

Ähnlichkeit von Werkstoffen: Die Anwendung unterschiedlicher Wissensmodellierungstechniken für eine intelligente Komponente von WING

Michaela Ludwig

Informationswissenschaft, Universität Regensburg

Thomas Mandl

Informationszentrum Sozialwissenschaften, Bonn

8.1 Einleitung

Der folgende Beitrag beschreibt die Entwicklung einer Ähnlichkeitskomponente im Kontext des Projekts WING. Zunächst werden empirische Hinweise für die Notwendigkeit dieser intelligenten Zusatzkomponente zur Recherche nach ähnlichen Werkstoffen erörtert. Mehrere Systeme, die auf verschiedenen Wissensverarbeitungsmethoden basieren, wurden entwickelt und verglichen. Im einzelnen wird die Vorgehensweise bei einem regelverarbeitenden System (WING-SIM), einem statistisch basierten Ansatz (Cluster-WING) und einem auf neuronalen Netzen beruhenden System (NEURO-WING) dargestellt, wobei die Vor- und Nachteile der einzelnen Verfahren diskutiert werden. Das abschließende Kapitel erläutert die Integration der Komponente in das Gesamtsystem WING-M2 und die konzeptuelle Einbindung in das WOB-Modell.

8.1.1 Ähnlichkeitsproblematik

Ein Schwerpunkt des Projektes WING war die Entwicklung intelligenter Komponenten zur Unterstützung des Retrievalprozesses. Eine dieser Komponenten ist die im System WING-M2 als spezielles Werkzeug integrierte Ähnlichkeits-

komponente. Die Ähnlichkeit von Werkstoffen konnte als Spezialfall der nicht-exakten Anfragemöglichkeiten und des *soft computing* isoliert und als eigenständiges Werkzeug realisiert werden (cf. Mandl/Womser-Hacker 1995). Solche 'weichen' Verfahren spielen in modernen Informationssystemen eine immer größere Rolle, da sie dem Denken der Benutzer besser angepaßt sind als, etwa die 'harte' Boolesche Logik und damit 'natürlicher' zu bedienen sind.

Ähnlichkeitsbeziehungen sind elementar für die menschliche Kommunikation sowie für die grundlegenden kognitiven Prozesse. Wir ordnen Objekte, die wir wahrnehmen, bzgl. ihrer ähnlichen Eigenschaften. Demgegenüber ist Ähnlichkeit auf einem Computer nur schwer zu modellieren. Ein gutes Beispiel für ein realisiertes Ähnlichkeitstool ist die Arbeit von Escobedo et al. (1993), das ähnliche Konstruktionsteile für Flugzeuge erkennen kann und bei der Firma Boeing eingesetzt wird. Bei der Entwicklung tauchten typische Probleme bei der Modellierung von Ähnlichkeit auf: Zum einen ist Ähnlichkeit gradueller Natur. Sie kann außerdem auf unterschiedliche Bezugspunkte referieren. Überdies können Ähnlichkeitsurteile von Experten zu Experten völlig unterschiedlich ausfallen und Ähnlichkeit muß nicht transitiv sein.

8.1.2 Fachlicher Kontext Werkstoffinformation

Die Entwicklung einer Ähnlichkeitskomponente war empirisch motiviert. Bereits bei den ersten Befragungen von Werkstoffexperten zum Anforderungsprofil tauchte der Begriff *Ähnlichkeit* in natürlichsprachlich formulierten Anfragen auf.

Z. B.: „Ich suche Daten zu einem *ähnlichen* Werkstoff“ (Womser-Hacker 1994)

Bei näherer Untersuchung stellte sich heraus, daß die Frage nach einem ähnlichen Werkstoff häufig in einer bestimmten Retrievalsituation gestellt wird. In der Werkstoffdatenbank des Anwendungspartners MTU gibt es zahlreiche Lücken, d. h. zu einzelnen Werkstoffen sind bestimmte Kennwerte nicht gemessen worden. Beim Auftreten solcher Lücken weichen die Experten häufig auf einen ähnlichen Werkstoff aus und recherchieren dazu den gewünschten Kennwert, von dem sie dann Rückschlüsse auf den ursprünglichen Werkstoff ziehen (cf. Mandl/Roppel 1992). Zur Unterstützung dieser Strategie kann ein Ähnlichkeitstool eingesetzt werden.

Bereits bei den ersten Gesprächen mit Experten wurde die starke Kontextabhängigkeit von Ähnlichkeit deutlich. Die Bedingungen für einen ähnlichen

Werkstoff können bei Hochtemperaturstählen ganz anders aussehen als bei Stählen.

Das Wissen über Ähnlichkeit ist in der Datenbank nicht explizit vorhanden, sondern muß erst aus den Daten gewonnen werden. Für diesen Mehrwertprozeß wird eine geeignete Methode zur Wissensmodellierung gewählt.

8.2 Modellierungstechniken

Um Ähnlichkeit maschinell zu modellieren, kann eine kognitionswissenschaftlich fundierte Theorie zugrundegelegt werden. Eine andere Möglichkeit wäre zu versuchen, nur das vorliegende Zahlenmaterial auszuwerten. Um die für diesen Kontext beste Methode zu finden, wurden mehrere Systeme entwickelt und verglichen. Als Maßstab sollte die Akzeptanz der gewonnenen Ergebnisse durch Experten des Anwendungspartners MTU dienen, die bei der Bewertung von inhaltlichen Kriterien ausgehen.

In der Kognitionswissenschaft gibt es im wesentlichen zwei konkurrierende Theorien, wie menschliches Denken auf einer Maschine abgebildet werden kann: regelbasierte Systeme der KI und neuronale Netze. Eine integrierende Theorie beider Denkrichtungen wurde von Smolensky (1988) entworfen, der festhält, daß menschliches Denken durch neuronale Vorgänge gesteuert wird und sich deren Auswirkungen zwar weitgehend, aber nicht vollständig durch Regeln beschreiben lassen.

Entsprechend wurden im Projekt WING ein regelbasiertes System und ein auf neuronalen Netzen aufbauendes System realisiert. Da neuronale Netze oft mit statistischen Verfahren verglichen werden und es naheliegt zu versuchen, die Ähnlichkeitsbeziehungen aus Statistiken über die Kennwerte zu bestimmen, wurde auch ein auf Statistik beruhendes System erarbeitet. Die Entwicklung dieser drei Systeme wird im folgenden dargestellt.

8.2.1 WING-SIM: regelbasierter KI-Ansatz

Der erste Ansatz zur Modellierung von Ähnlichkeit benutzt Regeln, die beschreiben, unter welchen Bedingungen Werkstoffe ähnlich sind und bleibt damit im von Newell/Simon (1976) vorgegebenen Rahmen für Künstliche Intelligenz. Diese Forschungsrichtung geht davon aus, daß Kognition durch Symbolmanipulation modelliert werden kann.

Die Entwicklung mündete in eine prototypische Implementierung in dem System WING-SIM (cf. Mandl/Roppel 1992). Um das Problem der Kontextabhängigkeit zu lösen, wurden in Zusammenarbeit mit Domänenexperten Kontexte definiert, innerhalb derer Ähnlichkeit betrachtet werden kann, wie etwa Anwendung, Herstellungsverfahren und chemische Zusammensetzung. Für jeden dieser Kontexte wurden, ausgehend von den Eigenschaften, heuristische Regeln für Ähnlichkeit formuliert. Um die graduelle Natur von Ähnlichkeit abzubilden, wurden die Regeln verschieden stark definiert, wobei zuerst die strengsten und anschließend weniger strenge Regeln angewandt werden. Z. B. lautet die strengste Regel für den Bereich Anwendung: Zwei Werkstoffe sind ähnlich, wenn sie beide für drei gleiche Anwendungen eingesetzt werden. In WING-SIM erfolgte die Realisierung von vier Bereichen mit jeweils drei bis vier Regeln. Der Benutzer kann einen Werkstoff eingeben und vor der Ähnlichkeitssuche den für ihn relevanten Kontext als Kriterium für Ähnlichkeit auswählen.

Bei informellen Tests stellte sich heraus, daß die Ergebnisse nicht befriedigend sind. Dies liegt am Kern des Verfahrens, den heuristisch formulierten Regeln, die Ähnlichkeit nicht gut genug beschreiben. Durch die Zusammenarbeit mit Experten könnte versucht werden, die Regelbasis zu erweitern und inhaltlich motivierte Regeln zu finden. Jedoch zeigte sich bei Interviews, daß es den Experten äußerst schwer fällt, allgemeine Regeln zu formulieren oder ihre intuitiv getroffenen Entscheidungen nachträglich inhaltlich zu begründen. Weiterhin wird es bei größeren Regelapparaten äußerst schwierig, ein widerspruchsfreies System zu gewährleisten.

Dieses Phänomen findet man auch in anderen Wissensbereichen, v. a. bei Tätigkeiten, die man schnell und mit expertenhafter Sicherheit ausführt, wie etwa die Erkennung komplexer Muster. Solche Vorgänge lassen sich laut Smolensky (1988) grundsätzlich nicht durch regelfolgendes Schließen modellieren, sondern erfordern einen „*intuitive processor*“ (Smolensky 1988:5). Dementsprechend wurde im Projekt WING nach einer anderen Technik zur Wissensverarbeitung gesucht.

8.2.2 Cluster-WING: statistischer Ansatz

Die statistische Ähnlichkeitsermittlung von Werkstoffen basiert - im Gegensatz zu heuristischen Regeln und Expertenurteilen wie bei WING-SIM (cf. Mandl/Roppel 1992) - auf exakt ermittelten Daten. Die Werkstoffe können anhand vorhandener Daten zu charakteristischen Eigenschaften entweder mit Hilfe

spezieller Ähnlichkeitsmaße in eine Rangfolge, die die Ähnlichkeit mehrerer Werkstoffe zu einem gegebenen Werkstoff ausdrückt, gebracht oder durch die Anwendung clusteranalytischer Verfahren in Klassen eingeteilt werden.

8.2.2.1 Wirkungsweise der Clusteranalyse

Statistische Klassifikationen erfüllen lt. Eckes/Roßbach 1980:15 drei Funktionen:

- Sie sind *objektivierend*, da die Entscheidungskriterien der Clusterbildung nachvollziehbar und somit kritisierbar sind. Diese Nachvollziehbarkeit ist bei Expertenurteilen nicht gegeben, da diese aufgrund des Vorwissens und der Erfahrung des Beurteilenden intuitiv erfolgen.
- Clusteranalytische Verfahren haben eine *datenreduzierende* Funktion, da sie Objekte, die durch eine komplexe Datenfülle charakterisiert sind, in überschaubare Beziehungen bringen.
- Eine weitere Funktion von Klassifikationsverfahren sind deren *hypothesengenerierende* Eigenschaften. Durch die Kenntnis der Ähnlichkeitsbeziehung zweier Objekte können Rückschlüsse auf übergeordnete Strukturen oder Zusammenhänge ermöglicht oder gefördert werden.

Eine grundlegende Forderung an Objekte innerhalb einer Klasse (bei statistischen Verfahren) besteht darin, daß sie sich möglichst ähnlich sein sollen und möglichst unähnlich zu Objekten anderer Klassen. Weiterhin wird gefordert, daß sich jedes Objekt nur in einer Klasse befinden soll. Beide Prämissen gelten aber nicht für alle Verfahren - es gibt zum einen sogenannte „Clumping“-Methoden, bei denen sich Objekte auch in mehreren Klassen gleichzeitig befinden können, zum anderen werden bei Methoden des Fuzzy Clustering Zugehörigkeitswerte von Objekten zu Klassen berechnet. Des weiteren soll jedes Objekt der Ausgangsmenge die Zuweisung zu einer Klasse erhalten.

8.2.2.2 Einzelentscheidungen bei der statistischen Ähnlichkeitsermittlung

Im folgenden wird die statistische Ermittlung von Werkstoffähnlichkeiten im Zuge des allgemeinen Vorgehens bei Clusterverfahren erläutert:

- Auswahl der Objekte
Als Grundlage für die statistische Ähnlichkeitsermittlung dienen die in der MTU-Datenbank enthaltenen Werkstoffspezifikationen, da hiervon exakt ermittelte Daten vorhanden sind.

- Bestimmung charakteristischer Merkmale der Objekte
In der MTU-Datenbank sind zweierlei Typen von Daten enthalten - zum einen liegen dichotome Daten vor allem zu den Sprechbezeichnungen der Werkstoffe vor (z.B. Anwendungsgebiete, Schweißbarkeit, Lötbarkeit, usw.), zum anderen gibt es metrische Daten zu den einzelnen Kenngrößen der Werkstoffspezifikationen (z. B. E-Modul, Zugfestigkeit, Längenausdehnungskoeffizient, Zeitstandfestigkeit, Dichte usw.). Im Interesse des Klassifikationsergebnisses sollen die Klassifikationsmerkmale aber vom gleichen Typ sein - also nur dichotom oder nur metrisch (cf. Vogel 1975:51). Bei dem Prototypen WING-SIM (cf. Mandl/Roppel 1992) und dem Ähnlichkeitswerkzeug NEURO-WING (cf. Mandl 1995) wurden größtenteils binäre Daten der Sprechbezeichnungen gewählt. Bei der statistischen Ähnlichkeitsermittlung wurde dagegen angenommen, daß die spezifisch ermittelten metrischen Kennwertdaten die Werkstoffe hinreichend charakterisieren.

- Optional: Überlegungen zur Verfahrensweise bei fehlenden Werten
Idealerweise liegen für eine Clusteranalyse die Daten der charakteristischen Merkmale der Objekte vollständig vor. Diesem Wunschbild kann aber in der Realität nicht immer entsprochen werden, so sind z. B. in der MTU-Datenbank kaum lückenlose Datenreihen vorhanden. Bei der Anwendung von Klassifikationsverfahren sind jedoch vollständige Datensätze für die Berechnung der wechselseitigen Distanzen zwischen den Objekten nötig. Diesem Problem kann man beikommen, indem man fehlende Werte z. B. mittels statistischer Schätzverfahren errechnet. Dedler (1993:5) führt hierfür drei Techniken an:
 - Eliminierung: hier werden ganze Datensätze von der Klassifikation ausgenommen.
 - Ersetzungstechniken: Fehlende Werte könnten durch den Median oder das arithmetische Mittel eines Datensatzes ersetzt werden, was aber zu Verfälschungen im Datensatz führt.
 - Parameterschätzung: hier werden fehlende Werte durch die vorhandenen Werte geschätzt.Das günstigste Verfahren bei den vorliegenden Werkstoffdaten ist, fehlende Werte zu schätzen und Datensätze mit zuviel fehlenden Werten aus der Klassifikationsprozedur auszuschließen. Das zu verwendende Schätzverfahren hängt vom zugrundegelegten Modell ab. Bei fehlenden E-Modulwerten hat sich beispielsweise eine quadratische Spline-Interpolation als günstig herausgestellt. Die Auswahl des richtigen Modells für die Parameterschätzung ist wichtig, da dies das Klassifikationsergebnis entscheidend

beeinflussen kann, was auch nachgewiesen werden konnte (cf. Ludwig 1994b).

- **Optional: Standardisierung der Werte**
Bei manchen Ähnlichkeitsmaßen werden Einzeldistanzen zwischen Objekten quadriert, was zu einer internen Gewichtung dieser Distanzen proportional zu deren Größe führt. Ist dieser Effekt nicht wünschenswert, so empfiehlt sich, die Werte zu standardisieren. Bei den vorliegenden Werkstoffdaten wird jedoch die interne Gewichtung beibehalten.
- **Bestimmung eines geeigneten Ähnlichkeitsmaßes**
Die Auswahl eines Ähnlichkeitsmaßes wird eingegrenzt durch folgende Überlegungen: Soll die interne Gewichtung der Daten erhalten bleiben? Sind die Daten normalverteilt? Welches Klassifikationsverfahren soll angewendet werden?
Die Entscheidung bei der Bestimmung von Werkstoffähnlichkeiten fällt zugunsten der *quadrierten euklidischen Distanz* aus, bei der die interne Gewichtung der Daten beibehalten wird. Dabei wird von einer annähernden Normalverteilung ausgegangen. Außerdem lassen sich manche Klassifikationsverfahren nur mit diesem Maß sinnvoll interpretieren.
- **Optional: Gewichtung von Merkmalen**
Manchmal werden unterschiedliche Gewichtungen an die erhobenen Merkmale vergeben, was aber bei den MTU-Kennwertgrößen unterlassen wird. Es gibt außerdem die Möglichkeit, interne Gewichtungen auszugleichen, damit Ausreißer nicht so stark ins Gewicht fallen. Bei den Werkstoffdaten empfiehlt sich, die natürliche Gewichtung der Merkmale beizubehalten. Ausreißer können dann ausgeschlossen werden, und es ist möglich, mit den übrigen Objekten eine erneute Clusteranalyse durchzuführen.
- **Auswahl des Clusterverfahrens**
Hier muß zunächst die Auswahl einer Gruppe von Clusterverfahren erfolgen. Eine gebräuchliche Einteilung ist z. B. die Unterscheidung von graphentheoretischen, hierarchischen, partitionierenden und Optimierungsverfahren. Die verbreitetsten Methoden sind dabei die partitionierenden und vor allem die hierarchischen Verfahren. Bei den hierarchischen Verfahren findet eine weitere Unterscheidung zwischen hierarchisch-divisiven und hierarchisch-agglomerativen Techniken statt, wobei letztere die größte Verbreitung gefunden haben. Diese Methoden fusionieren sukzessive diejenigen Objekte, die die geringste Distanz zueinander aufweisen.
Bei den MTU-Daten wurde entschieden, mehrere agglomerative Verfahren

zu verwenden, und anschließend die am besten geeignete Methode zu finden. Dieses Vorgehen ist nicht unüblich, da die Datenstruktur nicht *a priori* auf ein spezielles Verfahren schließen läßt.

- **Bestimmung der geeigneten Klassenanzahl**
Die Bestimmung einer geeigneten Klassenanzahl, d. h. die Angabe, wieviele Klassen gebildet werden sollen, stellt ein grundlegendes Problem dar. Bei nur 25 Objekten müßten z. B. $1 + 2 + 3 + \dots + 25 = 325$ Klassifikationsergebnisse betrachtet werden, um die geeignetste Klassenanzahl herauszufinden. In der Regel behilft man sich deshalb suboptimal durch das Zeichnen eines Struktogramms, das auf der Abszisse die einzelnen Fusionschritte und auf der Ordinate die Klassifikationsbewertung abbildet. Wenn in der Verlaufskurve ein deutlicher Knick zu sehen ist, so spricht dies dafür, daß an dieser Stelle ein relativ heterogenes Cluster gebildet wurde. Die Klassenanzahl an der Knickstelle sollte deshalb nicht verwendet werden, sondern irgendeine höhere Anzahl.
Bei den Werkstoffspezifikationen wurden die Objekte, die zu einer Kenngröße vorlagen, durch 4 geteilt, was dann die Klassenanzahl ergab. Von MTU-Werkstoffexperten wurde eine Klassengröße von 4 als günstig angesehen, wobei die sich ergebende Klassenanzahl aber durch das Zeichnen von Struktogrammen ergänzend überprüft wurde.
- **Optional: Evaluierung der Ergebnisse**
Es gibt unterschiedliche Verfahren Klassifikationsergebnisse zu evaluieren. Statistische Möglichkeiten sind z. B. die Anwendung von Verzerrungsmaßen, die - mit erheblichem Aufwand - den Informationsverlust bei der Gruppenaufteilung berechnen. Andere statistische Verfahren ermitteln die stabilste Clusterlösung. Eine dieser Techniken wurde auch zur Evaluierung der Werkstoffklassen verwendet, nämlich das Maß *Pairbonds*, das von Boorman/Arabie (1972) entwickelt wurde. Dabei erwies sich das Verfahren *Complete Linkage* als relativ stabil, *Single Linkage* als relativ unstabil bei der Clusterermittlung auf der Grundlage der E-Modul-Daten.

8.2.2.3 Fazit: Cluster-WING

Bei clusteranalytischen Verfahren sind - wie gezeigt - eine Reihe von Einzelentscheidungen zu treffen, die aber wegen der großen Vielzahl nicht alle evaluiert werden können. Die gefundene Verfahrens-Kombination kann also höchstens zufällig die optimalste sein.

Bei den vorliegenden MTU-Werkstoffdaten ergibt sich ein Paradoxon: zum einen will man ähnliche Werkstoffe finden, um Lücken im Datenbestand zu überbrücken, andererseits bereiten gerade diese Lücken große Probleme bei der Clusteranalyse, weshalb man Methoden zur Parameterschätzung heranziehen muß.

Ein weiteres Problem betrifft die Übertragung der getroffenen Einzelentscheidung der Stichprobe auf die Gesamtheit der Werkstoffe, bei der jede Alternative neu überdacht werden muß, da sich zwangsläufig Verschiebungen ergeben.

Als günstig könnte sich eine Kombination von clusteranalytischen Verfahren und neuronalen Netzen erweisen, bei der die Ergebnisse der Clusteranalyse - bei einer gewissen Anzahl von Werkstoffen - als Trainingsdaten für ein neuronales Netz Verwendung finden.

8.2.3 NEURO-WING: Neuronaler-Netzwerk-Ansatz

Aufgrund der Probleme bei WING-SIM wurde - zeitgleich mit Cluster-WING - ein weiteres kognitionswissenschaftlich motiviertes System zur Erkennung von Werkstoffähnlichkeiten entwickelt. NEURO-WING basiert auf künstlichen neuronalen Netzen, die sich zur Modellierung von intuitivem Wissen, das sich nicht in Regeln fassen läßt, gut eignen. Einen empfehlenswerten Überblick über neuronale Netze bietet Zell (1994).

8.2.3.1 Funktionsweise neuronaler Netze

Neuronale Netze beruhen auf dem Vorbild des menschlichen Gehirns und sind lernfähig. Sie bestehen aus zahlreichen kleinen Prozessoren oder Neuronen, zwischen denen gewichtete Verbindungen (connections) liegen. Jedes Neuron ist in der Lage, über die Verbindungen Impulse an andere Neuronen zu senden. Neuronen, die Impulse empfangen, können daraus nach einer definierten Rechenvorschrift wiederum ihre eigene Aktivierung berechnen und diese erneut weitergeben. Durch eine große Anzahl derartiger lokaler Vorgänge entsteht ein Gesamtverhalten, das interpretiert werden kann. Inzwischen gibt es zahlreiche verschiedene Modelle, die sich in Aufbau und Funktionsweise unterscheiden. Im folgenden wird das sehr häufig eingesetzte Backpropagation-Modell beschrieben, das auch den Kern von NEURO-WING bildet, und das Musterassoziationen aufgrund von Beispielen lernt.

Beim Backpropagation-Netz sind die Neuronen in Schichten organisiert, wobei es eine Input-, eine Output- und meist ein oder zwei Zwischenschichten gibt. Dabei ist grundsätzlich jedes Neuron einer Schicht mit jedem Neuron der nachfolgenden Schicht verbunden. Bei der Musterverarbeitung wird ein Eingangsvektor an der Eingangsschicht angelegt, d. h. die Neuronen der ersten Schicht werden mit den Werten des Vektors aktiviert. Dann durchläuft die Aktivierung das Netz in Richtung Ausgangsschicht. Dort wird nach dem letzten Verarbeitungsschritt die Aktivierung der Neuronen abgelesen und ergibt den Outputvektor.

Diese Abbildung von einem Vektor auf einen anderen bildet die Grundlage für den Lernprozeß. Dazu werden zunächst Beispiele für die gewünschte Abbildung benötigt. Dies ist im vorliegenden Anwendungsfall kein Problem, da die Werkstoffexperten Beispiele für ähnliche Werkstoffe geben können, nicht aber die dahinterliegenden Gesetzmäßigkeiten erklären können. Das Netz lernt nun diese Abbildung und kann sie auf unbekannte Muster generalisieren.

Beim Lernvorgang wird ein Beispielmuster angelegt und ein Outputvektor errechnet, der dann mit dem gewünschten Vektor verglichen wird. Durch ein mathematisches Verfahren wird die Differenz zwischen dem target und dem tatsächlichen Output ins Netz zurückgespeist, und dabei verstellen sich die Gewichtungen der Verbindungen derart, daß der nächste Output näher am target liegt. In den Verbindungen liegt also das formbare Gedächtnis des Netzes.

Nachdem viele Muster dem Netz sehr oft gezeigt wurden, konvergiert es typischerweise und kann alle Muster der Trainingsmenge mit nur geringem Fehler reproduzieren. Nun muß anhand dem Netz noch nicht vorgelegter Paare von Vektoren überprüft werden, inwieweit das Netz generalisiert, also auch unbekannte Muster richtig abbildet. Dieses Verfahren entspricht der Annäherung einer Funktion durch einzelne Punkte in einem vieldimensionalen Raum.

Die Qualität der Abbildung kann nicht durch massives Training beliebig verbessert werden. Im Verlaufe des Trainings stellt sich typischerweise ein overfitting-Effekt ein, d. h. das Netz verringert zwar weiterhin den Fehler in der Trainingsmenge, der Fehler in einer unbekanntem Testmenge steigt aber wieder. Aufgrund dieses Problems muß das Training ständig überwacht und in einem Zustand der maximalen Generalisierung beendet werden.

8.2.3.1 Entwicklung von NEURO-WING

Für NEURO-WING wurde Ähnlichkeit in dem Kontext 'Anwendungen eines Werkstoffs' definiert. Werkstoffe sind demnach ähnlich, wenn sie für ähnliche Anwendungen eingesetzt werden. Es standen die Daten von 72 Werkstoffen zur Verfügung, wobei 48 als Trainings- und 24 als Testmenge benutzt wurden. Als Input-Vektor dienen 22 wichtige Kennwerte des Werkstoffs und als gewünschter Output ein binärer Vektor über die Eignung des Werkstoffes für 22 Anwendungen. Dabei bedeutet 1, der Werkstoff wird eingesetzt, und 0, der Werkstoff wird nicht eingesetzt. Ziel war es also, ein Netz zu trainieren, das von den Kennwerten eines Werkstoffes dessen Anwendungsprofil errechnet. Die tatsächlichen Anwendungen stellen das Expertenwissen dar, das vom Netz abgebildet wird.

Für die Bewertung des Netzes wurde der Output mit den binären Vorgaben verglichen, wobei alle Ausgaben kleiner 0,1 als Treffer für 0 und alle Werte größer 0,3 als Treffer für 1 gewertet wurden. Dies erwies sich als sinnvoll, da die Grundmenge wesentlich mehr 0 als 1 enthielt und das Netz als Ausgabe öfter einen Wert nahe 0 lieferte.

Bei der Entwicklung eines Backpropagation-Netzes müssen zahlreiche Parameter optimiert werden, die größtenteils nur heuristisch und empirisch ermittelt werden können. Darunter fallen etwa die Zahl der Zwischenschichten, die Anzahl der Neuronen in diesen, die Zahl der Verbindungen, die Aktivierungsfunktion der Neuronen, die Lernfunktion und die Lernregel, wobei die Funktionen häufig noch zusätzliche Parameter besitzen. In zahlreichen Versuchen wurde ein optimales Ergebnis nach 450 Lernepochen bei einem Netz mit acht versteckten Neuronen in einer Schicht erzielt, wobei die Verbindungen anfangs zufallsgesteuert auf Werte im Intervall $[-0,3 / 0,3]$ initialisiert wurden. Das Netz konnte 48% aller 1er und 84% aller 0er richtig erkennen.

Die Trefferquote bei 0 konnte durch eine inhaltlich motivierte Modifikation der Netzarchitektur um 5% erhöht werden. Eine Stelle im Input-Vektor repräsentiert die Gruppe, zu der ein Werkstoff gehört, und ist damit besonders wichtig für die Ähnlichkeiten, da man davon ausgehen kann, daß Ähnlichkeiten hauptsächlich innerhalb von Gruppen auftreten. Um dieses Wissen auszunutzen, wurde dem entsprechenden Neuron in der Eingangsschicht durch eine sogenannte shortcut connection, wie sie Le Cun (1989) vorschlägt, größeres Gewicht eingeräumt. Das Neuron, das den wichtigen Parameter repräsentiert, erhält zusätzlich zu den schon vorhandenen Verbindungen zu den Neuronen der Zwi-

schenschicht noch Verbindungen die direkt zur Output-Schicht laufen. Eine ausführlichere Beschreibung der Netzarchitektur bietet Mandl (1994).

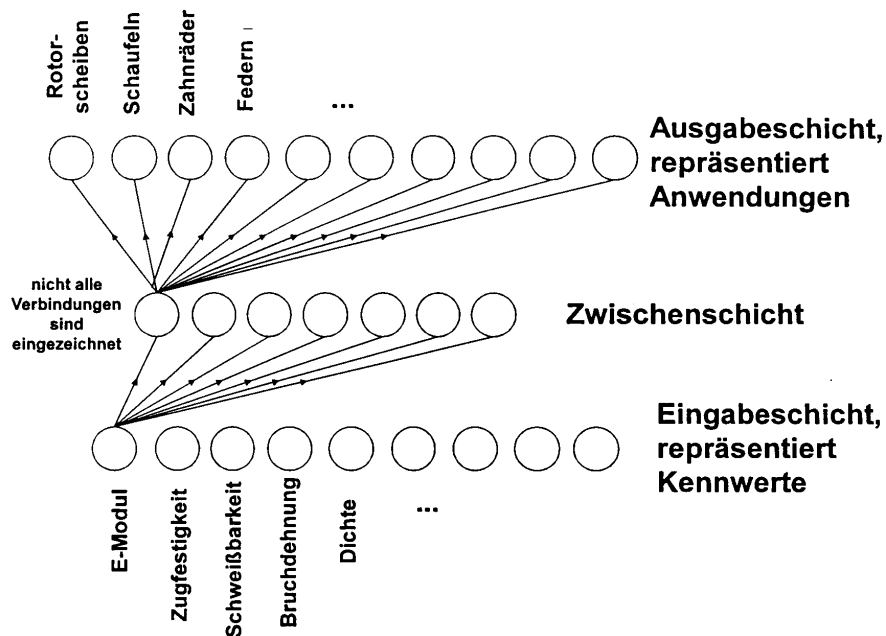


Abb. 8.1: Schematische Darstellung der verwendeten Netzwerkarchitektur

Durch Ausnutzen der Gruppenbeziehungen zwischen den Anwendungen bei der Repräsentation konnten ebenfalls Verbesserungen erzielt werden, jedoch ist diese Repräsentation weniger mächtig und kann einige Anwendungsprofile nicht ausdrücken.

8.2.3.3 Verbesserung durch Expertenfeedback

Die Fehler des Netzes wurden einer genaueren inhaltlichen Analyse unterzogen. Dabei stellte sich heraus, daß einige Fehler aufgrund der Datenlage durchaus nachvollziehbar waren, d. h. die Kennwerte reichten für die Differenzierung nicht aus. Um dem zu begegnen, nannten Experten zwei weitere wichtige Kennwerte, deren Aufnahme in die Input-Vektoren erfolgte. Weiterhin wurden zwei Anwendungen gestrichen, die im Korpus kaum vorkommen und viele Fehler produziert hatten.

Dieses Expertenwissen konnte durch Änderungen an den Daten und der Netzarchitektur also leicht integriert werden. Bei weiteren Versuchen ergaben sich wesentlich höhere Trefferraten bei einer Anordnung mit einer Zwischenschicht mit 14 Neuronen nach 1400 Lernepochen, wobei die oben erwähnte shortcut connection beibehalten wurde und sich weiterhin positiv auswirkte. Dieses Wissen stellt neben den Trainingsbeispielen eine zweite Art von Expertenwissen auf höherer Ebene dar, das durch Veränderungen an der Repräsentation und der Architektur in das Netzwerk integriert werden konnte. Die folgende Tabelle gibt einen Überblick über die Performanz der einzelnen Netze.

Tab. 8.1: Ergebnisse in Abhängigkeit von der Netzwerkarchitektur

| | Anwendungen (1er) | Nicht-Anwendungen (0er) |
|---------------------------------------|-------------------|-------------------------|
| Architektur mit einer Zwischenschicht | 48% | 84% |
| mit <i>shortcut-connections</i> | 48% | 89% |
| nach Integration von neuem Wissen | 79% | 91% |

Bei den verbleibenden Fehlern wurden Werkstoffe oft Anwendungen zugeordnet, für die sie zwar geeignet wären, aber nicht eingesetzt werden, weil sie dafür zu hochwertig und teuer sind. Solange die Ähnlichkeit über die Eignung definiert ist und in das Netz keine Preisangaben eingegeben werden, sind solche Fehler unvermeidlich. Insgesamt wurde die erreichte Leistungsfähigkeit des neuronalen Netzes von der MTU als positiv bewertet. Eine detaillierte Darstellung der Entwicklung von NEURO-WING findet sich in Mandl (1994).

8.2.4 Fazit: Modellierungstechniken

Da die drei Systeme teilweise auf verschiedenen Daten aufsetzen und verschiedene Facetten innerhalb der Ähnlichkeit von Werkstoffen bearbeiten, sind sie nicht im Detail auf der Ebene einzelner Werkstoffe vergleichbar. Zur Bewertung der verschiedenen Verfahren wird die jeweilige Bewertung des Gesamtergebnisses herangezogen. Dabei werden v. a. NEURO-WING und Cluster-WING betrachtet, da das regelbasierte WING-SIM insgesamt nur unbefriedigende Resultate erbrachte und auch theoretisch inadäquat zu sein scheint.

Es wird in der Literatur häufig herausgestellt, daß neuronale Netze und statistische Verfahren auf die gleichen mathematischen Wurzeln zurückgreifen (cf.

Ripley 1993). So erfordern beide Methoden in der praktischen Anwendung eine Vielzahl von heuristischen Entscheidungen über Parameter. Wie gezeigt wurde, ist bei neuronalen Netzen eine inhaltliche Entscheidung über das Netzdesign zumindest manchmal möglich. Weiterhin können bei einem Backpropagation-Verfahren die Auswirkungen der Parametersetzungen jederzeit an der Testmenge beobachtet werden. Insgesamt eignen sich neuronale Netze besser für die Integration von Expertenwissen.

NEURO-WING benutzt also Expertenwissen und Daten, während die anderen Systeme jeweils nur auf eine dieser Quellen zugreifen; WING-SIM nur auf Wissen und Cluster-WING nur auf Daten und erst bei der Evaluierung auf Expertenwissen. Dies bedeutet natürlich auch, daß bei neuronalen Netzen der Aufwand für die involvierten Experten höher ist.

Weiterhin kann von einem neuronalen Netz die Kontextabhängigkeit besser gelöst werden, da für jeden Kontext, im Extremfall für jeden Benutzer ein eigenes Netz trainiert werden kann.

Ein großer Nachteil von neuronalen Netzen liegt naturgemäß in ihrer mangelnden Erklärungsfähigkeit. Die Entscheidungen sind, wie ja auch menschliche Expertenurteile, kaum nachträglich rational zu begründen.

Bei der Entscheidung für eine Modellierungstechnik ist neben diesen theoretischen und für die Implementation relevanten Aspekten die Leistung des Verfahrens ausschlaggebend. Die Ergebnisse wurden in informellen Tests den Domänenexperten der MTU vorgelegt, die die Ähnlichkeitsurteile auf inhaltlicher Basis bewerteten. Da die Ergebnisse von NEURO-WING besser beurteilt wurden und bei der Entwicklung mit neuronalen Netzen einige Vorteile auftreten, wird bei der prototypischen und kommerziellen Version von WING ein neuronales Netz für das Ähnlichkeitswerkzeug eingesetzt.

8.3 Integration in die WING-M2-Benutzungsoberfläche

Im Werkstoffinformationssystem WING-M2 (cf. Kap. 5) ist ein Ähnlichkeitswerkzeug implementiert, das auf NEURO-WING basiert. Zwischen allen von NEURO-WING gelieferten Vektoren, von denen jeder ein Anwendungsprofil eines Werkstoffes repräsentiert, wurden mit einem Vektorähnlichkeitsmaß Ähnlichkeitswerte berechnet und in einer Matrix über alle Werkstoffe zusammengefaßt. Auf diese Werte greift das Werkzeug während des Dialogs zu.

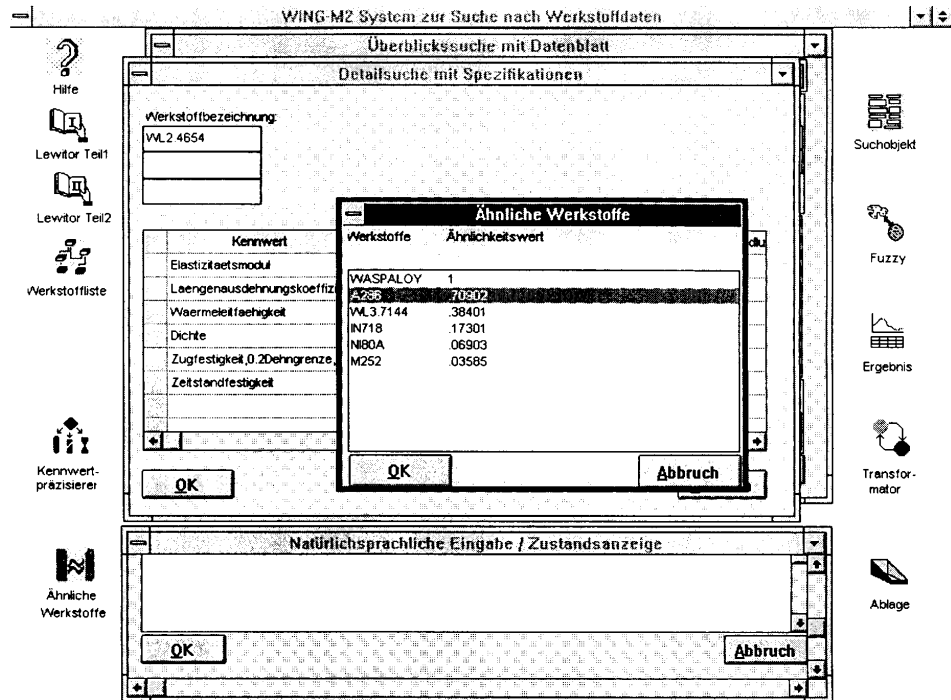


Abb. 8.2: WING-M2 mit geöffnetem Ähnlichkeitswerkzeug

Das Ähnlichkeitswerkzeug ist als eigenständiges Werkzeug im Sinne des WOB-Modells (cf. Kap. 4) realisiert. Es reagiert auf die in Abschnitt 8.2 geschilderte Situation einer Nullantwort. Stößt der Benutzer auf eine Datenlücke, so weist ihn das Werkzeug darauf hin und bietet ihm an, den gleichen Kennwert zu einem ähnlichen Werkstoff zu recherchieren. Will der Benutzer mit dieser Strategie weiterarbeiten, erhält er eine sortierte Liste von ähnlichen Werkstoffen mit Angabe ihres Ähnlichkeitsgrades. Daraus kann er einen in seinem Kontext ähnlichen Werkstoff auswählen und die Suche damit fortsetzen.

Das Ähnlichkeitswerkzeug ist aber nicht auf diese spezielle Retrievalsituation beschränkt. Will es ein Benutzer unabhängig davon für einen assoziativen Suchschritt einsetzen, so verhält es sich analog und liefert eine Liste ähnlicher Werkstoffe in jedem Fall, in dem der Fokus klar auf einem Werkstoff liegt. Der Benutzer kann also auch in der Werkstoffliste und in der Überblickssuche mit Datenblatt eine Suche nach ähnlichen Werkstoffen auslösen. Details der Implementierung finden sich in Mandl (1995).

Das Ähnlichkeitswerkzeug gliedert sich harmonisch in das Konzept des WOB-Modells ein und kann auch für andere WOB-Informationssysteme zur iterativen Fortsetzung des Retrievals entwickelt werden. Im Rahmen des WOB-Modells stellt sich dann die Frage nach der Parametrisierbarkeit des Werkzeugs. Es könnten dazu neben der Anwendung andere Ähnlichkeitskontexte implementiert werden. Der Benutzer kann dann einen Kontext auswählen oder dieser wird durch Dialogbeobachtung systemseitig sinnvoll gesetzt. Die Kontexte könnten für jeden einzelnen Benutzer durch weitere Wissensmodellierung erstellt werden. Als Input für eine derartige Weiterentwicklung bietet sich die Wahl des Benutzers in der systemseitig vorgeschlagenen Liste ähnlicher Werkstoffe an. Durch die Auswahl weist der Benutzer dem Werkstoff implizit eine hohe Ähnlichkeit für seinen Kontext zu.

8.4 Ausblick

In der weiteren Entwicklung wird eine stärkere Zusammenarbeit der intelligenten Komponenten von WOB-Informationssystemen angestrebt. Im Kontext von WING bietet sich insbesondere eine Kooperation des Ähnlichkeitstools mit FUZZY-WING (cf. Kap. 9) an. Beide Werkzeuge repräsentieren vages Wissen, jedoch auf verschiedenen Ebenen. NEURO-WING beschreibt vage Beziehungen zwischen Werkstoffen (im Gegensatz zu starren Beziehungen, wie etwa der hierarchischen Typisierung), während FUZZY-WING vages Wissen über Eigenschaften meist in Form von numerischen Werten modelliert. Die Zusammenhänge liegen auf der Hand. So können Werkstoffe, die in einem bestimmten Kontext in einer vagen Attributierung übereinstimmen, in diesem Kontext ähnlich genannt werden. Beide Werkzeuge können also von einer einheitlichen Wissensbasis profitieren, wobei sie gerade Wissen über Kontexte austauschen können. Dazu müssen Formalismen gefunden werden, die den Zugriff von Fuzzy Logic und neuronalen Netze auf eine gemeinsame Wissensbasis erlauben.