abschnitt angedeuteten Probleme sind Hilfsmittel für den Menschen von zunehmender Bedeutung. Früher waren es Bücher und Zeitschriften, heute sind es Rechner und Rechnernetze. Nur explizites Wissen kann in Büchern und Maschinen gespeichert werden. Das meiste Wissen ist an natürliche Sprachen in gesprochener oder geschriebener Form gebunden, wobei Englisch als Lingua Franca eine Sonderrolle spielt. Neben den natürlichen Sprachen gibt es Kunstsprachen wie Esperanto, die Gestensprache der Taubstummen oder die Programmiersprachen. Hier sind die Sprachelemente sowie die Grammatik den jeweiligen Bedürfnissen angepasst.

² Da Kinder auch meist nicht in der Lage sind, das was sie lernen zu bewerten, könnte man in diesen Fällen auch von (tradiertem) Wissen reden.

Gibt es eine Schriftform der Sprache, kann damit Wissen besonders gut dokumentiert, d. h. festgehalten werden. Früher war dies die einzige Art, um Wissen personen-unabhängig zu übertragen, also von einem Ort zum anderen und von einem Zeitraum in einen späteren. Heute gilt dies für die Medien Ton und Film ebenfalls. So kann man z.B. auch die Sprache der Eingeborenen von Papua-Neuguinea dokumentieren, ohne ihnen zuerst eine Schriftform zuordnen zu müssen. Welche Art der Speicherung am geeignetsten ist, hängt stark davon ab, welche Verarbeitung beabsichtigt ist. Zunächst ist entscheidend, ob ausschließlich an eine Verwendung durch Menschen oder durch Maschinen und Menschen gedacht ist. Ton- und Filmarchive fallen in die erste Kategorie.

Das Internet, oder genauer das Web, überholt derzeit die größten Bibliotheken der Welt, was den Umfang des Wissens betrifft, der vorgehalten wird. Schüler wie Studenten können Lehrmaterial zu fast jedem Thema finden³. Wissensprodukte werden in allen üblichen Medienformen angeboten, nämlich als Text, Bild, Ton und Film. Ein Nachteil des Web ist, dass es weniger gut strukturiert ist als die meisten Bibliotheken. Außerdem gibt es keine Trennung zwischen wissensrelevantem und nicht-wissensrelevantem Material. Das liegt an der Art der Beschaffung. Während Bibliotheken meist mit öffentlichen Mitteln dafür ausgestattet wurden, nur das anzuschaffen, was ein bestimmtes Experten- oder Beratergremium für nützlich hält, ist dies beim Web nicht der Fall. Jeder kann anbieten, was er anbieten möchte. Wer darüber zu sehr klagt oder es gar ändern will, muss sich fragen lassen, welche besseren, aber gleich kostengünstigen Alternativen vorgesehen werden sollen.

Wie Bibliotheken für ihre Bücher und Zeitschriften haben viele Web-Dateien außer ihrem inhaltlichen Text, den Bildern und Tönen, auch bibliografische Daten zu bieten. Dazu gehören Autor, Titel, Schlüsselworte und Kurzfassung. Man nennt diese Daten auch Metadaten. Eine Initiative, die das Ziel hat, die bisher von Bibliothekaren verwandten Metadaten zu vereinfachen, nennt sich Dublin Core⁴. Dieser Vorschlag soll Autoren dazu ermuntern, alle Metadaten gleich mit der ersten Bereitstellung einer Veröffentlichung selbst zur Verfügung zu stellen. Bei größeren Kompendien oder Zeitschriftensammlungen sind Querverweise (Hypertext-Netze) sinnvoll. Ein Beispiel dafür bietet der multimediale Brockhaus [2], und zwar basierend auf dem System Hyperwave [7]. Die dort gezeigten Wissensnetze wirken allerdings etwas willkürlich, da sie von den vom Autor des Texts vorgesehenen Links abhängig sind.

Seit einigen Jahren betreibt das Web-Koordinierungskomitee (W3C) den Plan, dem Web mehr beschreibende Information hinzuzufügen. Die Initiative hat den Namen Semantisches Web [1]. Die Grundtechnologien für das Semantische Web sind XML (eXtended Markup Language) und RDF (Resource Definition Facility). XML lässt es zu, dass Web-Seiten durch zusätzliche Merkmale (tags) annotiert werden. Mit RDF können dann weitere Aussagen codiert werden, so z.B. „ist Bruder von“. Damit wird eine semantische Beziehung zwischen Web-Einträgen hergestellt, ähnlich wie zwischen den einzelnen Begriffen in Abb. 1. Der nächste

Schritt besteht darin, sogenannte Ontologien zu verwenden. Das ist eine etwas großspürige Bezeichnung für Thesauren und Glossare. Darin können Unter- und Oberbegriffe sowie Synonyme definiert werden. Eine Suchmaschine kann dann Mehrdeutigkeiten (Polyseme) auflösen, wie die zwischen Bank als Möbelstück und Bank als Finanzinstitut. Ob diese Bemühungen einen signifikanten Einfluss haben werden, ist noch nicht abzusehen.

7 Wissensnutzung durch Maschinen

Die wichtigste Form der Nutzung von Wissen ist Suchen. Für das überreiche Angebot an Texten, Bildern und Musik gibt es im Internet jeweils spezielle Kataloge und Suchmaschinen. Das Prinzip von beiden ist in [5] beschrieben. Fast immer wird mit Schlüsselworten gesucht, sei es im Inhalt der Seite oder in der zusätzlichen Beschreibung (tags). Die Erfahrungen der letzten Jahre haben zu einer starken Polarisierung der Angebote geführt. Sie hat sich vor allem am Beispiel von Yahoo und Google gezeigt. Während Yahoo weiterhin eine Katalogisierung der Inhalte von Hand vornimmt, beruht der Erfolg von Google auf einer Suchmaschine mit einem besonders cleveren Bewertungsverfahren, nach dem für die gefundenen Objekte eine Reihenfolge (engl. Ranking) ermittelt wird. Dazu wird zunächst Information verwandt, die ohne inhaltliche Analyse von Texten gewonnen werden kann, allerdings mit erheblichem Rechenaufwand. Es werden dabei Gleichungen mit bis zu 500 Millionen Unbekannten gelöst⁵. Das Ergebnis wird anschließend durch eine inhaltliche Analyse und Bewertung einzelner Seiten verbessert. Aufgrund seines guten Ranking-Verfahrens hat Google praktisch alle anderen Suchmaschinen aus dem Internet verdrängt. Yahoos Ansatz ist dagegen sehr personalintensiv und kann daher nicht mit dem Wachstum des Internet Schritt halten. Yahoo ergänzt daher inzwischen seine Suchergebnisse durch den Aufruf von Google.

Eine andere Form der Wissensnutzung, die mittels Maschinen versucht wird, ist das deduktive Schließen, auch Inferenz genannt. Diese Anwendung läuft meist unter der Bezeichnung Expertensysteme. Hierfür ist es erforderlich, Wissen formal darzustellen. Das für die Anwendung benötigte Wissen wird dazu in eine Sprache übertragen, in der sowohl die Namen der Objekte als auch die Beziehungen oder Relationen zwischen diesen leicht erkennbar sind, so dass sie maschinell, d.h. von Programmen, ausgewertet werden können. Es gibt dazu viele verschiedene Ansätze. Die bekanntesten basieren auf Fakten und Regeln. Fakten beziehen sich auf Objekte und geben bestimmten zugehörigen Attributen einen Wert. Regeln bestehen aus Prämissen und Folgerungen. Die Prämisse besagt, wann eine Regel zur Anwendung kommt. Andere Ansätze zur Wissensrepräsentation bedienen sich der mathematischen Logik (PROLOG) oder der Beschreibung von Objekten mittels semantischer Netze wie in Abb. 1 [11]. Es ist dies eine grafische Repräsentation von semantischen Beziehungen, die mittels einer formalen Sprache (KL-ONE) dargestellt werden können. Vermutlich gibt es inzwischen noch mächtigere Sprachen.

Es wird kaum möglich sein, mehr als nur ganz kleine Ausschnitte des für uns relevanten Wissens in dieser Weise formal zu beschreiben. Den wenigen Schulbeispielen, wo dies gelungen ist, stehen jederzeit Berge von Wissen gegenüber, die nur

³ Übersichten findet man u.a. unter http://dir.yahoo.com/Education/Distance_Learning/

⁴ <http://dublincore.org/>

⁵ <http://www.google.com/press/overview.html>

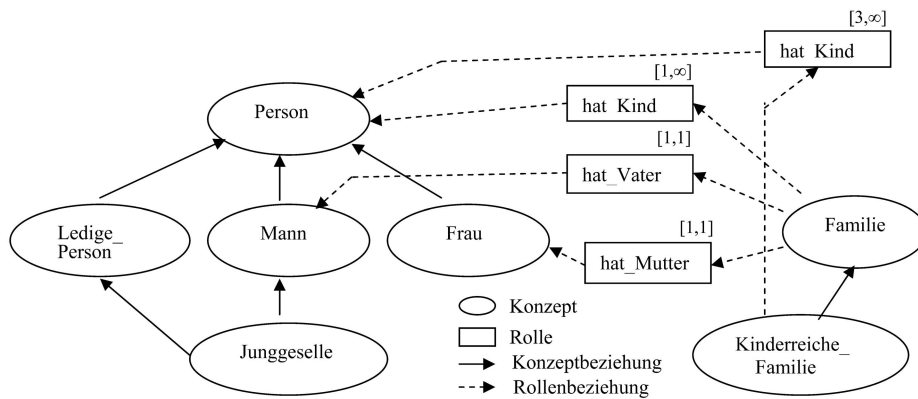


Abb. 1. Darstellung von Familienbeziehungen in KL-ONE [11]

in Form von Texten, Bildern oder gesprochener Sprache vorliegen (so wie dieser Artikel). Es ist kaum vorstellbar, dass sich dies in Zukunft wesentlich ändert. Waren in der ersten Begeisterungsphase für Expertensysteme noch viele Informatiker bereit, sich die Berufsbezeichnung Wissensingenieur zuzulegen, so ist davon kaum mehr die Rede. Dabei ist sehr zu wünschen, dass sie sich auch weiterhin intensiv um die maschinelle Verarbeitung von Wissen (und nicht nur von Information) kümmern, und zwar nicht nur mittels einer bestimmten, in der Fachwelt gerade populären Methode.

Wissens-Management befasst sich primär mit den organisatorischen Fragen aus der Sicht von Unternehmen. Es geht darum, wer wann welches Wissen benötigt und wie es beschafft werden kann [3]. Da hierbei aktuelles und privates Wissen im Vordergrund stehen, denkt man nur an elektronische Medien. Es steht außer Zweifel, dass für den Erfolg dieser Tätigkeit eine klare Unterscheidung von Information und Wissen sehr wichtig ist. Auch sollte dabei beachtet werden, dass Wissen auf Plattentürmen oder CDs erst in Köpfe transferiert werden muss, um nutzbar zu sein.

8 Können Maschinen denken?

Kognitive Leistungen dienen den Lebewesen zum Anpassen an ihre Umwelt sowie dem Menschen zur Selbsterkenntnis. Denken ist dabei die hervorstechendste Leistung. Es ist eine Leistung, die wir normalerweise nur dem Homo sapiens (auch Cro-Magnon-Mensch genannt) zusprechen. Dabei gibt es von seinen Vorfahren, dem Homo erectus und dem Neandertaler, bereits große Mengen von Geräten und Werkzeugen, die man als überlieferte Dokumente ihres Denkens ansehen muss. Der Frage, ob Maschinen denken können, kann eine Informatikerin oder ein Informatiker kaum ausweichen, sobald man sich mit maschineller Wissensverarbeitung beschäftigt. Sie ist so naheliegend, dass auch Laien sie immer wieder stellen.

Eine mögliche Antwort geht davon aus, dass man den Prozess des Denkens differenziert betrachtet. Die einfachste Form des Denkens ist die Suche im Gedächtnis. Darauf folgt das Klassifizieren und das analoge Schließen. Erst das Denken in sprachlichen Kategorien erleichtert das Abstrahieren und Generalisieren. Zuletzt kommen das deduktive Schließen und das Vorausdenken. Bei letzterem verlassen wir die Realität und extrapolieren aus der Vergangenheit in die Zukunft. Wir stellen uns dann Fragen der Art „Was wäre, wenn?“. Es ermöglicht

vorausschauendes Planen. Bei Rechnern heißt dies Simulieren. Die meisten maschinellen Ansätze beziehen sich auf das Suchen und das deduktive Schließen. Die größten Erfolge gibt es zweifellos beim Suchen. Hier sind Maschinen den Menschen bereits heute eindeutig überlegen. Beim Klassifizieren, Generalisieren und Abstrahieren tun sich Maschinen schon erheblich schwerer als Menschen. Als höchste Leistung des Menschen gilt die Selbstreflexion, d.h. das Nachdenken über unsere Denkvorgänge. Dies haben Plato und Aristoteles, die Väter der europäischen Philosophie, als erste unternommen. Hiermit sind Maschinen eindeutig überfordert, nicht nur heute, sondern vermutlich auch in Zukunft. Simulieren ist ähnlich wie Suchen. Hier können Maschinen die Leistungen des Menschen um Größenordnungen übertreffen, vorausgesetzt, wir verfügen über entsprechende Theorien und Modelle.

Ein anderer Weg, um zu entscheiden, ob Maschinen denken können, ist der berühmte Turing-Test. Danach gilt diese Frage als beantwortet, sobald das Antwortverhalten eines Rechners nicht mehr von dem eines Menschen zu unterscheiden ist. Es gibt sogar einen diesbezüglichen Wettbewerb (den Loebner-Preis⁶), für den jedes Jahr ein Zwischenpreis vergeben wird. Er wird bereits für kleine Teilerfolge („den menschlichsten Rechner“) vergeben. Vermutlich wird es auch dabei bleiben. Wie das Schachspielen zeigt, sind die Chancen mit dem Menschen gleichzuziehen, umso besser, je enger der Blickwinkel ist. Sobald Breitenwissen und soziales Verhalten gefragt sind, tun sich Rechner heute und auch in Zukunft enorm schwer. Ob dies nur von der Menge und der Art des erforderlichen Wissens abhängt, oder ein generell unlösbares Problem ist, sei dahingestellt.

Die Frage, ob Maschinen Texte oder die Welt (je) so verstehen, wie dies Menschen tun, ist damit nicht berührt. Es ist eigentlich auch keine sehr sinnvolle Frage. Moderne Computer können zwar immer größere Mengen von Wissen vorhalten. Sie können damit leicht jede konventionelle Bibliothek überbieten. Auch ist es in ihrem Falle leichter herauszubekommen, was hinter den Buchdeckeln steht als bei einer konventionellen Bibliothek. Jeder Rechner kann nämlich Texte, Bilder und Sprachaufzeichnungen mit einer Gründlichkeit und Ausdauer analysieren, wie dies kein Mensch kann oder will. Dennoch wird er nicht dieselbe Art von Gedanken mit dem Gelesenen verbinden wie ein Mensch, es sei denn er hätte genau denselben Werdegang gehabt und hätte dabei dieselben Erlebnisse als wichtig angesehen.

⁶ <http://www.loebner.net/Prizef/loebner-prize.html>

9 Pilotcharakter der maschinellen Wissensverarbeitung

Die Wissensverarbeitung ist nur in technischer Hinsicht als eine Weiterentwicklung der Informationsverarbeitung anzusehen. Anders als etwa von den Förderern des japanischen Fifth-Generation-Projekts erwartet, ist sie nicht zum Hauptstrom ihrer wirtschaftlichen Entwicklung geworden, sondern eher zu einem elitären Seitenzweig. In wirtschaftlicher Hinsicht dominieren die geschäftliche Datenverarbeitung sowie Unterhaltung und Werbung, und werden es weiter tun. Dennoch können die Fortschritte und die Erfahrungen, die auf diesem Gebiet erzielt werden, auf alle anderen Anwendungen ausstrahlen. Dabei steht besonders die inhaltsbezogene (semantische) Beschreibung und Strukturierung von Wissen im Vordergrund. Andererseits können die Anbieter von Wissensprodukten sehr viel von dem lernen, was etwa die Musikindustrie bezüglich ihrer Abrechnungsverfahren oder dem Schutz des geistigen Eigentums als praktikable Lösungen zustande bringt. Auch hier kann ein Anwendungsbereich der Informatik andere befruchten.

Literatur

1. Berners-Lee T, Hendler J, Lassila O (2001) The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American* May 17
2. Der Brockhaus Multimedial (2003) Brockhaus, Mannheim
3. Endres A (2003) Die Wissensgesellschaft und ihr Bezug zur Informatik. *Informatik-Spektrum* **26**(3):195–200
4. Endres A (2003) Der Informationsbegriff – eine informatikorientierte Annäherung. *Informatik – Forsch. und Entw.* **18**(2):88–93
5. Endres A, Fellner DW (2000) *Digitale Bibliotheken. Informatik-Lösungen für globale Wissensmärkte.* dpunkt, Heidelberg
6. Klix F (1993) *Erwachendes Denken. Geistige Leistungen aus evolutionspsychologischer Sicht.* Spektrum, Heidelberg
7. Maurer H (2003) Wissensmanagement. Ein Schritt nach vorne oder nur ein neues Schlagwort. *Informatik-Spektrum* **26**(1):26–33
8. Nonaka I, Takeuchi H (1997) *Die Organisation des Wissens.* Campus, Frankfurt
9. Polanyi M (1969) *Knowing and Being.* The University of Chicago Press, Chicago 1969
10. Popper K (1994) *Vermutungen und Widerlegungen: Das Wachstum der wissenschaftlichen Erkenntnis.* Mohr, Tübingen
11. Puppe F (1991) *Einführung in Expertensysteme.* Springer, Heidelberg