



Psychologische Theorien sind Information-Retrieval-Verfahren

Manfred Wettler

Universität Paderborn
Fach Psychologie
33098 Paderborn
wettler@psycho.upb.de

Abstract

Psychological models of learning, forgetting, and memory processes are presented. These models describe human information processing as the formation and retrieval of associations. They use the same equations as the algorithms which are used for Information Retrieval and for Data Mining. A possible explanation for these correspondences is presented.

Zusammenfassung

Es werden psychologische Modelle des Lernens, des Vergessens und des Gedächtnisses vorgestellt. Menschliche Informationsverarbeitung wird darin durch die Bildung und den Abruf von Assoziationen erklärt. Die in den psychologischen Theorien verwendeten Gleichungen finden sich in auch in den im Information Retrieval und im Data Mining verwendeten Algorithmen. Es wird eine mögliche Erklärung für diese Übereinstimmungen diskutiert.

1 Einleitung

Die Beantwortung einer Anfrage durch die Suchmaschine Google verläuft in zwei Schritten. Zunächst werden die Dokumente im www bestimmt, in denen die in der Anfrage verwendeten Wörter vorkommen. Darauf werden diese Dokumente entsprechend ihrer Wichtigkeit in eine Rangreihe geordnet. Dazu verwendet Google den sog. PageRank-Algorithmus. Dieser richtet sich nicht nach den Inhalten der Seiten sondern nach der Verweisstruktur: Je häufiger auf ein Dokument verwiesen wird, desto wichtiger sei es. Dabei wird nicht nur die Anzahl der Verweise auf ein Dokument berücksichtigt, sondern auch die Wichtigkeit der verweisenden Dokumente und die Anzahl ihrer Verweise auf andere Dokumente. Die folgende vereinfachte Gleichung definiert den



von Google verwendeten Algorithmus zur Bestimmung des Rangplatzes eines Dokumentes (Brin & Page, 1998; Brin, Page, & Winograd, 1998):

$$R(u) = \sum R(v)/N_v$$

(1)

Der Rangplatz (Wichtigkeit) eines Dokumentes u entspricht der Summe aus den Rangplätzen aller anderen Dokumente v , die auf u verweisen, dividiert durch die Anzahl der Verweise des jeweiligen Dokumentes. Die so definierte Wichtigkeit entspricht der Häufigkeit, mit der ein hypothetischer Zufallssurfer die Seite aufrufen würde. Ein solcher Zufallssurfer wählt zunächst eine zufällige Seite im *www* aus, klickt dann auf einen zufällig ausgewählten Verweis auf dieser Seite, wiederholt diesen Vorgang bei den nächsten Seiten bis er genug hat und mit einer anderen, ebenfalls zufällig ausgewählten Seite neu startet.

Ferber (2003) hat darauf hingewiesen, dass dieser durch den PageRank-Algorithmus beschriebene Prozess einem spreading activation-Verfahren entspricht, wie es in psychologischen Lern- und Gedächtnismodellen postuliert wird. So ist nach Anderson (2001) die assoziative Aktivierung (Wichtigkeit) eines Begriffs u die Summe der Aktivierungen aller mit ihm assoziierten Begriffe v , dividiert durch die Anzahl der von v ausgehenden Assoziationen.

$$Akt(u) = \sum Akt(v) / N_v$$

(2)

Durch den Divisor N_v kann dem von Lewis & Anderson (1976) beobachteten Fächereffekt Rechnung getragen werden, der besagt, dass die Summe der von einem Begriff ausgehenden Aktivierungen konstant ist, d.h. unabhängig von der Anzahl der Verbindungen zu anderen Begriffen. Je größer die Anzahl N_v der von einem Begriff v ausgehenden Assoziationen, desto schwächer werden die einzelnen mit ihm verbundenen Begriffe aktiviert.

Gleichung (1) beschreibt die Gewichtung von Dokumenten in Abhängigkeit von der Verweisstruktur im *www* und Gleichung (2) beschreibt assoziative Aktivierungen von Begriffen in Abhängigkeit von der assoziativen Struktur einer Person. Die Beziehungen zwischen den Termen sind in den beiden Gleichungen identisch.

In den folgenden Abschnitten werden weitere Beispiele gegeben, in denen Übereinstimmungen zwischen psychologischen Beschreibungen und informationswissenschaftlichen Algorithmen bestehen. Dabei soll auch der Frage nachgegangen werden, wie es zu diesen Übereinstimmungen kommt.

2 Lernen

Wenn beliebige seelische Gebilde gleichzeitig oder in naher Aufeinanderfolge das Bewusstsein erfüllt haben, so ruft hinterher die Wiederkehr einiger Glieder des früheren Erlebnisses Vorstellungen auch der übrigen Glieder hervor, ohne dass für sie die ursprünglichen Ursachen gegeben zu sein brauchen.

(Ebbinghaus)

Nach der klassischen Assoziationstheorie ist Lernen das Verknüpfen von wahrgenommenen Ereignissen, die in unmittelbarer zeitlicher Nähe zueinander aufgetreten sind. Auf der Grundlage dieses Gesetzes sind eine Reihe von mathematischen Modellen zur Beschreibung von Lernprozessen entwickelt worden. In neuerer Zeit hat eine Variante des Assoziationsgesetzes, die sog. Hebb'sche oder Delta-Regel bei der Entwicklung von lernenden neuronalen Netzen eine starke Popularität gewonnen.

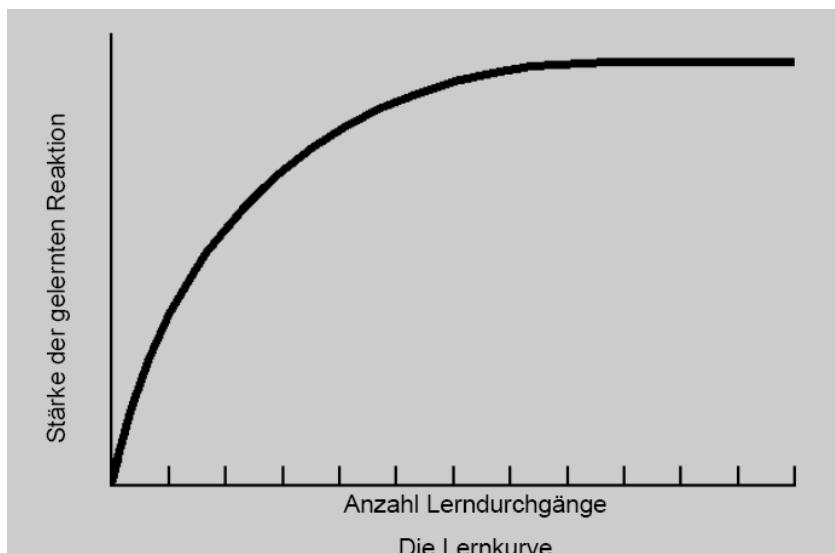


Abb. 1: Die Lernkurve

Verschiedene Theorien innerhalb des assoziationspsychologischen Ansatzes unterschieden sich im Hinblick auf die Lernsituationen, auf die sie sich beziehen, im Hinblick darauf, was gelernt wird (Lernen von Signalen, Lernen

etwas zu tun usw.) und im Hinblick auf die Annahmen über den Einfluss bestehender Assoziationen auf den Lernprozess. Übereinstimmung besteht über die folgenden drei Annahmen:

1. Die Stärke assoziativer Verknüpfungen schwankt zwischen Null und einem endlichen Wert, der maximalen Assoziationsstärke. In Bezug auf den Schwankungsbereich der Assoziationsstärken werden unterschiedliche Annahmen vertreten. Bei der folgenden Gleichung (3) können die Stärken von Assoziationen zwischen 0 und +1 schwanken.
2. Wenn zwei Ereignisse gemeinsam auftreten, dann erhöht sich die Stärke der Assoziation um einen konstanten Anteil des maximal möglichen Lernbetrages. Durch diese zweite Annahme kann der typische Verlauf von Lernkurven erklärt werden, die abbilden, dass bei gleich bleibendem Lernstoff zunächst viel und im weiteren Verlauf immer weniger gelernt wird.
3. Wenn eines der beiden Ereignisse ohne das andere auftritt, dann verringert sich die Assoziationsstärke um einen konstanten Anteil ihrer bisherigen Größe. Vergessen wird also nicht dadurch erklärt, dass die Spuren der gelernten Inhalte sich im Laufe der Zeit abschwächen, sondern durch Interferenzen, d.h. das Überlagern von früher gelerntem durch spätere Erfahrungen.

Diese drei Annahmen sind in der folgenden Gleichung (3) realisiert:

$$ASS_{u \leftarrow v} (n) = ASS_{u \leftarrow v} (n-1) + (\lambda - ASS_{u \leftarrow v} (n-1)) * \theta \quad (3)$$

Dabei sind:

$ASS_{u \leftarrow v} (n)$ die Stärke der Assoziation von einem Wort v zu einem Wort u zum Zeitpunkt n

λ der Wert der maximalen Assoziationsstärke, meist 1, wenn u und v gemeinsam aufgetreten sind, und 0, wenn nur eines der beiden Wörter, u oder v , aufgetreten ist.

θ ein freier Parameter, die sog. Lernkonstante.

Aus Gleichung (3) folgt nach Estes (1950), dass bei langfristigem Lernen die Stärke der Assoziation von einem Ereignis v zu einem zweiten Ereignis u sich der bedingten Wahrscheinlichkeit von u , gegeben v , annähert. Assoziatives Lernen führt also dazu, dass die Zukunft vorhersehbar wird. Dies gilt für alle Arten von Lernprozessen gleichermaßen; der Hund in Pawlow's Laboratorium lernt vorherzusagen, ob er Futter bekommt und der Student der Astronomie lernt vorherzusagen, wann es eine Mondfinsternis geben wird.

Was nützt es dem Hund, einen bedingten Reflex zu lernen, die bedingte Wahrscheinlichkeit, dass nach dem Ton das Fressen kommt? Im Labor wenig, aber draußen, wo es wenig Nahrung und viele Konkurrenten gibt, wo das Rudel sich ein Kaninchen teilt, da bekommt derjenige Hund am meisten ab, der am schnellsten frisst, weil er früher als die andern gewusst hat, dass sie das Kaninchen bekommen, bei dem schon vorher die Spucke im Maul zusammengelaufen ist. Assoziativ gelernt zu haben heißt, sich an frühere Ereignisse zu erinnern, die geschehen sind, als es so wie jetzt war. Dies ist jedem nützlich, auch dem Kaninchen, wenn es gelernt hat, wo die Hunde lauern.

3 Wortassoziationen

Um die Gültigkeit assoziativer Lerntheorien zu überprüfen und um damit Verhaltensvoraussagen zu machen, ist es notwendig die Lerngeschichte möglichst vollständig zu kennen. In Situationen des täglichen Lebens trifft dies höchstens ausnahmsweise zu. Niemand kann von sich oder von anderen angeben, wann sie wofür belohnt oder bestraft wurden. Meine Erwartungen darüber, wie das Wetter wird, was es zu essen geben wird oder wie sich andere verhalten werden, sind das Produkt meiner ganzen Lerngeschichte, das heißt einer Vielzahl von nicht rekonstruierbaren Ereignissen. Aus diesem Grund arbeitet die Lernpsychologie seit Ebbinghaus (1885) mit künstlichem Lernmaterial, sinnlosen Silben zum Beispiel. Dies hat den Nachteil, dass die Übertragbarkeit der Ergebnisse auf Alltagssituationen fraglich bleibt.

Ein verhältnismäßig gut untersuchter Bereich assoziativen Lernens, der zum einen für das Verhalten in außereperimentellen Situationen relevant und zugleich quantitativ leicht überprüfbar ist, sind die sog. freien Wortassoziationen. Eine freie assoziative Antwort ist das erste Wort, das einer Person einfällt, nachdem sie ein anderes Wort gehört hat, den so genannten assoziativen Stimulus. Auf den Stimulus „Mann“ produzieren die meisten Leute die assoziative Antwort „Frau“ oder auf den Stimulus „hell“ das Wort „Licht“ oder „dunkel“. In der psychologischen Literatur der letzten hundert Jahre finden sich viele Belege dafür, dass die Wahrnehmung, das Lernen und das Erinnern von gesprochenen und geschriebenen Wörtern von der Stärke der Assoziationen zwischen den Wörtern abhängen. Assoziationen zwischen Wörtern, so William James (1890), sind die grundlegenden Verknüpfungen durch die der Fluss der Gedanken gesteuert wird.

Assoziative Strukturen können durch sog. assoziative Netze beschrieben werden. Jeder Knoten eines solchen Netzes steht für ein Wort und die Stärke

der Verbindung zwischen den Knoten entspricht der Assoziationsstärke. Die traditionelle von Deese (1966) verwendete Methode zur Konstruktion solcher Netze besteht darin, dass man die Antworten im freien Assoziationsexperiment zugrunde legt. Dazu wird zunächst ein Vokabular definiert, d.h. eine Menge von Wörtern, deren assoziative Verbundenheit untersucht werden soll. Danach wird jedes dieser Wörter einer möglichst großen Zahl von Personen vorgelegt mit der Aufforderung, das erste Wort zu nennen, das ihnen einfällt. Die Ergebnisse dieser Versuche können in einer quadratischen Tabelle festgehalten werden, die angibt, wie häufig jedes Wort als assoziative Antwort auf jedes andere Wort gegeben worden ist.

Eine zweite Methode besteht darin, Assoziationen aufgrund der statistischen Beziehungen zwischen Wörtern in Texten zu rekonstruieren. Dies ist möglich, weil einerseits die assoziativen Beziehungen zwischen Wörtern aufgrund der Häufigkeiten ihres gemeinsamen Auftretens in der Sprache gelernt worden sind und weil zugleich die Wahrscheinlichkeiten, dass Wörter in naher zeitlicher Aufeinanderfolge geäußert werden, von den Assoziationen des Sprechers bzw. des Autors abhängen. Für unsere korpusbasierten Berechnungen der Wortassoziationen verwenden wir auf CDs gespeicherte Zeitungstexte und Belletristik. Ein gemeinsames Auftreten zweier Wörter wird immer dann angenommen, wenn sie innerhalb eines Abstandes von 10 anderen Wörtern auftreten (Rapp, 1996).

Es gibt drei Verfahren wie Wortassoziationen als Funktion des gemeinsamen Auftretens von Wörtern in Texten berechnet werden können. Beim ersten Verfahren werden die bedingten Wahrscheinlichkeiten direkt aus den vorher ausgezählten Gesamthäufigkeiten des gemeinsamen Auftretens berechnet. Dazu wird eine Term \times Term-Matrix erstellt und normalisiert (Rapp, 1996; Wettler, 2002). Bei dem zweiten, von Landauer & Dumais (1997) verwendeten Verfahren wird eine Term \times Dokument-Matrix erstellt. Das dritte Verfahren ist eine direkte Simulation assoziativen Lernens. Dabei werden die Texte Wort für Wort eingelesen und nach jedem neuen Wort, das gelesen wird, die ganze Assoziationsmatrix nach Gleichung (3) neu berechnet (Wettler & Seidensticker, 2003). Bei drei allen Methoden sind die aus den Texten berechneten Wortassoziationen in der Regel von denjenigen menschlicher Teilnehmer im freien Assoziationsversuch nicht unterscheidbar.

Die erste Anwendung solcher Wortnetze war im Bereich des Information Retrieval. Dabei ging es allerdings nicht darum, neue Algorithmen für die Recherche in Datenbanken zu entwickeln sondern um die kognitionspsychologische Beschreibung der Suchfragenformulierung von

professionellen Datenbankrechercheuren. Das Material dieser von Ferber, Wettler & Rapp (1995) durchgeführten Untersuchung bestand aus 40 schriftlichen Anfragen von Endnutzern wissenschaftlicher Literaturdatenbanken und den Protokollen der entsprechenden, von professionellen Datenbankrechercheuren durchgeführten Recherchen. Für jede der 40 Recherchen wurden zunächst in einem assoziativen Wortnetz die Knoten für alle in den Problembeschreibungen der Endnutzer verwendeten Begriffe aktiviert. Diese Aktivitäten wurden nach einem spreading activation-Verfahren auf alle Knoten des Netzes verteilt. Es zeigte sich, dass es sich bei denjenigen Begriffen, die bei diesem Verfahren am stärksten aktiviert wurden, zugleich um die von den Rechercheuren verwendeten Suchbegriffe handelt. Es kann sowohl vorhergesagt werden, welche Begriffe aus den Problembeschreibungen die Rechercheure weglassen, als auch, welche zusätzlichen Begriffe sie in die Recherche aufgenommen haben.

Die beschriebene Untersuchung hatte eine psychologische Fragestellung, die Aufklärung der psychischen Mechanismen, welche die Auswahl der Begriffe bei der Suchfragenformulierung steuern. Dabei zeigte sich, dass dieses Verhalten mithilfe der Lerntheorie von Estes (1950) vorausgesagt werden kann. Da es sich bei den untersuchten Personen um professionelle Rechercheure mit langer Berufserfahrung handelt, darf man annehmen, dass Programme, mit denen das Verhalten von Rechercheuren simuliert wird, unabhängig von ihrer psychologischen Geltung Erfolg versprechende Kandidaten für die Suche nach Dokumenten sind.

4 Vergessen

Lernprozesse führen dazu, dass die unter den aktuellen Bedingungen wahrscheinlich eintretenden Ereignisse erinnert werden. Lernen ist also im Hinblick auf das allgemeine Informationsbedürfnis die Zukunft vorherzusagen optimiert. Ähnliches gilt auch für das Vergessen.

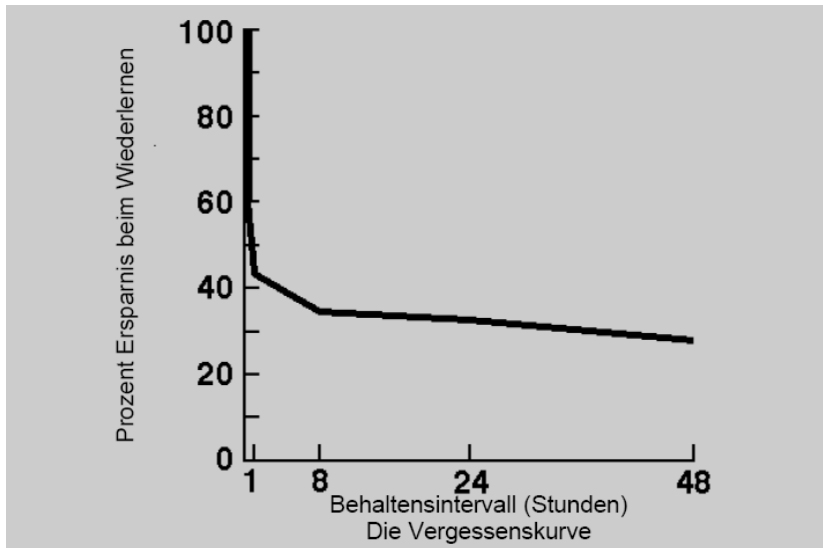


Abb. 2: Die Vergessenskurve

Ebbinghaus fand in Selbstversuchen 1885, dass unmittelbar nach dem Erlernen am meisten vergessen wird, und dass mit zunehmendem Intervall zwischen Erlernen und Reproduzieren immer weniger vergessen wird. Er folgerte daraus, dass pro Zeiteinheit ein konstanter Anteil der erinnerten Inhalte vergessen wird, dass die Stärke des Vergessens demnach proportional zum Logarithmus des Intervalls seit dem Erlernen sei. In einer Nachanalyse der Daten von Ebbinghaus fanden Anderson & Schooler (1991), dass Vergessen in einer Exponentialfunktion verläuft, das heißt, der Logarithmus des vergessenen Materials ist proportional zum Logarithmus der Zeit seit dem Erlernen.

Damit lassen sich Beobachtungen erklären, die den intuitiven Vermutungen über den Verlauf des Vergessens widersprechen. Dazu gehört der sog. Intervall (spacing)-Effekt. Im Allgemeinen werden Ereignisse umso schlechter erinnert, je länger sie her sind. Das ist jedoch nicht immer so. Nehmen wir an, wir hätten ein Ereignis, eine Person, oder einen Film zwei Mal gesehen, das letzte Mal vor einem Jahr. Wenn diese zweite Wahrnehmung kurz nach dem ersten Mal stattgefunden hat, dann wird das Ereignis schlechter erinnert, als wenn die erste Wahrnehmung schon länger her ist. Durch diesen Intervalleffekt kann es also dazu kommen, dass ältere Gedächtnisspuren besser erinnert werden als jüngere.

Wie die Lernfunktion so ist auch der Verlauf von Vergessensprozessen eine Anpassung an die Kontingenzen in der Umwelt. Anderson & Schooler (1991) untersuchten dies anhand von drei Arten von Ereignissen. Als erstes berechneten sie die Wahrscheinlichkeiten, dass einzelne Begriffe in einer bestimmten Ausgabe der New York Times erwähnt wurden in Abhängigkeit

von der Anzahl der Tage im vergangenen Jahr, in denen diese Begriffe in dieser Zeitung erwähnt wurden, und in Abhängigkeit von der Anzahl Tage, seit denen die Begriffe zum letzten Mal erwähnt wurden. Sie fanden, dass diese „objektiven“ Wahrscheinlichkeiten der Stärke der Erinnerung entsprachen, die beim Vergessen nach der Exponentialfunktion resultiert. Vergessen verläuft also derart, dass die Stärke der erinnerten Spur proportional zur Wahrscheinlichkeit ist, dass der erinnerte Inhalt gebraucht wird. Ähnliche Berechnungen wurden mithilfe des CHILDES Korpus von MacWhinney & Snow (1990) durchgeführt, einer Sammlung von Protokollen sprachlicher Interaktionen von Kindern im Stadium des Spracherwerbs. Auch hier wurden die Wahrscheinlichkeiten des Auftretens von Begriffen in Abhängigkeit von der bisherigen Häufigkeit und von dem Intervall seit der letzten Erwähnung untersucht. Es zeigte sich, dass die Wahrscheinlichkeiten des Auftretens einzelner Wörter proportional zur berechenbaren Stärke der Erinnerung an die entsprechenden Wörter ist. In einer dritten Untersuchung wurden Statistiken über die Absender von elektronischen mails erstellt. Auch hier zeigte sich, dass die Wahrscheinlichkeit, dass eine mail von einem bestimmten Absender eintrifft, der Stärke der Erinnerung an diese Person entspricht, die unter der Annahme einer exponentiellen Vergessensfunktion aufgrund früherer mails vorausgesagt wird.

Wie das Lernen, so ist auch das Vergessen ein Anpassungsprozess, der uns Vorhersagen zukünftiger Ereignisse ermöglicht. Auch hier ist die direkte Übertragbarkeit psychologischer Gesetze auf den Bereich von Information und Dokumentation gegeben. Nehmen wir an, ein Bibliothekar möchte die Bücher der Bibliothek so einordnen, dass er bei der Ausleihe möglichst wenig laufen muss. Er sollte, so Burrell (1980), diejenigen Bücher in seine Nähe stellen, die er aufgrund früherer Ausleihen am besten erinnert. Nach dem gleichen Prinzip sollte und wird auch in Datenbanken nach Dokumenten gesucht.

5 Synthese

Es sind drei allgemeine psychologische Gesetze beschrieben worden und für jedes dieser Gesetze ein entsprechender Algorithmus von Information Retrieval-Systemen. In Abschnitt 1 handelte es sich um den in Gleichung (2) definierten spreading activation-Prozess und den in Gleichung (1) beschriebenen PageRank-Algorithmus. Die Abschnitte 2 und 3 haben gezeigt, dass mithilfe des Assoziationsgesetzes berechnet werden kann, welche Suchwörter bei der Dokumentsuche verwendet werden sollen. In Abschnitt 4

sind es die Vergessensfunktion und die Wahrscheinlichkeiten, dass Dokumente benötigt werden. Wie kommt es zu diesen Übereinstimmungen?

Die sinnvolle Beschäftigung mit Information Retrieval und Data Mining setzt voraus, dass wir Kriterien dafür haben, was eine gute Recherche sei. In der traditionellen Forschung wird dies durch den Vergleich einer von Fachleuten ausgewählten und einer durch eine Suchwortliste eingegrenzten Menge von Dokumenten realisiert. Dies ist ein sicheres und für den Fortschritt der Forschung notwendiges Verfahren. Es ist jedoch extensional in dem Sinne, dass über die inhaltliche Beziehung zwischen Suchproblem und Dokumentenmenge nichts ausgesagt wird. Dieser Mangel findet sich auch in den Arbeiten über Benutzermodelle, in denen versucht wird, die Güte von Recherchen auf dem Umweg über Rekonstruktionen des „eigentlichen“ Informationsbedürfnisses der Benutzer zu bestimmen. Eine befriedigende Antwort auf die Frage nach der „guten Suche“ setzt eine Modellierung des domains voraus, des entsprechenden Wissensbereiches.

Die erwähnten Übereinstimmungen zwischen psychologischen Beschreibungen von Lern- und Vergessensprozessen einerseits und im Information Retrieval verwendeten oder verwendbaren Algorithmen andererseits zeigt jedoch, dass beide Prozesse durch das gleiche allgemeine und bereichsübergreifende Ziel bestimmt werden: vorauszusagen, was kommt.

Dies scheint mir eine wichtige Konsequenz aus dem von Kuhlen (2000) geforderten pragmatischen Primat der Informationswissenschaft: Wissen wird in Handlungsbezügen zu Information. Eine wichtige Determinante von Handlungen ist die Erinnerung an Gewesenes, die uns vorherzusagen erlaubt, was kommen wird. Nehmen wir an, eine Person interessiere sich für das Züchten von Rosen und sie gebe diese beiden Begriffe in einer Suchmaschine ein. Dann bekommt sie Dokumente, wie man Rosen pflanzt, einpflanzt, düngt usw. Die Dokumente beschreiben also, was passiert, wenn man Rosen züchtet, und das ist es, was sie wissen wollte. Der Benutzer möchte erfahren, was kommt. Dieses Bedürfnis wird am besten befriedigt, wenn Informationssysteme dem menschlichen Gedächtnis nachgebildet sind.

6 Literaturverzeichnis

Anderson, J. R. (2001³), *Kognitive Psychologie*. Heidelberg: Spektrum.

Anderson, J. R. & Schooler, L.J. (1991), Reflections of the environment in memory. *Psychological Science*, 2, 396-408.

- Bahrick, H. P. (1979), Maintenance of knowledge: Questions about memory we forget to ask. *Journal of Experimental Psychology: General*, 108, 296-308.
- Brin, S. & Page, L. (1998), The anatomy of a large-scale-hypertextual web search engine. Proceedings of the Seventh International Web Conference <http://www7.scu.edu.au/programme/fullpapers/1921/com1921.htm> [3/16/2004]
- Brin, S., Page, L., & Winograd, T. (1998), The PageRank citation ranking: Bringing order to the web. <http://www-db.stanford.edu/~backrub/pageranksub.ps> [3/16/2004]
- Burrell, Q.L. (1980), A simple stochastic model of library loans. *Journal of Documentation*, 41, 100-115.
- Ebbinghaus, H. (1885), *Über das Gedächtnis*. Leipzig: Duncker.
- Estes, W.K. (1950), Toward a statistical theory of learning. *Psychological Review*, 57, 94-117.
- Ferber, R. (2003), *Information Retrieval. Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web*. Heidelberg: dpunkt.verlag.
- Ferber, R., Wettler, M. & Rapp, R. (1995), An associative model of word selection in the generation of search queries, *Journal of the American Society for Information Science*, 46, 685-699
- James, W. (1890), *Principles of Psychology*. New York: Dover Publications
- Deese (1966), *The Structure of Association in Language and Thought*. Baltimore: The Johns Hopkins University Press.
- Kuhlen, R. (2000), *Studie "Nicht explizites Wissen" – aus der Sicht der Informationswissenschaft*. http://www.inf-wiss.uni-konstanz.de/People/RK/gutachten/gutachten_kuhlen0500_3.pdf [3/17/2004]
- Landauer, T.K. & Dumais, S.T. (1997), A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, **104**, 211-240.
- Lewis, C.H. & Anderson, J.R. (1976), Interference with real world knowledge. *Cognitive Psychology*, **7**, 311-335.
- MacWhinney, B. & Snow, C. (1990), The child language data exchange system: an update. *Journal of Child Language*, **17**, 457-472.
- Rapp, R., (1996), *Die Berechnung von Assoziationen: ein korpuslinguistischer Ansatz*. Hildesheim: Olms.
- Wettler, M. (2002), Free word associations and the frequency of co-occurrence in language use. In: P. Sedlmeier & T. Betsch (eds.), *Frequency Processing and Cognition*. Oxford: Oxford University Press, 471-484.
- Wettler, M. & Seidensticker, P. (2003) Learning free word associations from texts. *Proc. 25th Annual Meeting of the Cognitive Science Society*, Boston.