

Graphisches Faktenretrieval mit vager Anfrageinterpretation

Christian Wolff

Christa Womser-Hacker

Institut für Informatik
Universität Leipzig

Informationswissenschaft
Universität Regensburg

Zusammenfassung

Zwei neue Techniken im Bereich der Faktenrecherche werden vorgestellt: Ein *fuzzy logic*-basierter Interpretationsmechanismus für Aufbau und Übersetzung von Anfragen, die vage Konzepte (linguistische Variablen) enthalten und ein graphisches Recherchewerkzeug für numerische Daten, das dem Benutzer die direkte Manipulation von Liniendiagrammen erlaubt. Dabei kann der Benutzer entweder auf bereits visualisierte Daten im Diagramm Bezug nehmen oder mit Hilfe graphischer Primitive im Diagramm eine völlig neue Anfrage aufbauen. Beide Techniken werden kombiniert, um eine flexible Anfrageinterpretation durch das System zu ermöglichen, die auch Domänencharakteristika (hier: Werkstoffinformation) berücksichtigt. Die vage Interpretation graphischer Anfragen dient als Schnittstelle zu einem Standard-RDBMS, in dem die Daten verwaltet werden.

Abstract

In this paper the combination of two new techniques in fact retrieval is discussed: A fuzzy logic based interpretation mechanism allowing the construction and translation of queries containing vague concepts (linguistic variables) and a graphical retrieval tool for factual data in which the user can specify queries by directly manipulating line charts of data sets in the database or drawing typical search areas in the coordinate system from scratch. The combination of both techniques results in a flexible and dynamic interpretation of graphical query constructs and optimizes the interface between the visual representation of factual information and its formal representation needed to query a standard RDBMS.

1 Einführung

Wir beschreiben die Verknüpfung zweier unterschiedlicher Techniken zur Behandlung vager Anfragen an Faktendatenbanken: Die direkte Manipulation visueller Anfragekonstrukte und die unscharfe Interpretation formaler Anfragen, die aus natürlichsprachlichen oder visuellen Anfragedefinitionen stammen. Indirekt soll dadurch auch deutlich werden, daß im Information Retrieval (IR) seit langem etablierte Konzepte wie *partial match* und qualitätsorientierte Ausgabeaufbereitung (*ranking*) im Kontext der Faktenrecherche von Bedeutung sein können.

Die jeder Abbildung von Benutzerinteressen auf formale Anfragekonstrukte inhärente Unschärfe ist ein zentrales Thema der Information Retrieval Forschung (vgl. SALTON 1989). Diese Unsicherheit resultiert aus der für unstrukturierte Daten, wie sie im Textretrieval

vorliegen, nicht eindeutig zu definierenden Beziehung zwischen der internen Wissensrepräsentation eines Information Retrieval-Systems (IRS) einerseits und den Anfragen andererseits. Daran macht man traditionell den wesentlichen Unterschied zwischen Text Retrieval (*partial match paradigm*) und Faktenretrieval (*exact match paradigm*) fest (vgl. VAN RIJSBERGEN 1979:2ff., FRAKES 1993:9).

Die Qualität einer Datenbank bestimmte sich lange nach der Exaktheit und Vollständigkeit der in der Datenbank enthaltenen Information. Es liegt auf der Hand, daß nur wenige Wissenschaftsgebiete vollständige und exakte Daten liefern können, mithin sind Lücken und Inkonsistenzen im Datenbestand ein typisches Phänomen realistischer Datenbanken. Um einen solchen Datenbestand adäquat erschließen zu können, hat man versucht, gängige Datenmanipulationssprachen (*structured query language*, SQL) um Komponenten zu erweitern, die auf Techniken zur Einbeziehung von Unsicherheit wie *fuzzy logic* (vgl. BOSC & PIVERT 1992) oder Wahrscheinlichkeitstheorie (vgl. FUHR & HOFFMANN 1991) basieren. Nichtsdestoweniger werden Information Retrieval und Datenbanktechnologie nach wie vor als Forschungsdomänen mit unterschiedlichen theoretischen Ansätzen betrachtet, was die Repräsentation und den Abgleich von Anfragen und Daten angeht. Die Integration der Behandlung von Vagheiten ist ein Gebiet, auf dem die Datenbanktechnologie von einer langen Tradition in der IR-Forschung profitieren kann, da dort *best* oder *partial match*-Ansätze den wichtigsten theoretischen Modellen zugrunde liegen (etwa bei vektorialen und probabilistischen Ansätzen im IR, vgl. SALTON 1989, FUHR 1989).

Die beiden hier beschriebenen Werkzeuge, FUZZY-WING und WING-GRAPH wurden im Rahmen des Projektes WING-IIR (Werkstoffinformationssystem mit graphischer/natürlichsprachlicher Benutzeroberfläche und Intelligentes Information Retrieval) an der Universität Regensburg entwickelt (vgl. KRAUSE & WOMSER-HACKER 1997). Ursprünglich auf einen Vergleich und eine Mischung graphischer und natürlichsprachlicher Benutzerschnittstellen ausgelegt, entstand in diesem Projekt ein multimodales Werkstoffinformationssystem mit einer objekt-orientierten Benutzerschnittstelle und integrierten intelligenten Komponenten. Dabei wurden unterschiedliche Systemprototypen im rapid prototyping Verfahren und mit intensiver empirischer Auswertung durch Benutzertests erstellt (vgl. KRAUSE et al. 1993/1994). Die dem System zugrundeliegende Datenbank ist eine strukturell vollständige Untermenge der Datenbank des Projektpartners MTU, eines zur Deutschen Aerospace (DASA) gehörenden Triebwerksproduzenten. Die Datenbank enthält vornehmlich Kennwertdaten aus Versuchen mit Hochtemperaturwerkstoffen (hochwarmfeste Stähle, Titanlegierungen) wie Elastizitätsmodul oder Kriechfestigkeit.

2 Unscharfe Anfragemodalitäten in Faktendatenbanken

Der bereits genannte Projektkontext von WING-IIR folgt einem empirischen Ansatz. Daraus motiviert sich die Einbeziehung vager Anfragemodalitäten. Unter einem vagen Anfragemodus ist ein Dialogmodus zu verstehen, bei dem die Anfragerepräsentation nicht direkt auf eine formale Datenbankabfrage abgebildet werden kann. Dies ist sowohl beim Einsatz sog. linguistischer Variablen (*gut*, *hoch* etc.) in natürlichsprachlichen Dialogsystemen und bei der Ähnlichkeitssuche als auch bei der Interpretation graphischer Anfragen der Fall.

2.1 Anwendung von fuzzy logic-Techniken

Die Anwendung von Techniken der *fuzzy logic* im Information Retrieval ist naheliegend, da *fuzzy logic* und IR auf eine gemeinsame theoretische Basis zurückgreifen: Unscharfe Mengen beruhen auf der Idee, daß es unterschiedliche Grade der Zugehörigkeit eines Elementes zu einer Menge geben kann. Dieser Zugehörigkeitsgrad läßt sich durch den Wert einer Zugehörigkeitsfunktion bestimmen. Unscharfe Mengen wurden etwa von RADECKI 1979, BUELL & KRAFT 1983 und MIYAMOTO 1990 für das Dokumentenretrieval theoretisch fundiert und operationalisiert.

Im Rahmen von WING-IIR wurde ein Werkzeug entwickelt, das es erlaubt, Datenbankrecherchen mit unscharfen Begriffen wie *hoch*, *niedrig*, *gut* etc. durchzuführen. Das System FUZZY-WING (vgl. KRAUSE, WOLFF & WOMSER-HACKER 1994, WOMSER-HACKER 1997) benutzt eine Wissensbasis mit domänenspezifischem Wissen und Inferenzregeln. Zusätzlich ist die Suche nach jeweils ähnlichen Elementen (Werkstoffen) in der Datenbasis erlaubt, wobei der dabei verwendete Ähnlichkeitsbegriff domänenspezifisch definiert ist (z.B. durch Eignung zweier Werkstoffe für die gleiche Anwendung bzw. für das gleiche Bauteil, vgl. MANDL & WOMSER-HACKER 1995). Der Benutzer erhält eine nach Anfrage-Inhaltsähnlichkeit sortierte Ausgabe, aus der er Werkstoffe selektieren kann. Durch Setzen von Grenzwerten kann er den Ausgabeumfang selbst einschränken. Eine ausführliche Diskussion der Benutzeroberfläche des Systems kann hier nicht geleistet werden, vgl. aber KRAUSE & WOMSER-HACKER 1997.

2.2 Graphisches Faktenretrieval

Während FUZZY-WING ursprünglich für die Interpretation vager natürlichsprachlicher Anfragen oder vager Ausdrücke in graphischen Abfrageformularen entwickelt wurde, widmet sich ein weiteres Werkzeug im Kontext von WING-IIR der unmittelbaren Interpretation visueller Anfragen für das Faktenretrieval.

Die graphische Aufbereitung von Fakten hat eine lange Tradition in dem Bemühen, die Bedeutung von Information verständlich zu machen (vgl. KOSSLYN 1989, TUFTE 1983, 1990, 1997). Standard-Softwarewerkzeuge bieten eine breite Palette von Darstellungsformaten für die Informationsvisualisierung (Linien-, Balken-, Tortendiagramme etc.), mit denen sich die Ergebnisse von Datenbankrecherchen unmittelbar darstellen lassen. Benützt man solche Werkzeuge zur Nachbereitung von Anfrageergebnissen, so ergibt sich eine lineare Dialogstruktur, die für je ein Anfragebedürfnis mit der visuellen Aufbereitung abgeschlossen ist. Dies bedeutet, daß ungeachtet der auch im Faktenretrieval zyklischen Natur des Dialogprozesses keine Fortsetzung der Anfrage im adäquaten Darstellungsmedium möglich ist: Der Benutzer fragt die Werkstoffdatenbank anhand numerischer Kriterien (hier: Werkstoffeigenschaften) ab, erhält ein tabellarisch aufbereitetes Ergebnis und visualisiert dieses als Diagramm, um es interpretieren zu können. Ergibt sich aus der Chartanalyse eine Folgefrage, ist ein Moduswechsel nötig, d.h. wiederum muß ein SQL-Statement aufgebaut oder eine alphanumerische Abfragemaske ausgefüllt werden.

An diesem Punkt setzt das graphische Retrievalwerkzeug WING-GRAPH an, das es dem Benutzer erlaubt, die visuelle Datendarstellung selbst zu manipulieren und als Anfrage zu

Graphisches Faktenretrieval mit vager Anfrageinterpretation

verwenden. Dies hat zur Folge, daß nicht nur Anfrageergebnisse, sondern auch die Anfragen selbst im gewünschten visuellen Zielmedium erfolgen, bei WING-GRAPH als Linien- oder Balkendiagramm. Zur Differenzierung unterschiedlicher Anfragetypen stehen dem Benutzer sechs unterschiedliche graphische Anfragemodi zur Verfügung. Abb. 1 zeigt die Benutzeroberfläche von WING-GRAPH mit einer Beispielanfrage.

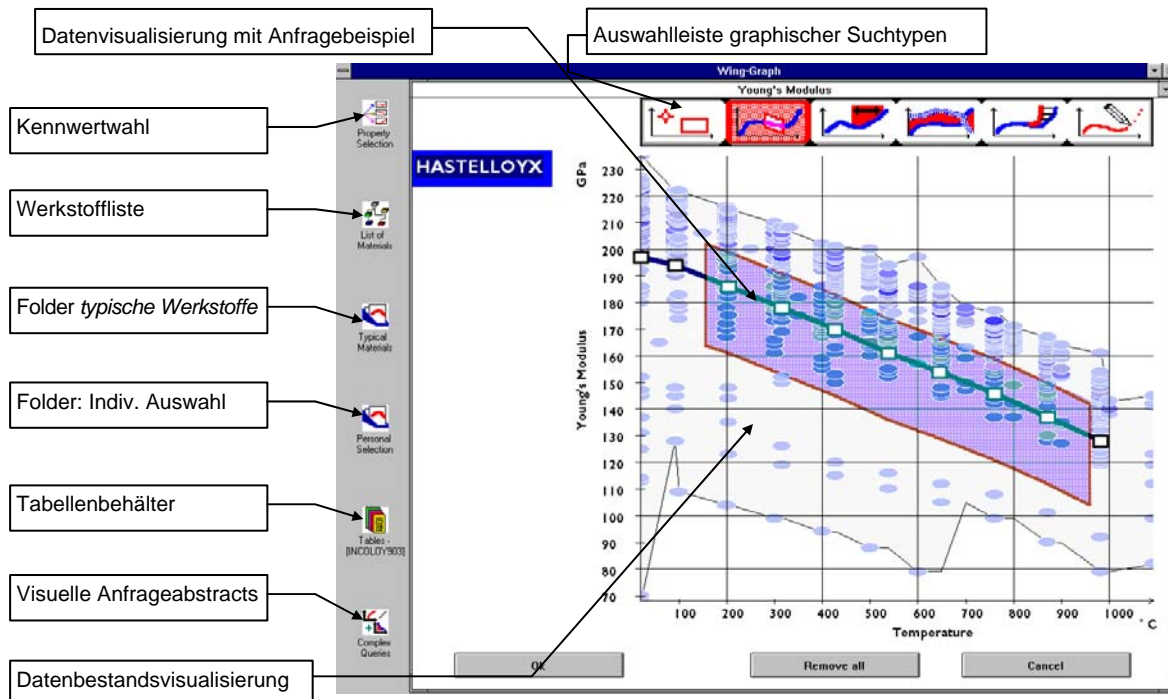


Abb. 1: WING-GRAPH - Interface mit Beispielanfrage (Streuband relativ zu Werkstoff HASTELLOYX)

2.2.1 Graphische Suchtypen als visuelle Sprache

Die in WING-GRAPH verwendeten Anfragetypen bauen auf dem kognitiven Modell der Graphenwahrnehmung und -interpretation von Stephen PINKER auf (vgl. PINKER 1990). Dieses kognitionspsychologische Modell wird für die Operationalisierung in WING-GRAPH an die Gegebenheiten des Mensch-Maschine-Dialogs angepaßt, insbesondere um einen Handlungsaspekt erweitert, der die Modellierung eines Zyklus aus *Anfrage - Ergebnisinterpretation - Anfrage* erlaubt (vgl. WOLFF 1996:141ff). Aus den Parametern dieses Modells ergeben sich für die Interaktion mit Liniendiagrammen drei Grundtypen von Anfragen in Abhängigkeit von der Relation zwischen Vorlageinformation und Informationsbedürfnis des Benutzers: Abstrakte Suchtypen nehmen keinen Bezug auf Vorlageinformation, relative Suchtypen beziehen sich auf eine oder mehrere Werkstoffkurven oder auf die direkte Modifikation einer Vorlagekurve und bei dem produktiven Verfahren trägt der Benutzer eine Werkstoffkurve selbst als Suchspezifikation in das Diagramm ein (vgl. Tab. 1).

Die Suchtypen von WING-GRAPH konstituieren eine visuelle Sprache, die durch kombinierte Anwendung der elementaren Suchtypen eine flexible Recherche in der Datenbank ermöglichen und den Gedanken eines zyklischen Recherchedialogs im Zielmedium realisieren.

Abstrakte Suchtypen:

- Definition von Suchpunkten und -bereichen

Relative Suchtypen:

- Definition von Suchbereichen in Bezug zu einer Vorlagekurve (Streubereich)
- Besser-/Schlechter-Relation in Bezug zu einer Vorlagekurve
- Vergleich zweier Werkstoffkurven
- Modifikation einer Vorlagekurve

Produktives Verfahren:

- Benutzergenerierte Kurvenhypothese

Tab. 1: Suchtypen in WING-GRAPH

2.2.2 Abbildung visueller Anfragen in einer formalen Abfragesprache

Es ist offensichtlich, daß die Abbildung graphischer Anfragen auf eine formale Datenmanipulationssprache ein zentraler Aspekt bei der Realisierung des graphischen Retrieval ist. Die folgenden Formeln zeigen den logischen Zusammenhang zwischen einem graphischen Anfragekonstrukt und seiner Übersetzung in SQL am Beispiel eines Streubandes um eine Vorlagekurve. Das Ergebnis einer Anfrage ist eine Menge W von Werkstoffen, deren Meßwertreihen dann visualisiert werden. Aus jedem Anfragekonstrukt werden Bereichsgrenzen für die Punkte abgeleitet, an denen Meßwerte in der Datenbasis vorliegen können. Die nachfolgende Formalisierung zeigt die Übersetzungslogik für ein vom Benutzer um eine Vorlagekurve gezeichnetes Streuband, in dem die gesuchten Werkstoffkenndaten liegen sollen: Für jeden Stützpunkt werden alle Meßreihen selektiert, die die Bedingung des Streubereichs explizit erfüllen.

Ergebnis :

Menge W der Werkstoffbezeichner, denen ein geordneter

n – Paartupel $T(x_1 y_1, \dots, x_i y_i, \dots, x_n y_n)$ zugeordnet ist, für den gilt :

$\exists(x_j y_{lower_j} / y_{upper_j})$ für $dengilt : \exists(x_i y_i) \in T$ mit

$$\left(\begin{array}{l} (y_{lower_j} \leq y_i \leq y_{upper_j} \wedge x_i = x_j, x_i, x_j \in M_S) \\ \wedge (y_{lower_{j+1}} - y_{lower_j} = y_{upper_{j+1}} - y_{upper_j} = y_{j+1} - y_j) \\ \wedge (y_{upper_j} - y_{lower_j} = y_{upper_{j+1}} - y_{lower_{j+1}} = \delta_y) \end{array} \right) \wedge$$

$\exists(x_j y_{lower_j} / y_{upper_j})$ für $dengilt : \exists x_i y_i \in T$ mit

$$\left(\begin{array}{l} (y_{lower_j} \leq y_i \leq y_{upper_j} \wedge x_i = x_j, x_i, x_j \in M_S) \\ \wedge (y_{lower_{j+1}} - y_{lower_j} = y_{upper_{j+1}} - y_{upper_j} = y_{j+1} - y_j) \\ \wedge (y_{upper_j} - y_{lower_j} = y_{upper_{j+1}} - y_{lower_{j+1}} = \delta_y) \end{array} \right)$$

Graphisches Faktenretrieval mit vager Anfrageinterpretation

wobei

y_j Meßwert der Referenzkurve am Stützpunkt x_j

y_{upper_j}, y_{lower_j} untere / obere Suchgrenze am Stützpunkt x_j

δ_y Breite des Streubands,

$1 < j \leq k \leq m$, k = Stützpunktanzahl des Suchbereichs,

m = Stützpunktzahl der Referenzkurve

M_S = geordnete Menge der Meßstützpunkte, $x_l < x_{l+1}$, $x_l, x_{l+1} \in M_S$

Das Verfahren ermittelt alle Meßreihen, die *wenigstens einen* Meßwert innerhalb des Streubandes aufweisen, d.h. ein Werkstoff gelangt auch in die Ergebnismenge, wenn seine zugehörige Kennwertkurve nur einen Teil des Streubandes durchläuft. Umgekehrt werden aus der Ergebnismenge alle Datenreihen ausgefiltert, an wenigstens einem Stützpunkt außerhalb des Streubands liegen:

$$W = \bigcup_I^k W_j - \bigcup_I^k \bar{W}_j$$

W_j (\bar{W}_j) Menge der Werkstoffbezeichner, denen ein geordneter

n -Paartupel $T(x_1 y_1, \dots, x_i y_i, \dots, x_n y_n)$ zugeordnet ist,

der genau ein Wertepaar $x_i y_i$ aufweist, das die j te Suchbedingung

(*explizit nicht*) erfüllt

Die Implementierung dieser Interpretation visueller Anfragen erfordert die Auswertung der graphischen Anfragerrepräsentation und die Übersetzung nach SQL, wobei aufgrund der hohen Zahl einzelner Meßwerte komplexe SQL-Statements generiert werden. Das nachfolgende - einfache - Beispiel zeigt eine SQL-Repräsentation eines Streubandes:¹

```
SELECT DISTINCT SPZFK FROM KNW11 WHERE (1 = 1) AND
((
  (TMPRT = 100 AND EMOD BETWEEN 160 AND 169)
  OR (TMPRT = 200 AND EMOD BETWEEN 155 AND 164)
  OR (TMPRT = 300 AND EMOD BETWEEN 149 AND 158)
  OR (TMPRT = 400 AND EMOD BETWEEN 144 AND 153)
  OR (TMPRT = 500 AND EMOD BETWEEN 138 AND 147)
) AND SPZFK NOT IN
(SELECT SPZFK FROM KNW11 WHERE
  (TMPRT = 100 AND NOT EMOD BETWEEN 160 AND 169)
  OR (TMPRT = 200 AND NOT EMOD BETWEEN 155 AND 164)
  OR (TMPRT = 300 AND NOT EMOD BETWEEN 149 AND 158)
  OR (TMPRT = 400 AND NOT EMOD BETWEEN 144 AND 153)
  OR (TMPRT = 500 AND NOT EMOD BETWEEN 138 AND 147)
)) ORDER BY SPZFK;
```

Ein wesentliches Argument für den Einsatz eines graphisches Suchverfahrens ist die Beobachtung, daß einerseits dieser Modus der Recherche das Benutzerbedürfnis nach einer visuellen Interpretation erfüllt, andererseits dabei notwendigerweise Abfragen generiert werden, die so komplex sind, daß sie alphanumerisch (direkt in SQL, über eine Eingabemaske etc.) kaum zu konstruieren wären.

Eine Analyse des formalen Aufbaus der Anfragen zeigt, daß der hier gewählte Ansatz für komplexe Domänen wie die Werkstoffwissenschaft noch zu allgemein ist, da jeweils abso-

¹ Der Zusammenhang von Anfragekomplexität und Datenbankperformanz wird hier nicht näher diskutiert; es ist offensichtlich, daß bei sehr umfangreichen Meßwertreihen heuristische Verfahren zur Komplexitätsreduktion eingesetzt werden müssen, vgl. WOLFF 1996:267ff.

lute Interpretationsbandbreiten verwendet werden. In der Berechnung der Streubandbreite ist der Wert von δ_y zunächst ein fester prozentualer Wert (in Abhängigkeit vom in der Datenbank verfügbaren Wertespektrum), d.h. die feste Breite des Streubands wird in einen Suchbereich im SQL-Statement übersetzt (... WHERE ... EMOD BETWEEN UNTERE_GRENZE AND OBERE_GRENZE...), der für alle Einzelbedingungen den gleichen Absolutwert hat. Tatsächlich beeinflussen aber eine Reihe von Faktoren wie ökonomische Randbedingungen, gewünschte Werkzeuggruppe, Temperaturbereich oder das zu konstruierende Werkstück die Definition einer Anfrage. Die inhärenten Vagheiten der Interpretation einer graphischen Darstellung durch den Benutzer müssen vom System mit in Betracht gezogen werden. Um den Übersetzungsprozeß für graphische Anfragen an diesem Punkt zu dynamisieren, erscheint die Verbindung mit FUZZY-WING und damit einer unscharfen Interpretation der Anfrageklauseln sinnvoll. Damit ist gewährleistet, daß eine visuelle Anfrage nicht direkt in feste Suchbereiche übersetzt wird, sondern auch vom System einer Interpretation unterworfen wird. Dieser Interpretationsvorgang wird im Detail in Kap. 3 geschildert.

2.2.3 Benutzertest

Ein Benutzertest, bei dem WING-GRAPH von Werkstoffingenieuren anhand einer Aufgabenliste getestet wurde, ergab eine grundsätzliche Bestätigung für das graphische Faktenretrieval als Suchmodus.² Dies läßt sich auf zwei Phänomene zurückführen: Zum einen kann man mit einem bekannten Datenvisualisierungsformat recherchieren, zum anderen sind Anfragen möglich, die typische Informationsbedürfnisse unmittelbar umsetzen und in einer formalen Anfrage kaum zu formulieren wären. Der industrielle Kooperationspartner setzt derzeit die positiven Erfahrungen mit diesem Retrievalmodus in einer Reimplementierung um. Gleichzeitig deckte der Benutzertest aber Schwächen der zunächst starren Interpretationssemantik auf, die sich als zu wenig differenziert für ein polyparametrisches Anwendungsgebiet erwies (vgl. WOLFF 1996:219ff). Diese offensichtlichen Nachteile soll der hier vorgelegte Ansatz einer Kombination von vager Anfrageinterpretation mit der graphischen Recherchemöglichkeit ausgleichen helfen.

3 Wissensbasierte Vagheitskonzepte für das Graphische Faktenretrieval

Die oben angeführten Nachteile einer starren Interpretation graphischer Anfragekonstrukte wurden durch die Beobachtung unterstützt, daß Benutzer ihr Informationsbedürfnis meist mittels sprachlicher Vagheiten verbalisieren und daß auch das Denken von vagen Kategorien bestimmt wird. Dies gilt auch für die Interpretation graphischer Darstellungen („Suche nach einem Werkstoff, dessen Kennwertkurve im **hohen** Temperaturbereich **günstiger** verläuft“) und stellt einen weiteren Grund dar, graphische Suchkonstrukte unscharf zu interpretieren. Aus diesen Gründen werden unscharfe Interpretationen vager Terme herangezogen,

² Der Benutzertest umfaßte durch Videoaufnahmen protokollierte Aufgabenlösungen durch 10 Werkstoffexperten. Zu den Details der Testprotokolle und ihrer Auswertung vgl. MARX & SCHUDNAGIS 1997, WOLFF 1996.

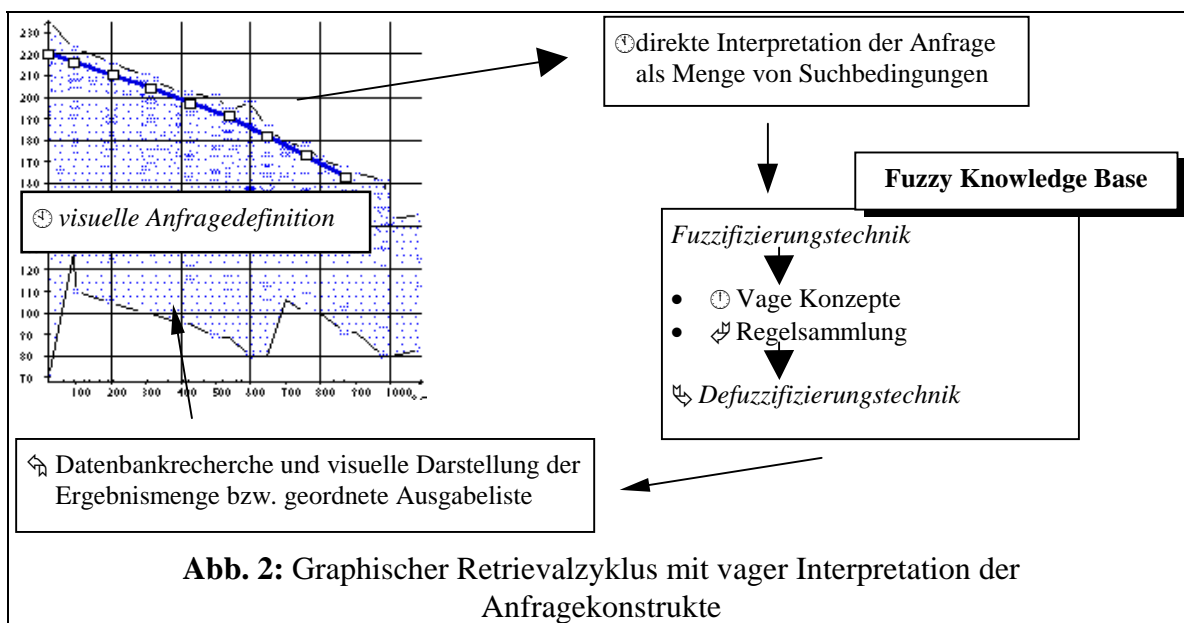
um die Semantik visueller Repräsentationen auszudrücken. Es gilt, ein Konzept für die Implementierung von Toleranzbereichen zu entwickeln, das sich mit den Vorstellungen der Benutzer deckt. Im folgenden wird ein Modell eines derartigen integrativen Ansatzes dargestellt.

3.1 Verfahren der unscharfen Interpretation visueller Recherchen

Unter Rückgriff auf vage Interpretationen läuft der graphische Retrievalzyklus wie folgt ab:

1. Der Benutzer erstellt ein graphisches Suchkonstrukt (Suchkurve, Streuband etc.) bezüglich eines Werkstoffkennwerts im vorgegebenen Koordinatensystem (vgl. Abb. 1).
2. Diese visuelle Anfragedefinition wird vom System auf die Punkte reduziert, für die aktuelle Werte in der Datenbank vorkommen können. Dies geschieht mit Hilfe von Domänenwissen; ein SQL-Statement kann wie in Kap. 2.2.2 dargestellt aufgebaut werden.
3. Fuzzifizierung: Für jedes Wertepaar einer Suchbedingung werden die Einzelwerte der abhängigen und unabhängigen Variablen auf unscharfe Konzepte abgebildet (z.B. Temperatur als unabhängige und Elastizitätsmodul als abhängige Variable).
4. Regelarbeitung zur Bestimmung der Suchtoleranzen: Inferenzregeln geben an, welche Quantifizierung des gesuchten Toleranzbereichs für einen bestimmten Kontext, d.h. eine Kombination der Inputvariablen, gültig ist.
5. Defuzzifizierung: Hier erfolgt eine Rücktransformation der vagen Konzepte in exakte Wertebereiche, wobei als Maß jeweils ein bestimmter prozentualer Anteil vom maximalen Wertespektrum der Datenbank an einem Meßpunkt dient (s. u. Tab. 3). Mit Hilfe dieser Anteilswerte läßt sich das endgültige SQL-Statement für die Recherche generieren.
6. Die Suche wird durchgeführt, wobei die Suchergebnisse nach absteigendem Ähnlichkeitsgrad mit den unscharfen Konzepten gerankt werden. Falls die Ergebnismenge klein (< 10) ist, werden die Werkstoffkurven unmittelbar visualisiert.

Abb. 2 illustriert diesen Prozeß:



Zur Adaption dieses Verfahrens an eine bestimmte Domäne wurden empirische Analysen herangezogen, um die Werte der Zugehörigkeitsfunktionen zu den *fuzzy*-Konzepten sachlich adäquat ermitteln zu können: In Experimenten mit Werkstoffexperten mußten diese vage Aufgabenstellungen mit einem traditionellen DBMS (*exact match*) lösen (z. B.: „Gesucht wird ein Werkstoff, der in *hohen* Temperaturbereichen einen *niedrigen* E-Modul und eine *hohe* Festigkeit aufweist.“). Die Übersetzung der vagen Terme in exakte Wertebereiche durch mehrere, verschiedenen Abteilungen angehörende Benutzer, führte zu einer Verteilung, die es zuließ, allgemeine Vagheitskonzepte daraus abzuleiten. Vorhandene Unterschiede konnten durch Kontextfaktoren wie z. B. das Aufgabengebiet und die Abteilungszugehörigkeit erklärt werden. Um diesen unterschiedlichen Sichtweisen Rechnung zu tragen, wurde im System die Möglichkeit vorgesehen, die Vagheitskonzepte manipulierbar und damit adaptierbar zu machen. Zusätzlich fließt bei der Steuerung der Toleranzbereiche im graphischen Retrieval Domänenwissen mit ein. Der Regelapparat basiert auf einem durch Werkstoffexperten formulierten Toleranzkonzept.

3.2 Die Modellierung von Toleranzen auf der Basis von fuzzy logic

Es konnte gezeigt werden, daß eine exakte Übernahme der von Benutzern erstellten graphischen Suchkonstrukte zu kurz greift. Deshalb wird ein dynamischer Ansatz zur Modellierung der Toleranzbereiche vorgeschlagen, der seine Motivation aus der Analogie zum Expertenverhalten zieht. Menschen entscheiden oft auf der Basis vager Kategorien und sind nicht immer in der Lage, *exakte* Verbalisierungen vorzunehmen. Je nach Kontext können verschiedene Toleranzdefinitionen angewendet werden. Der entscheidende kontextuelle Faktor betrifft in diesem Fall bestimmte Kurvenbereiche, d.h. die Beziehung zwischen abhängiger und unabhängiger Variablen. Beispielsweise ist das Konzept *große Dichte bei hohen Temperaturen* hinsichtlich der Toleranzbildung anders zu handhaben als in niedrigen Temperaturbereichen. Das bedeutet, daß Toleranzen in Abhängigkeit von den Kurvenbereichen, auf die sie angewendet werden sollen, definiert werden müssen. Dazu dient ein *fuzzy logic*-Ansatz:

Während des Fuzzifikationsprozesses werden exakte Meßwerte, d.h. die unmittelbare numerische Interpretation des graphischen Suchkonstrukts, *linguistischen Variablen* zugeordnet. Diese Terme (*hoch, mittel, niedrig* etc.) repräsentieren unscharfe Mengen, jede Werkstoffeigenschaft besitzt dabei für sie spezifische Zuordnungskriterien. Auf diese Weise ist es möglich, graduelle Zugehörigkeiten zu modellieren, wie sie anhand eines Beispiels in Abb. 3 gezeigt werden:

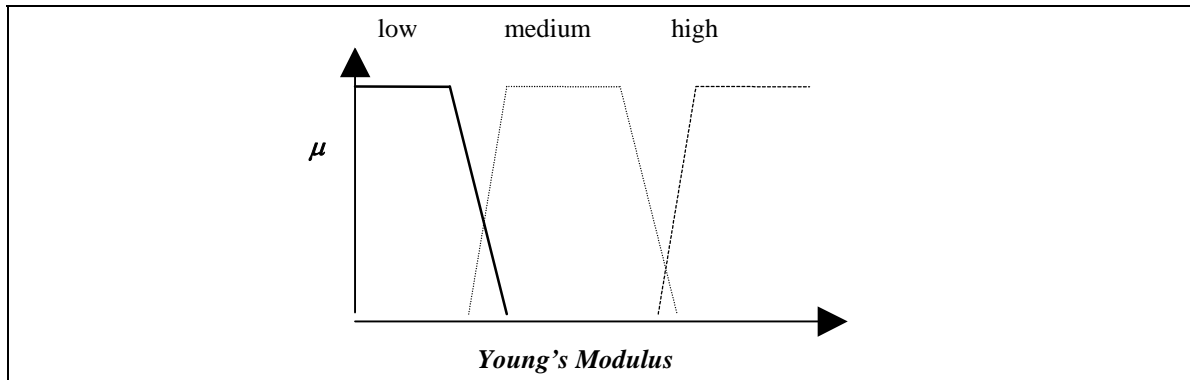


Abb. 3: Vage Konzepte und Termmengen einer linguistischen Variablen

Für jeden Kennwert der Datenbasis und jeden Parameter werden linguistische Variablen und entsprechende unscharfe Terme kreiert. Analog wird eine Variable *Toleranz* mit den unscharfen Termen *sehr groß*, *groß*, *mittel*, *gering* definiert. Aus ihnen ist abschließend in der Defuzzifizierungsphase die Quantifizierung des Toleranzkonzepts abzuleiten. Die Abbildung der linguistischen Variablen, die einer Suchbedingung zugeordnet ist, auf eine Toleranzklasse erfolgt durch Inferenzregeln. Nach diesem Schema wird jede kontextuelle Situation und deren „logische“ Folgerung ausgedrückt, z.B.:

WENN Elastizitätsmodul *hoch*
 UND
 Temperatur *hoch*
 DANN Toleranz *sehr gering*

Die Prämissen (*WENN*-Teile) der Regeln beschreiben die Inputvariablen, d.h. die möglichen Situationen, während die Konklusionen (*DANN*-Teile) die Reaktionen auf die möglicherweise komplexen Situationen ausdrücken. Die unscharfe Inferenz basiert auf der Aggregation (Berechnung der *WENN*-Teile) und der Komposition (Berechnung der *DANN*-Teile). Besteht der Prämissenteil aus mehr als einer Bedingung, werden spezifische Kombinationsoperatoren eingesetzt. Analog zur Booleschen Logik fanden zunächst folgende elementaren Operatoren Verwendung (vgl. KLIR & FOLGER 1988):

AND $\mu_{A \wedge B} = \min \{ \mu_A, \mu_B \}$
OR $\mu_{A \vee B} = \max \{ \mu_A, \mu_B \}$
NOT $\mu_{\neg A} = 1 - \mu_A$

Mittels derartiger Operatoren läßt sich die Gültigkeit des Bedingungssteils berechnen.³ Das Ergebnis stellt den Grad dar, mit dem die jeweilige Regel für die aktuelle Situation gültig ist. Falls alternative Regeln zur gleichen Konklusion führen, werden die Einzelergebnisse durch die sog. *MAX/MIN*-Inferenz (*fuzzy ODER*-Verknüpfung) zusammengeführt. Als Ergebnis steht zunächst ein vager Wert der linguistischen Variablen *Toleranz* zur Verfügung, der durch Defuzzifizierung auf exakte Wertebereiche abgebildet werden muß. Die Defuzzifizierung reagiert in umgekehrter Weise zur Fuzzifizierung, d.h. sie führt von der linguistischen Variablen zu exakten Wertebereichen. Da sich vage Konzepte überschneiden können, können mehrere Konzepte eine Gültigkeit > 0 besitzen. Hier kommt wiederum Exper-

³ Komplexere Operatoren wie z.B. der *Gamma*-Operator ermöglichen die Einbeziehung von kompensatorischen Aspekten. Hier werden wiederum empirische Untersuchungen die Auswahl geeigneter Parameter unterstützen, vgl. MAYER et al. 1993:46ff.

tenwissen zum Tragen. Menschliche Experten lösen derartige Situationen, indem sie Kompromisse zwischen den beteiligten Konzepten bildet. Dieses Verhalten wird durch die *Center-of-Gravity*-Methode zur Defuzzifizierung modelliert (vgl. MAYER et al. 1993:80ff).

3.3 Ein Anwendungsbeispiel

Das folgende Beispiel illustriert die Anwendung einer vagen Interpretation von Toleranzen anhand des bezüglich der Temperatur gemessenen Werkstoffkennwerts *Elastizitätsmodul*. Durch die Verbindung zwischen vager Kennwert- und Parameterinterpretation und der domänenspezifischen Handhabung der Toleranz entsteht eine zweidimensionale Regeltabelle. Die entsprechenden Interpretationen für jedes Wertepaar werden unter Einbeziehung von Expertenwissen definiert (vgl. Tab. 2) und auf prozentuale Wertebereiche abgebildet (vgl. Tab. 3). Diese Abbildung kann durch weitere empirische Analysen optimiert werden.

	<i>hohe</i> Temperatur	<i>mittlere</i> Temperatur	<i>niedrige</i> Temperatur
<i>hoher</i> Elastizitätsmodul	sehr niedrig	niedrig	mittel
<i>mittlerer</i> Elastizitätsmodul	mittel	mittel	hoch
<i>niedriger</i> Elastizitätsmodul	hoch	sehr hoch	sehr hoch

Tab. 2: Toleranzklassen in Abhängigkeit von Kennwert und Temperatur

<i>Linguistische Variable</i>	<i>effektiver Toleranzbereich (prozentualer Anteil am lokalen Wertespektrum in der Datenbank)</i>
sehr niedrig	1 %
niedrig	5 %
mittel	10 %
hoch	15 %
sehr hoch	30%

Tab. 3: Numerische Interpretation der Toleranzklassen (Defuzzifizierung)

Beispiel: Ein Suchpunkt, der einen Elastizitätsmodul von 220 GPa bei einer Temperatur von 1000 ° Celsius repräsentiert, wird auf folgendes vages Toleranzkonzept abgebildet (Tab. 4); die Gewichtung der einzelnen Konzepte berechnet sich anhand der für diesen Kennwert definierten Zugehörigkeitsfunktion (Abb. 3).

	<i>hohe</i> Temperatur	<i>mittlere</i> Temperatur	<i>niedrige</i> Temperatur
<i>hoher</i> Elastizitätsmodul	0,9 sehr niedrig	0,8 niedrig	0 mittel
<i>mittlerer</i> Elastizitätsmodul	0,7 mittel	0,4 mittel	0 hoch
<i>niedriger</i> Elastizitätsmodul	0 hoch	0 sehr hoch	0 sehr hoch

Tab. 4: Instantiierung der Korrelation für das Beispiel

Tritt der höchste Wert einer linguistischen Variablen mehrfach auf, wird ein *fuzzy*-Akkumulationsoperator zur Konfliktlösung herangezogen. Der Wert jeder *fuzzy*-Interpretation dient als Basis für die Berechnung der aktuellen Toleranzklasse. In obigem Beispiel erhält der Datenpunkt (220 GPa, 1000 Grad Celsius) 0,9 als höchsten Zugehörigkeitswert und wird damit mit einer *sehr niedrigen* Toleranz verbunden. Die Defuzzifizierung bestimmt die Toleranz für jede Bedingung, indem auf die Prozentwerte der Tab. 3 zurückgegriffen wird.

Abschließend findet unter Einbeziehung der berechneten Toleranzbereiche für jedes Wertepaar die endgültige Datenbanksuche statt. Werden weniger als 10 Werkstoffe gefunden, erfolgt die Visualisierung des Ergebnisses unmittelbar in Anschluß an die Suche. Erfüllen mehr als 10 Werkstoffe die Bedingungen, können Benutzer aus sortierten Ergebnislisten auswählen, wobei als Sortierkriterium die gewichtete Distanz zwischen den Meßpunkten der Ausgangskurve und jeder Ergebniskurve herangezogen wird.

4 Implementierung und Ausblick

WING-GRAPH ist in C++ implementiert und läuft unter Windows 3.1 und Windows95/NT. Als Datenbankengine dient Gupta SQLBase, die Ansteuerung der Datenbank erfolgt wahlweise über das proprietäre SQLBase SQL API bzw. über eine ODBC-Schnittstelle (mit Performanzverlusten). Wir haben uns bei der Implementierung für ein Standard-RDBMS entschieden, weil dies am besten mit den Vorgaben des industriellen Kooperationspartners zu vereinbaren war und nur ein solches System die für die Verarbeitung der komplexen Anfragen notwendige Performanz aufweist. Aus diesen Gründen sind neuere Datenbankkonzepte zunächst nicht berücksichtigt. Die FUZZY-WING-Komponente verwendet die DataEngine ADL-Klassenbibliothek, mit deren Hilfe die Zugehörigkeitsfunktion der unscharfen Bewertung für jeden Kontext (d.h. Werkstoffkennwert) spezifiziert und ausgewertet werden kann.

Der vage Interpretationsmechanismus für das graphische Faktenretrieval zeigt, wie sich das Phänomen vager natürlichsprachlicher und graphischer Anfragen an relationale Datenbanksysteme behandeln läßt. Damit wird deutlich, daß typische IR-Konzepte wie *ranking* und *partial match* auch im Bereich Faktendatenbanken fruchtbar zu machen sind.

Die Weiterentwicklung des Systems beinhaltet in technischer Hinsicht die Portierung nach Java, um das graphische Suchverfahren auch im Internet anbieten zu können (Datenbank-anbindung über das JDBC-Interface). Gleichzeitig ist klar, daß eine inhaltliche Generalisierung notwendig ist, um das System domänenunabhängig zu machen und es für beliebige Anwendungskontexte, in denen die Visualisierung von numerischer Information eine Rolle spielt, verfügbar zu machen. Ein formularbasiertes Integrationswerkzeug soll es dem Benutzer ermöglichen, neue Datenbestände zu integrieren, Domänencharakteristika zu spezifizieren und die Bewertungsfunktion der unscharfen Interpretation anzupassen.

5 Literatur

- BOSC, P., PIVERT, O. (1992). „About Equivalences in SQLf, A Relational Language Supporting Imprecise Querying.“ In: TERANO, T. et al. (edd.) (1992). Fuzzy Engineering Toward Human Friendly Systems. Amsterdam, 309-320.
- BUELL, D.A., Kraft, D. (1983). „Fuzzy Sets and Generalized Boolean Retrieval Systems.“ In: International Journal of Man-Machine Studies 19 (1983), 45-56.
- FRAKES, W. (1993). „Introduction.“ In: FRAKES, W.; BAEZA-YATES, R. (edd.) (1993). Information Retrieval: Data Structures and Algorithms. Englewood Cliffs/NJ: P T R Prentice Hall.

- FUHR, N. (1989). „Models for Retrieval with Probabilistic Indexing.“ In: *Information Processing & Management* 25(1) (1989), 55-72.
- FUHR, N.; HOFFMANN, T. (1991). „A Prototype for Integrating Fact and Text Retrieval.“ In: KILLENBERG, H.; KUHLEN, R.; MANECKE, H.-J. (edd.) (1991). *Wissensbasierte Informationssysteme und Informationsmanagement. Proc. 2. Internationales Symposium für Informationswissenschaft, ISI '91. Konstanz: UVK Informationswissenschaft, 94-103.*
- KLIR, G.J.; FOLGER, T.A. (1988). *Fuzzy Sets, Uncertainty, and Information. Englewood Cliffs/NJ: Prentice Hall.*
- KOSSLYN, St. (1989). „Understanding Graphs and Charts.“ In: *Applied Cognitive Psychology* 3 (1989), 185-226.
- KRAUSE, J. et al. (1993/94). *Multimodality and Object Orientation in an Intelligent Materials Information System. In: Journal of Document and Text Management* 1(3) (1993), 256-275 and 2(1) (1994), 48-66.
- KRAUSE, J., WOLFF, Ch., WOMSER-HACKER, Ch. (1994). *Multimodality and Vagueness in the Context of a Graphical, Object-Oriented Materials Information System. In: Proceedings of RIAO '94, Intelligent Multimedia Information Retrieval Systems and Management. New York, S. 585-597.*
- KRAUSE, J.; WOMSER-HACKER, Ch. (1997). *Vages Information Retrieval und graphische Benutzungsoberflächen. Beispiel Werkstoffinformation. Konstanz: UVK Informationswissenschaft.*
- LARKIN, Jill H.; SIMON, Herbert A. (1987). „Why a Diagram is (Sometimes) Worth Ten Thousand Words.“ In: *Cognitive Science* 11 (1987), 65-99.
- MANDL, Th., WOMSER-HACKER, Ch. (1995). „Softcomputing“-Verfahren zur Behandlung von Ähnlichkeit und Vagheit in objektorientierten Informationssystemen.“ In: KUHLEN, R., RITTBERGER, M. (edd.) (1995). *Hypertext - Information Retrieval - Multimedia. Synergieeffekte elektronischer Informationssysteme. Proceedings HIM '95, Konstanz: UVK Informationswissenschaft, 277-291.*
- MARX, J.; SCHUDNAGIS, M. (1997). „Überblick über die WING-IIR-Benutzertests und methodisches Vorgehen.“ In: *KRAUSE & WOMSER-HACKER (1997) 43-58.*
- MAYER, A., MECHLER, B., SCHLINDWEIN, A., WOLKE, R. (1993). *Fuzzy Logic. Bonn et al.: Addison-Wesley.*
- MIYAMOTO, S. (1990). *Fuzzy Sets in Information Retrieval and Cluster Analysis. Dordrecht: Kluwer.*
- PINKER, St. (1990). „A Theory of Graph Comprehension.“ In: FREEDLE, Roy (ed.) (1990). *Artificial Intelligence and the Future of Testing. Hillsdale/NJ: Lawrence Erlbaum, 73-126.*
- RADECKI, T. (1979). „Fuzzy Set Theoretical Approach to Document Retrieval.“ In: *Information Processing & Management* 15 (1979), 247-259.
- SALTON, G. (1989). *Automatic Text Processing. The Transformation, Analysis, and Retrieval of Information by Computer. Reading/MA: Addison-Wesley.*
- TUFTE, E.R. (1983). *The Visual Display of Quantitative Information. Cheshire/Conn.: Graphics Press.*
- TUFTE, E.R. (1990). *Envisioning Information. Cheshire/Conn.: Graphics Press.*
- TUFTE, E.R. (1997). *Visual Explanations. Cheshire/Conn.: Graphics Press.*
- VAN RIJSBERGEN, C.J. (1979²). *Information Retrieval. London: Butterworths.*

Graphisches Faktenretrieval mit vager Anfrageinterpretation

WOLFF, Ch. (1996). Graphisches Faktenretrieval mit Liniendiagrammen. Konstanz: UVK Informationswissenschaft.

WOLFF, Ch. (1997). Graphisches Retrieval in Wing-Graph.“ In: KRAUSE & WOMSER-HACKER (1997), 113-143.

WOMSER-HACKER, Ch. (1997). „FUZZY-WING: ein Werkzeug für Faktenanfragen mit vagen Kriterien.“ In: KRAUSE & WOMSER-HACKER (1997), 185-203.