

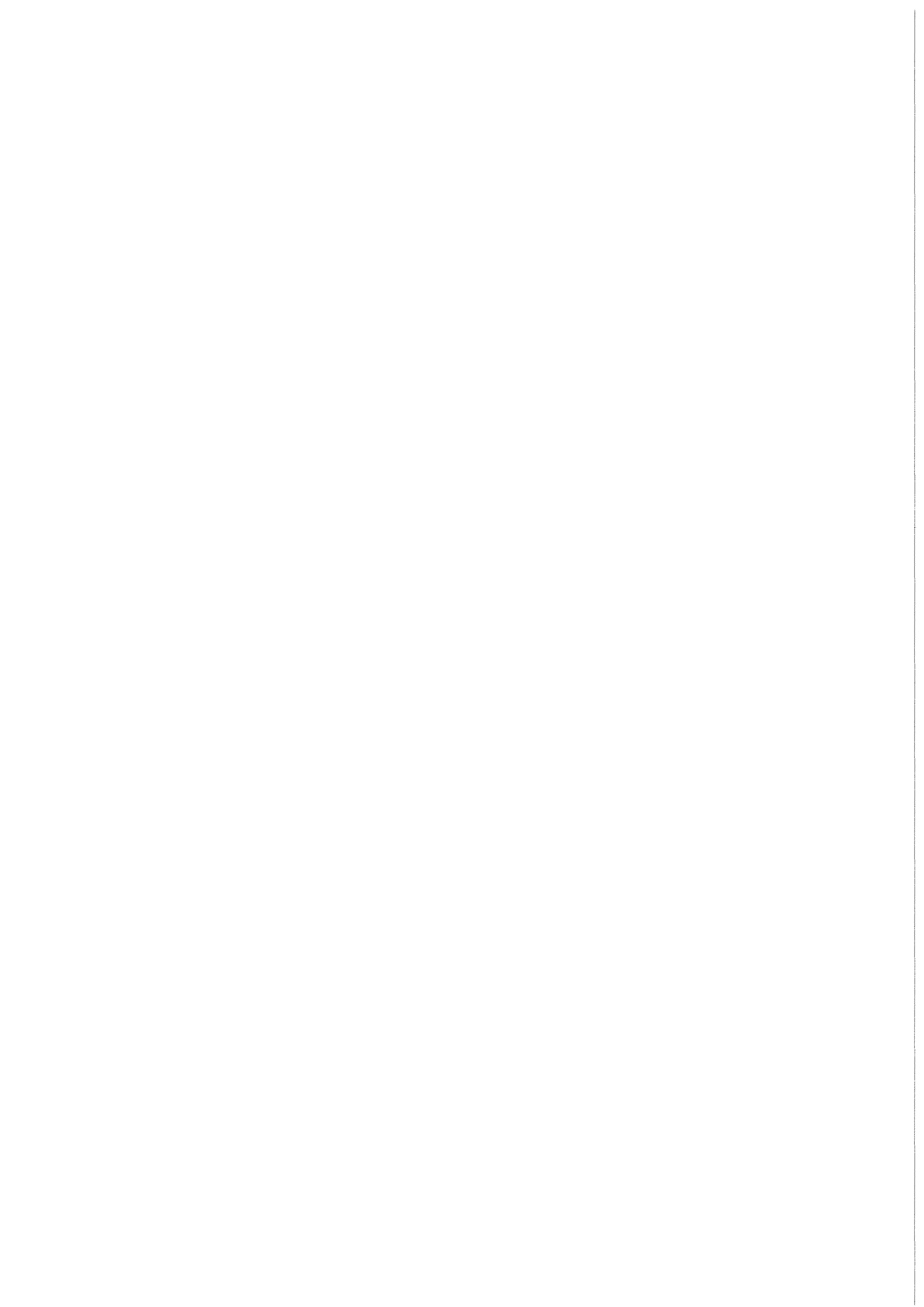


KfK 4600
August 1989

Eine interaktive Wissenserwerbskomponente für ein wissensbasiertes Atlasensystem

U. Clausen
Institut für Datenverarbeitung in der Technik
Projekt Schadstoffbeherrschung in der Umwelt

Kernforschungszentrum Karlsruhe



KERNFORSCHUNGSZENTRUM KARLSRUHE

Institut für Datenverarbeitung in der Technik
Projekt Schadstoffbeherrschung in der Umwelt

KfK 4600

Eine interaktive Wissenserwerbskomponente für
ein wissensbasiertes Altlastensystem

Uwe Clausen

Kernforschungszentrum Karlsruhe GmbH, Karlsruhe

Als Manuskript vervielfältigt
Für diesen Bericht behalten wir uns alle Rechte vor

Kernforschungszentrum Karlsruhe GmbH
Postfach 3640, 7500 Karlsruhe 1

ISSN 0303-4003

Zusammenfassung

Eine interaktive Wissenserwerbskomponente für ein wissensbasiertes Altlastensystem

Es werden drei Kategorien von Wissenserwerb für wissensbasierte Systeme erläutert. Das Anwendungsgebiet Altlasten und der vorhandene Prototyp des Systems XUMA (Expertensystem Umweltgefährlichkeit von Altlasten) werden unter diesem Aspekt untersucht. Die Unterschiede zwischen etablierten Architekturen im Bereich 'Heuristische Klassifikation' und der Architektur von XUMA werden erörtert. Basierend auf Regelmäßigkeiten der Wissensbasis des Prototyps wird eine anwendungsspezifische Komponente entworfen. Das System kennt Objekte der Ebenen Regel, Aussageformen in Regeln und Regelgruppen. Den verschiedenen Klassen von Objekten werden jeweils menüorientierte Methoden zu ihrer Darstellung oder Modifikation zugeordnet. Das Wissen kann in für den Fachexperten aussagekräftiger Sprache dargestellt werden. Aufbauend auf den Regelgruppen unterstützt eine Zugangsheterarchie die Strukturierung der Wissensbasis nach wichtigen Begriffen. Das System wurde in Common Lisp und mit Flavorn auf TI Explorer II implementiert.

Abstract

A knowledge acquisition facility for a knowledge-based system on risk assessment of contaminated sites

Three categories of knowledge acquisition are explained. Expertise in 'Risk assessment of hazardous waste sites' and a prototype of the system XUMA (Expert system environmental hazard of contaminated sites) are analysed from the knowledge acquisition point of view. The differences between established architectures for heuristic classification and the architecture of XUMA are considered. Based on regularities in the prototype's knowledge base an application-dependend knowledge acquisition facility is designed. The structure for knowledge entry or modification by menu-oriented methods is based on objects on the three levels rules, forms within rules and groups of rules. The knowledge can be represented in terms familiar to domain experts. The concept of rule groups provides an access heterarchy and thus supports structuring the knowledge base. The system was implemented in Common Lisp and the Flavor-System on a TI Explorer II.

Inhaltsverzeichnis

1. Einleitung	1
2. Wissenserwerb für wissensbasierte Systeme	2
2.1 Begriffe	2
2.2 Wissenserwerb und Phasenmodelle	2
2.3 Kategorien des Wissenserwerbs	3
2.4 Wissenserwerb mit Wissensingenieur	5
2.5 Automatischer Wissenserwerb	6
2.6 Direkter Wissenserwerb	7
2.7 KADS-Modell	11
2.8 Hybride Entwicklungsumgebungen	12
2.9 Schlußfolgerungen	13
3. Altlasten	15
3.1 Einleitung	15
3.2 Begriffe	15
3.3 Bedeutung der Altlastenproblematik	16
3.4 Sanierung	17
3.5 Beurteilung und Bewertung von Altlasten	18
3.6 Wissen im Altlastenbereich	20
4. Das Altlasten-System XUMA	22
4.1 Ziele	22
4.2 Aufbau	22
4.2.1 Analysenplanerstellung	23
4.2.2 Erfassung von Analysen	24
4.2.3 Beurteilung	24
4.2.4 Bewertung	25
4.3 Implementierung	25
4.4 Das Werkzeug ART	25
4.4.1 Überblick	25
4.4.2 Regeln in ART	26
4.4.3 Schemata in ART	26
4.4.4 Bewertung von ART	27
4.5 Untersuchungen der Wissensbasis in XUMA	28
4.6 Einordnung des Systems	32

5. Anforderungen und Systementwurf der Wissenserwerbskomponente . . .	34
5.1 Anforderungen	34
5.2 Grundlegende Konzepte	36
5.3 Klassen von Objekten	37
5.4 Struktur und Ablaufmodell	38
5.4.1 Auswahl	38
5.4.2 Eingabe	40
5.4.3 Test	41
5.4.3 Bestätigung	41
5.5 Weitere Konzepte	41
5.6 Zusammenfassung	42
5.7 Umfang	43
5.8 Erweiterungsmöglichkeiten	43
6. Detailentwurf und Realisierung der Wissenserwerbskomponente	44
6.1 Flavor-System	44
6.2 Methoden der Wissenserwerbskomponente	45
6.2.1 Menü-Darstellung von Regeln	45
6.2.2 Aktualisierung der Wissensbasis	47
6.2.3 Eingabe und Modifikation von Aussagen	49
6.3 Konventionen bei Namen und Strukturen	51
6.4 Abbildung zwischen ART-Notation und Instanzen	52
6.5 Menü-Oberfläche	53
6.6 Zugangsstruktur	53
6.7 Besonderheiten	55
6.8 Anwendungsbeispiele	56
7. Zusammenfassung	61
8. Literaturverzeichnis	63

1. Einleitung

Die vorliegende Arbeit umfaßt die Problemanalyse und den Entwurf einer interaktiven Wissenserwerbskomponente für ein wissensbasiertes Altlasten-System sowie die Beschreibung der dazugehörigen Implementierung. Sie wurde im Rahmen eines gemeinsamen Vorhabens¹⁾ des Instituts für Datenverarbeitung in der Technik (IDT) des Kernforschungszentrums Karlsruhe und des Instituts für Altlastensanierung der Landesanstalt für Umweltschutz (LfU) in Karlsruhe durchgeführt. Ziel dieses Vorhabens ist die Entwicklung eines wissensbasierten Altlastensystems, das die Sachbearbeiter in der LfU und in den Wasserwirtschaftsämtern bei der Beurteilung der Umweltgefährlichkeit schadstoffbelasteter Standorte unterstützt und die wenigen Fachexperten auf diesem Gebiet entsprechend entlastet.

Wichtiges Charakteristikum wissensbasierter Systeme ist die Trennung des Schlußfolgerungsmechanismus von dem zur Problemlösung verwendeten Wissen. Dieses Wissen wird in einer Wissensbasis explizit repräsentiert. Durch Änderung und Erweiterung der Wissensbasis soll eine einfache Anpassung an neue Erkenntnisse möglich sein. Die gängigen hybriden Entwicklungsumgebungen für wissensbasierte Systeme unterstützen die Änderung und Erweiterung der Wissensbasis aber nur unzureichend. Sie sind für die in der Regel damit nicht vertrauten Fachexperten eines Anwendungsgebietes nicht unmittelbar nutzbar, so daß ein langfristig erfolgreicher Einsatz eines wissensbasierten Systems gegenwärtig nur durch die Entwicklung einer spezifischen Wissenserwerbskomponente erreicht werden kann. Entwurf und Realisierung einer solchen Komponente für ein Altlasten-System ist Gegenstand der vorliegenden Arbeit.

Im zweiten Kapitel wird zunächst der Stand der Forschung im Bereich Wissenserwerb für wissensbasierte Systeme erläutert. Das dritte Kapitel stellt kurz die Problematik "Altlasten" dar und untersucht sie als Anwendungsgebiet für wissensbasierte Systeme. Es folgt eine Analyse der derzeitigen Ausbaustufe des Altlasten-Systems XUMA im Hinblick auf den Entwurf und die Integration einer interaktiven Wissenserwerbskomponente. Im fünften Kapitel wird der Systementwurf dargestellt. Über den Detailentwurf und die Realisierung in der Programmiersprache Common Lisp unter Verwendung der objektorientierten Spracherweiterung Flavor wird dann im sechsten Kapitel berichtet.

1) Das Vorhaben wird unter der Nummer PD 87053 durch das Projekt Wasser-Abfall-Boden des Landes Baden-Württemberg gefördert.

2. Wissenserwerb für wissensbasierte Systeme

In diesem Kapitel werden einige für diese Arbeit wichtige Begriffe erläutert. Es werden drei prinzipielle Kategorien von Wissenserwerb und ihre spezifischen Vor- und Nachteile beschrieben. Eine Auswahl wichtiger Systeme wird vorgestellt.

2.1 Begriffe

Die Beschäftigung mit wissensbasierten Systemen ist ein Teilbereich der Künstlichen Intelligenz. Bezüglich des zweiten Begriffes schließt sich der Verfasser der Definition von [Stoyan 88] an und versteht die Künstliche Intelligenz als "die Wissenschaft, deren Ziel die Simulation von menschlichem Verhalten auf Rechenanlagen ist". Dabei gilt die Künstliche Intelligenz als Teilgebiet der Informatik mit Einflüssen aus anderen Wissenschaften, wie etwa der Kognitionsforschung.

Von einem wissensbasierten System spricht man, wenn das zur Problemlösung verwendete Wissen von der Schlußfolgerungskomponente getrennt und explizit in einer Wissensbasis repräsentiert wird. Diese Definition trifft genauso für den Begriff Expertensystem zu. Da jedoch immer nur bestimmte Aspekte der Fähigkeiten von Experten in Systemen nachgebildet werden können, wird im angelsächsischen Sprachraum auch der Terminus "knowledge based system" (wissensbasiertes System) verwendet. In der vorliegenden Arbeit wird dieser Begriff bevorzugt. Die Begriffe Wissenserwerb und Wissensakquisition werden im folgenden synonym verwendet.

Mit hybriden Entwicklungsumgebungen werden Werkzeuge zur Entwicklung wissensbasierter Systeme bezeichnet, die neben einem Schlußfolgerungsmechanismus mehrere kombinierbare Formalismen zur Repräsentation von Wissen bereitstellen. Die mächtigsten und verbreitetsten Systeme sind zur Zeit KEE, ART und Knowledge Craft.

2.2 Wissenserwerb und Phasenmodelle

Wissenserwerb umfaßt nach [Puppe 88] "die Identifizierung, Formalisierung und Wartung des Wissens, das ein Expertensystem zum Problemlösen braucht". Aus

[Hayes-Roth et al. 83] stammt ein auf Buchanan zurückgehendes Standardphasenmodell (Abbildung 2-1), das für das zu erwerbende Wissen die Ausprägungen Anforderungen, Konzepte, Strukturen und schließlich Regeln kennt.

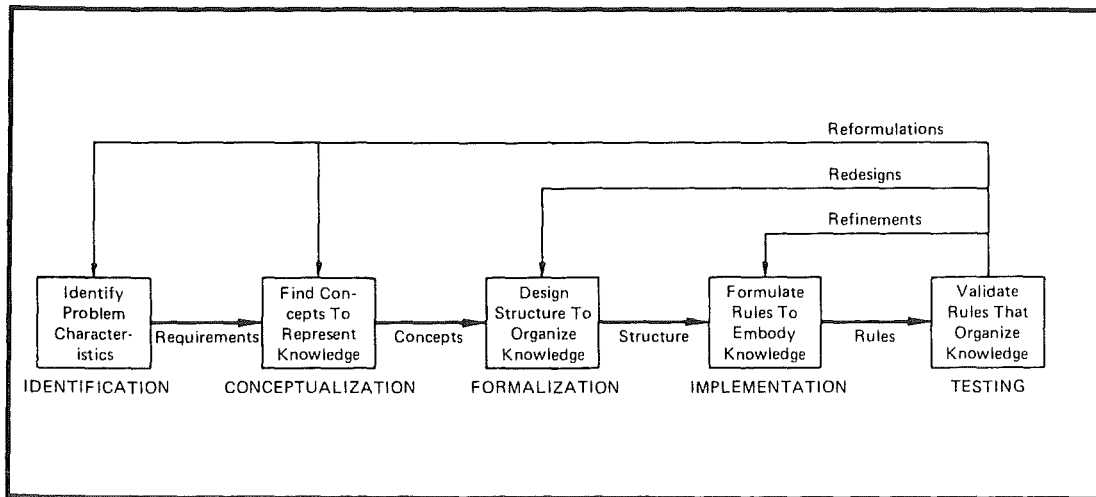


Abb.2-1 Buchanan-Phasenmodell für den Wissenserwerb (nach [Hayes-Roth et al. 83])

In diesem, wie auch in späteren Phasenmodellen des Wissenserwerbs geht es jeweils um das zur Lösung von bestimmten Aufgaben erforderliche Wissen. Diese Perspektive erklärt, daß unter Anforderungen gerade die Anforderungen dieser Aufgaben verstanden werden. Mit Konzepten sind die wesentlichen Prinzipien gemeint, nach denen Wissen aus einem Anwendungsbereich für die Lösung einer solchen Aufgabe formalisiert werden kann. Die Formalisierung betont den Entwurfsaspekt, dessen Ergebnis konkrete Strukturen sind. Regeln sind eine wichtige, allerdings nicht die einzige Form zur Implementierung des Wissens in einer durch die Maschine nutzbaren Form. Dies ist auch eine Schwäche dieses Phasenmodells. [Puppe 88] schlägt eine Konkretisierung vor (Abbildung 2-2), welche die einzelnen Phasen um Durchführungsvorschläge und Angaben über Hilfsmittel ergänzt.

2.3 Kategorien des Wissenserwerbs

Es lassen sich drei Kategorien von Wissenserwerb unterscheiden :

- Wissenserwerb mit Wissensingenieur

Bei dieser auch als indirekter Wissenserwerb bezeichneten Kategorie wird das Wissen des Fachexperten in das System übertragen, indem ein Wissensingenieur ihn interviewt und die in der Regel verbalen Äußerungen auf eine geeignete Repräsentation und Problemlösungsstrategie abbildet. Dazu gibt

PHASEN	BESCHREIBUNG	DURCHFÜHRUNG	HILFSMITTEL
<pre> graph TD A[Problemcharakterisierung] --> B[Shell-Entwicklung] B --> C[Aufbau der Wissensbasis] C --> D[Wartung der Wissensbasis] D --> A C --> A B --> A </pre>	<p>Identifikation der Problemlösungsstrategie und der Wissensrepräsentation</p>	<p>Wissensingenieur befragt Experten</p>	<p>Interviewtechniken und Protokolle</p>
	<p>Bereitstellung eines Expertensystem-Shells mit komfortabler Wissenserwerbskomponente</p>	<p>Auswahl oder Neuentwicklung durch den Wissensingenieur</p>	<p>allgemeines Expertensystem-Werkzeug</p>
	<p>Formalisierung des Expertenwissens</p>	<p>Experte, eventuell unterstützt durch Wissensingenieur</p>	<p>Shell</p>
	<p>Tuning und Weiterentwicklung, Anpassung an neue oder geänderte Anforderungen</p>	<p>Experte, eventuell unterstützt durch automatische Analysetechniken</p>	<p>Falldatenbank</p>

Abb.2-2 erweitertes Phasenmodell (nach [Puppe 88])

es einige spezielle Vorgehensweisen, wie das strukturierte Interview, das Laut-Denken-Protokoll und die Introspektion, die das traditionelle (unstrukturierte) Interview ergänzen können [Kidd 87].

- automatischer Wissenserwerb

Mit Verfahren des maschinellen Lernens gewinnt das System Wissen automatisch aus Falldaten oder Fachliteratur. Neben der Leistungsfähigkeit der eigentlichen Verfahren ist die hier wesentliche Frage, ob die Ausgangsdaten in nutzbarer Form und ausreichender Menge und Qualität zur Verfügung stehen. Test und Validation einer durch automatischen Wissenserwerb erzeugten Wissensbasis ist oft problematisch, ebenso wie in bestimmten Anwendungsgebieten die Frage nach der Verantwortlichkeit für von solchem Wissen abhängige Entscheidungen.

- direkter Wissenserwerb

Der Fachexperte formalisiert selbst sein Wissen. Er gibt es durch Interaktion mit der Wissenserwerbskomponente des Systems ein, prüft und korrigiert es gegebenenfalls. Dies setzt eine angemessene, gut handhabbare Schnittstelle voraus.

In den weiteren Abschnitten dieses Kapitels werden die einzelnen Kategorien genauer untersucht.

2.4 Wissenserwerb mit Wissensingenieur

Ein Wissensingenieur ist eine Person, die einerseits soweit Einblick in das Anwendungsgebiet hat, daß sie sich unmittelbar mit dem Fachexperten darüber verständigen kann, und die andererseits in der Lage ist, die geeignete Repräsentation und Problemlösungsstrategie auszuwählen, um das gewonnene Wissen in ein auf Rechnern ablauffähiges System einzubringen. Es handelt sich dabei trotz weitgehender Verbreitung des Begriffes Wissensingenieur nicht um eine Tätigkeit etwa im Sinne eines eigenständigen Berufes. Hierzu sei auf den Beitrag [Crasemann et al. 88] verwiesen, der die Aufgaben des Wissensingenieurs gegenüber denen abwägt, die ohnehin Aufgaben des Informatikers bzw. Systemanalytikers sind. Die erforderliche Qualifikation geht weit über die Beherrschung der in wissensbasierten Systemen verwendeten Formalismen hinaus. Ein gute Darstellung der psychologischen Aspekte und interessante praktische Anmerkungen enthält [Davies et al. 88].

Die oben erwähnten Interviewtechniken unterscheiden sich in der Rolle des Wissensingenieurs. Er kann einerseits aktiv durch Fragen ein Interview steuern, oder andererseits die Äusserungen des Experten passiv analysieren. Ein weiteres Unterscheidungsmerkmal ist, ob und welche konkreten Anwendungsbeispiele zugrundegelegt werden. Das sogenannte 'Laut-Denken-Protokoll' ist ein unmittelbares Protokoll der vom Experten geäußerten Gedankenschritte bei der Lösung eines Testfalles. Davon wird die Introspektion, eine zusammenfassende Beschreibung des Vorgehens, unterschieden. Insbesondere in der Phase der Formalisierung von Wissen im Rahmen einer zuvor entworfenen Struktur bietet sich dann das strukturierte Interview an, in dem Einzelheiten ergänzt und Unklarheiten ausgeräumt werden können [Kidd 87].

Für ein wissensbasiertes System muß Wissen in genauer und detaillierter Form repräsentiert werden. In der zwischenmenschlichen Kommunikation wird dagegen vieles vorausgesetzt und - bewußt oder unbewußt - nicht explizit erwähnt, obwohl es für die Problemlösung relevant ist. Dies ist eine wesentliche Ursache für den Aufwand, den dieses Vorgehen beim Wissenserwerb in der Regel erfordert. Ein anderes Problem, das sich durch die Arbeit eines Wissensingenieurs als Mittler zwischen Fachexperte und System ergibt, ist, ob das Verständnis des Fachexperten auf der einen und die Interpretation und ent-

sprechende Formalisierung durch den Wissensingenieur auf der anderen Seite hinreichend übereinstimmen. Dieses Problem verlangt immer wieder Validationen, in denen etwa der Fachexperte mit Teilfunktionen des Systems und der Wirkungsweise des formalisierten Wissens konfrontiert wird. Der so insgesamt entstehende Aufwand macht dieses Verfahren für den Wissenserwerb, insbesondere im Hinblick auf langfristigen wirtschaftlichen Erfolg, fragwürdig.

Eine Untersuchung über den Wissenserwerb bei etwa 70 Projekten [Cullen et al. 88] veranlaßt die Autoren zu der These, daß es den vielzitierten 'Flaschenhals Wissenserwerb' gar nicht gibt, sondern das Finden der geeigneten Repräsentation das eigentliche Problem darstellt. Diese Einschätzung trifft von ihren Beobachtungen her sicher zu, übersieht aber, daß die so verstandene 'Wissensrepräsentation' ebenfalls Teil des Wissenserwerbs ist.

Abschließend kann zu dieser Kategorie 'Wissenserwerb mit Wissensingenieur' festgestellt werden, daß es in der Regel nicht schwierig ist, Aussagen des Experten über sein Wissen zu erhalten. Das Problem ist, die richtige Sorte von Wissen zu erhalten [Eshelman et al. 87].

Konzepte für Komponenten, die diese Form des Wissenserwerbs ersetzen oder zumindest den Wissensingenieur entlasten, werden in späteren Abschnitten erläutert. Wo solche Komponenten nicht einsetzbar sind, wird er jedoch auch weiterhin eine wichtige Funktion erfüllen. Dies betrifft in der Buchanan'schen Terminologie insbesondere die frühen Phasen des Wissenserwerbs.

2.5 Automatischer Wissenserwerb

Dieser Bereich umfaßt textverstehende Systeme und Verfahren des maschinellen Lernens. Beide Teilbereiche sind Gegenstand der Grundlagenforschung innerhalb der Künstlichen Intelligenz. Insbesondere textverstehende Systeme mit einem gewissen Allgemeingrad bezüglich Form und Inhalt des zu analysierenden Textes sind noch weit von der praktischen Einsetzbarkeit entfernt, da sie die Formalisierung schwieriger linguistischer und semantischer Phänomene voraussetzen.

Bei Lernverfahren muß zwischen speziellen Verfahren, wie sie etwa im Bereich der Diagnostik zur Adjustierung von Evidenzwerten aufgrund neuer Falldaten existieren und solchen, die kausale Modelle zugrundelegen und dann innerhalb eines gegebenen Vokabulars Regeln in der Wissensbasis verändern oder ergänzen können, sowie sehr viel allgemeineren Verfahren wie etwa ID3 unterschieden werden. Bei ID3 ([Michalski et al. 84], [Gross 88], [Jackson 88])

handelt es sich um einen Algorithmus zur symbolischen Induktion, der jeweils eine Menge von Attribut-Wert-Listen verarbeitet, die Beschreibungen positiver und negativer Beispiele eines zu lernenden Konzeptes darstellen. Das Verfahren liefert zu einer Menge solcher Beispiele jeweils einen gerichteten azyklischen Graphen, dessen Knoten sich auf Attribute beziehen und mit dem weitere in gleicher Form vorliegende Beispiele nach dem vorher gelernten Wissen klassifiziert werden können. Auch die Darstellung des Ergebnisses als Regel ist in einfacher Weise möglich. Neuere Verfahren verbessern dieses Prinzip, indem z.B. von den initialen Symbolen der Beschreibungen durch Vorwissen über den Gegenstandsbereich abstrahiert wird und Komponenten zur Erkennung und Elimination von Störfaktoren integriert werden [Drastal et al. 89].

Die sinnvolle Anwendung dieser Verfahren der symbolischen Induktion stellt bestimmte Bedingungen an das Anwendungsgebiet. Dazu gehört, daß das Beschreibungsvokabular, auf dem die Falldaten aufbauen, ausreichend präzise und mächtig ist, so daß die Konzepte, die gelernt werden sollen, prädikatenlogisch darauf aufbauen können. Wo Wissen sich aber nicht auf diskrete sondern kontinuierliche Größen bezieht und, wie etwa im Anwendungsgebiet der vorliegenden Arbeit, oft durch numerische Ausdrücke und Vergleiche zwischen solchen Ausdrücken sinnvoll repräsentiert wird, können Verfahren der symbolischen Induktion - wenn überhaupt - nur nach aufwendiger Vorverarbeitung angewandt werden. Dort aber, wo sie anwendbar sind, werden sich automatische Verfahren aufgrund der Zeit- und Kostenersparnis in Zukunft sicher verstärkt durchsetzen.

Grundsätzliches Problem, neben der Eignung der Verfahren selbst, ist jedoch, daß Falldaten nicht immer in ausreichender Menge oder in nutzbarer Form vorliegen. Außerdem werden oft Regeln erzeugt, die zwar für das System zur im Sinne der Lerndaten korrekten Problemlösung genutzt werden können, die sich aber von dem vom Fachexperten tatsächlich angewandten Vorgehen unter Umständen erheblich unterscheiden. Dadurch ist die Wissensbasis schwer verständlich. Zusätzliche Funktionen, mit denen gelerntes Wissen kontrolliert und dokumentiert werden kann, sind dann erforderlich.

2.6 Direkter Wissenserwerb

Als Mittelweg zwischen dem aufwendigen Wissenserwerb mit Wissensingenieur und dem im wesentlichen nicht praxisreifen automatischen Wissenserwerb kann die Verwendung einer interaktiven Wissenserwerbskomponente direkt durch den Fachexperten betrachtet werden. Diese dritte Kategorie wird als direkter

Wissenserwerb bezeichnet. Zu dieser für die vorliegende Arbeit wichtigsten Kategorie werden zunächst einige in der Forschung wichtige Ansätze bzw. Systeme erläutert.

Als erstes System dieser Art gilt TEIRESIAS ([Davis 79] [Davis 82]). Es ermöglicht dem Experten im Bereich bakterielle Infektionskrankheiten die Verfeinerung der Wissensbasis von MYCIN. Prinzipielle Beschränkung ist dabei, daß Wissenserwerb immer ausgehend von Fehlern des aktuellen Systems stattfindet. Im Dialog kann der Benutzer von TEIRESIAS unter Verwendung der Erklärungskomponente den Ableitungspfad zurückverfolgen und eine falsche oder fehlende Regel identifizieren. Eine neue Regel kann dann in natürlichsprachlicher Formulierung eingetippt werden. Die Eingabe wird nach Schlüsselwörtern durchsucht und mit Schemata eines Regelmodells verglichen. Bei verschiedenen möglichen Inter-

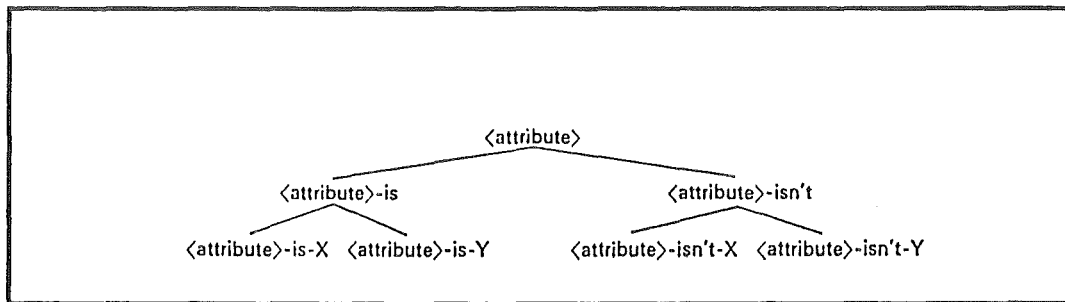


Abb. 2-3 Regelmodell in TEIRESIAS (nach [Davis 79])

pretationen ermittelt das System im Dialog mit dem Benutzer die gültige. Das eingesetzte Verfahren zum Verstehen von natürlicher Sprache ist sehr einfach und funktioniert nur, wenn bezüglich Struktur und Vokabular der Eingabe viele Annahmen zulässig sind. Die Abbildung 2-3 zeigt ein einfaches Regelmodell in TEIRESIAS.

Nach Eingabe einer Regel wird diese vom System auf Abweichung gegenüber dem Regelmodell geprüft und im Dialog können fehlende Aussagen ergänzt werden. Die Eingabeprozedur schließt damit ab, daß die neue Regel auf den Fall angewandt wird, von dem ausgehend die Eingabe in der oben beschriebenen Weise erfolgte. Damit wird geprüft, ob der Fehler tatsächlich behoben wurde.

Neues Wissen durch Rückverfolgung von Fehlern einzugeben, ist oft nicht der komfortabelste, manchmal ein gar nicht möglicher Weg. Da zudem die Eingabe für TEIRESIAS viel Tippaufwand des Benutzers erfordert, blieb es zwar als Forschungsanstoß wichtig, wurde aber nicht im tatsächlichen Wissenserwerb von Experten benutzt.

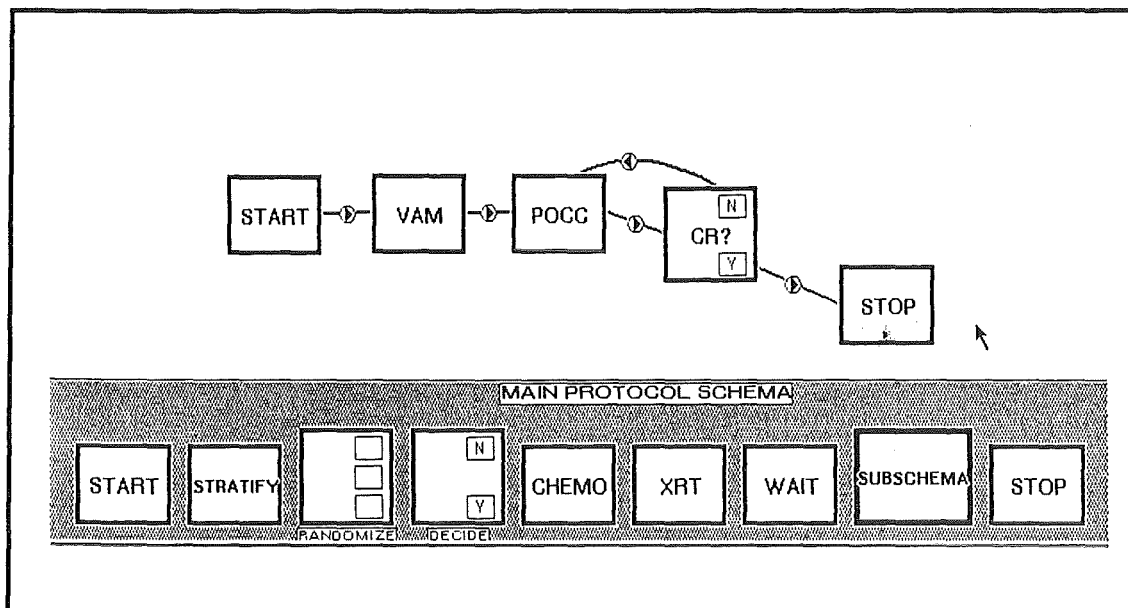


Abb. 2-4 Benutzeroberfläche des Wissenserwerbssystems OPAL, Eingabe von Therapieplänen (nach [Musen et al. 86])

Einen anderen Ansatz stellt das System OPAL, die Wissenserwerbskomponente von ONCOCIN, dar [Musen et al. 86]. Dieses System zum wissensbasierten Planen von Therapieabläufen bei Krebspatienten stammt ebenso wie MYCIN aus dem Labor für wissensbasierte Systeme der Medizininformatiker der Stanford Universität. OPAL ermöglicht Ärzten die Eingabe von Therapieplänen durch eine graphische, genau auf die Terminologie des Problems zugeschnittene Oberfläche. Durch diese Darstellung und die Vermeidung von Tastatureingaben erreicht OPAL eine hohe Benutzerfreundlichkeit und eine geringe Einarbeitungszeit. Die graphischen Einheiten repräsentieren Aktionen und Verzweigungen. Abbildung 2-4 zeigt die graphischen Einheiten, mit denen der Benutzer des Systems einen Therapieplan bearbeiten kann. Die in der Abbildung auftretenden Abkürzungen stellen für Ärzte verständliche Bezeichnungen von Therapieschritten bzw. den Beobachtungen dabei dar.

Die Verzweigungen sind mit Bedingungen versehen, die sich auf Laborwerte von Patienten oder mit Hilfe von Regeln daraus hergeleitetes Wissen beziehen. Auch die Eingabe dieser Regeln soll mit graphischen Hilfsmitteln unterstützt werden, wobei deren Implementierung noch nicht abgeschlossen ist, so daß es (bis Ende 1988) noch keine praktischen Erfahrungen mit diesem Teil der Wissenserwerbskomponente gibt.

Während es sich bei OPAL um eine auf ein ganz bestimmtes Problem spezialisierte Lösung handelt, gibt es einige allgemeinere Ansätze, welche die Merkmale eines Problemlösungstyps ausnutzen und direkten Wissenserwerb für Probleme dieses Typs ermöglichen wollen. Dies entspricht auch der in [Chandrasekaran 87] verallgemeinerten Hoffnung mit sogenannten 'generischen Problemlösungstypen' Aufgaben verschiedener Anwendungsgebiete, bei denen es Parallelen bei der angewandten Lösungsstrategie und der auftretenden Strukturen gibt, lösen zu können. Bislang finden sich Ansätze, die diesbezüglich erfolversprechend scheinen, ausnahmslos im Bereich des diagnostischen Problemlösens.

Arbeiten bei Boeing ([Boose 84], [Kitto et al. 87] und [Boose 88a]) gehen dabei von dem aus der kognitiven Psychologie stammenden Konstruktgitterverfahren zur Ermittlung von Diagnosen, Symptomen und ihrer Korrelationen aus. Das erste darauf aufbauende System für den Wissenserwerb war ETS. Es erfragte zunächst verschiedene Diagnosen, die zu einer Problemstellung gehören. Die zugehörigen Symptome werden nach dem Prinzip ermittelt, daß sie jeweils zwei Diagnosen von einer dritten unterscheiden helfen. In einem Konstruktgitter der Symptome und Diagnosen muß der Benutzer dann mit Gewichtungsfaktoren von 1 bis 5 die Korrelationen von einzelnen Diagnosen zu einzelnen Symptomen angeben, sowie die relative Wichtigkeit der Symptome. ETS erzeugt dann die Regeln zur Interpretation von Symptomen und ihre probabilistische Bewertung. Boose gibt an, daß über 500 Wissensbasen mit dem Werkzeug erstellt wurden [Boose 88a]. Dabei ist von Zeitdauern von z.T. zwei Stunden pro Wissensbasis die Rede, woraus gefolgert werden kann, daß es sich oft um triviale oder zumindest kleine Probleme gehandelt hat.

In AQUINAS, einer Weiterentwicklung, wird ETS um Hierarchien von Symptomen und Diagnosen, sowie um Frageformen für Multiple-Choice und für numerische Eingaben erweitert. [Puppe 88] weist daraufhin, daß bei ETS/AQUINAS auch spezifisch diagnostische Probleme, wie die Steuerung des Dialogs zur Symptomerfassung und die Feinheiten der Diagnosevorverarbeitung unberücksichtigt bleiben.

CLASSIKA ([Gappe 88a]), die Wissenserwerbskomponente von MED2, erlaubt erweiternd zu AQUINAS die Eingabe von zwei Werten im Symptom/Diagnosegitter, und zwar einen bezüglich des Vorhandenseins des Symptoms und einen bezüglich seines Nicht-Vorhandenseins. Außerdem sind den Tabelleneinträgen Regeln zugeordnet, die zusätzlich Bedingungen oder Ausnahmen enthalten können.

CLASSIKA ist speziell auf die Wissensarten der Diagnostikshell MED2 zugeschnitten. Die Komponente wird zur Zeit implementiert, so daß aus dem praktischen Einsatz noch keine Erfahrungen vorliegen.

Die Übersicht zeigt, daß die Idealvorstellung eines einfach handhabbaren und dennoch allgemeinen Werkzeugs für den Wissenserwerb von bisherigen Forschungsergebnissen nicht erfüllt wird. Erfolgversprechende Systeme bedienen sich der Ausdrucksmittel eines klar abgegrenzten Anwendungsgebietes (wie OPAL) oder bauen auf der gewichteten Korrelation von Diagnose-Symptom-Paaren auf, z.T. ergänzt mit Hilfsmitteln wie Frageklassen, Kontexte und Ausnahmen.

Dabei muß sich erst noch erweisen, wie breit der Bereich von Anwendungen ist, die dafür in Frage kommen. Daß die Systematik für diesen Forschungszweig insgesamt umstritten ist, wird u.a. daran deutlich, daß es wenige solcher systematischen Arbeiten gibt und der neueste dem Verfasser bekannte Versuch [Boose 88b] sich in der Diskussion jedoch nicht allgemeiner Zustimmung erfreuen konnte (vergleiche [Gappa 88b]).

2.7 KADS-Modell

Einen Anwendungsgebiet - und Problemlösungstyp übergreifenden Ansatz stellt KADS dar ([Wielinga et al. 86], [Schreiber et al. 88]). Die Abkürzung KADS stand ursprünglich für 'Knowledge Acquisition and Documentation Structuring'. Sie wird inzwischen auch als 'Knowledge Acquisition and Design Specification' verstanden.

<i>Layer</i>	<i>Elements</i>	<i>Relation with lower layer</i>
Strategy	Plans, meta rules	controls
Task	Goals, task structures	applies
Inference	Meta classes, knowledge sources	describes
Domain	Concepts, relations	

Abb. 2-5 Ebenen des KADS-Modells

Das zur Lösung von Aufgaben eines Bereiches erforderliche Wissen wird entsprechend seiner Funktion einer der vier Ebenen 'domain level', 'inference level', 'task level' und 'strategic level' zugeordnet. Jeweils höhere Ebenen verwenden das Wissen der tieferen Ebenen zur Lösung ihrer Aufgaben.

Die 'domain'-Ebene bildet die unterste Ebene und entspricht dem statischen Wissen innerhalb eines Problembereichs, das dem Lösungsprozeß zugrunde liegt. In wissensbasierten Systemen wird das Wissen dieser Ebene durch Konzepte wie Frames, Relationen oder andere statische Strukturen dargestellt.

Die zweite Ebene - 'inference level' - beschreibt mit 'Metaklassen' und 'Wissensquellen', welche Rolle die Konzepte auf unterster Ebene im Problemlösungsprozeß spielen. Inferenzen basieren dabei durch Generalisation und Verfeinerung auf den Relationen der 'domain'-Ebene und sind nicht mit dem auch als 'Inferenz' bezeichneten konkreten Schlußfolgerungsmechanismus eines wissensbasierten Systems zu verwechseln. Die 'task-level' enthält das Wissen, wie durch Kombination der einzelnen 'Wissensquellen' bestimmte Ziele erreicht oder verschiedene Aufgaben erledigt werden können. Die oberste Ebene - 'strategic-level' - enthält das strategische Wissen. Damit ist gemeint, wie durch Planung, Aktivierung und Überwachung der 'tasks' der darunterliegenden Ebene ein Problem gelöst wird. Da nicht viele wissensbasierte Systeme nach mehr als einer einzigen festen Strategie arbeiten, enthält diese Ebene meist nur den Aufruf einer 'task'

Der Ansatz ist sowohl für das Erarbeiten einer Forschungssystematik im Bereich Wissenserwerb als auch für das Software Engineering im allgemeinen interessant. Das Projekt, begonnen als ESPRIT-Projekt der Universität Amsterdam mit industriellen Partnern und nun bei der GMD in Sankt Augustin fortgeführt, hat jedoch nichts - und wird dies nach Aussagen Beteiligten auch in naher Zukunft nicht - ergeben, was auf Rechnern zum Zwecke des direkten Wissenserwerbs einsetzbar wäre.

2.8 Hybride Entwicklungsumgebungen

Die gängigen hybriden Entwicklungsumgebungen für wissensbasierte Systeme stellen zwar eine Vielzahl von Formalismen zur Repräsentation und Anwendung von Wissen zur Verfügung, erfüllen aber nicht die Forderung, direkt durch einen Fachexperten nutzbar zu sein. Die Differenz zwischen den Techniken und der Terminologie jeweils im Anwendungsgebiet und in den Entwicklungsumgebungen ist zu groß.

Erst einmal muß der Benutzer die Syntax und Semantik des Repräsentationsformalismus beherrschen. Daneben gibt es eine Reihe weiterer Dinge, deren Verständnis Voraussetzung für die Benutzung ist. Dazu gehören die Wirkungsweise des Schlußfolgerungsmechanismus, mögliche logische Verknüpfungen, die Verwendung von Variablen und die Ausdrucksmöglichkeiten für das Pattern Matching und ihre Semantik. Konkrete Probleme treten außerdem für den Fachexperten schon auf, sobald Voraussetzung für die Benutzung einer Entwicklungsumgebung das Wissen über Funktionen im Editor, im Dateisystem oder über Schnittstellen zu Sprachen oder anderen Werkzeugen ist. Auch bestimmte Dinge, die ein Systementwickler im "Hinterkopf" hat, gehören dazu, wie die Verwendung von Namen und ganz allgemein das Wissen über die Lokalisierung von Wissen. Diese Probleme sind gleichzeitig Anforderungen für die Entwicklung komfortabler Wissenserwerbskomponenten. Nur solche Komponenten, die ihre Benutzer von aus der Kenntnis des Anwendungsgebietes heraus nicht verständlichen Mechanismen entlasten, werden tatsächlich einsatzfähig sein.

Dies ist bei den gegenwärtig gängigen hybriden Entwicklungsumgebungen für wissensbasierte Systeme nicht der Fall. Wünschenswert für zukünftige Generationen dieser Entwicklungssysteme wäre, daß sie durch Integration von Bausteinen für den Wissenserwerb wie etwa generische Funktionen zur Darstellung und Modifikation von Wissen die Entwicklung spezifischer Komponenten für den Wissenserwerb als Teil der Systementwicklung besser unterstützten.

2.9 Schlußfolgerungen

Der automatische Wissenserwerb bietet nur in bestimmten Problembereichen und bei Verfügbarkeit entsprechend strukturierter Lerndaten erfolgversprechende Perspektiven. Insbesondere für die Aufgabe der vorliegenden Arbeit kommt er daher nicht in Frage.

Der Wissenserwerb mit Wissensingenieur ist oft die einzig mögliche, allerdings aufgrund des Aufwandes keine langfristig tragfähige Lösung.

Bei den durch den Wissensingenieur durchzuführenden systemanalytischen Aufgaben, wie etwa dem Erkennen der Problemstellung und der Wissensquellen, ist eine Automatisierung nur schwer vorstellbar. Auch die Konzeptionalisierung wird, außer bei wenigen Problemen, welche die sofortige Zuordnung von Kernkonzepten eines Problemlösungstyps zu Begriffen des Anwendungs-

gebietes erlauben, eine wichtige vom Wissensingenieur durchzuführende Tätigkeit bleiben.

Die Forschung auf dem Gebiet der Systeme des direkten Wissenserwerbs kann bislang keinen Durchbruch im Sinne eines übergeordneten - d.h. nicht für einen bestimmten Problemlösungstyp spezifischen - Ansatzes verzeichnen. Warum zum Thema der vorliegenden Arbeit die Konzepte der - in diesem Gebiet fortgeschrittensten - wissensbasierten Diagnostik nicht einsetzbar sind, wird u.a. im Abschnitt 4.6 ausgeführt.

Ein eigener Ansatz einer interaktiven Komponente für den direkten Wissenserwerb - auf Spezifika des Anwendungsgebietes bezogen - und seine Realisierung werden in den weiteren Kapiteln der Arbeit erläutert.

3. Altlasten

3.1 Einleitung

In diesem Kapitel wird das Anwendungsgebiet Altlasten kurz vorgestellt und eine allgemeine Einführung in das Vorgehen von Fachexperten in diesem Gebiet gegeben. Die Darstellung des Wissens und prinzipieller Vorgehensweisen in diesem Bereich dient als Grundlage für Entscheidungen beim Entwurf der Wissenserwerbskomponente, die dann in späteren Kapiteln ausgeführt werden.

3.2 Begriffe

Bei der Altlastenproblematik geht es um die Gefährdung von Mensch und Umwelt, die von Altablagerungen (Ablagerungen von häuslichen, industriellen oder gewerblichen Abfallstoffen) oder Altstandorten (ehemalige Industrie- oder Gewerbeflächen) mit Produktionsrückständen ausgehen kann. Mit dem Begriff "gefährverdächtige Flächen" werden Altablagerungen und Altstandorte zusammengefaßt.

In [BaWü 87] wird von "Altlast" gesprochen, wenn "nach fachlicher Beurteilung durch die zuständigen Behörden begründeter Verdacht besteht, daß von solchen Standorten eine Gefahr für die öffentliche Sicherheit und Ordnung oder Gefahren bzw. Beeinträchtigungen für die menschliche Gesundheit oder die Umwelt ausgehen können". Der Einfachheit halber wird in der vorliegenden Arbeit der Begriff Altlasten - wie auch überwiegend in der einschlägigen Literatur - im allgemeineren Sinne verwendet und damit auch auf altlastverdächtige Flächen bezogen.

Bei der Gefährdung der Umwelt werden die Schutzgüter Grundwasser, Oberflächenwasser, Boden und Luft unterschieden. In diesem Zusammenhang wird auch vom Gefährdungspfad gesprochen, also dem Weg, auf dem ein gefährlicher Stoff zum Schutzgut gelangt.

Aus der zusammenhängenden Betrachtung der Gefährlichkeit einzelner auf einer Fläche vorkommenden Stoffe ergibt sich die sogenannte Stoffgefährlichkeit. Die Gefährdung der Schutzgüter ist von dieser Stoffgefährlichkeit und von den örtlichen Verhältnissen abhängig.

Unterschieden wird auch zwischen der Bewertung und der Beurteilung einer Altlast. Dabei bezieht sich die Bewertung auf das quantifizierte Risiko, z.B. in Baden-Württemberg die Maßzahl r_0 für das sogenannte Risiko in Vergleichslage, während die Beurteilung ein umfassendes Gutachten darstellt, welches auch Aspekte eines konkreten Falles umfaßt, die einer solchen Quantifizierung nicht zugänglich sind.

Auf die Beurteilung und Bewertung einer Altlast müssen ggf. Maßnahmen zur Gefahrenminderung folgen. Dies kann eine Sanierungs- oder Sicherungsmaßnahme sein. Sanierungsmaßnahmen definiert [BaWü 87] als "Beseitigung oder Verringerung der Schadstoffmenge am Gefahrenherd und im kontaminierten Umfeld", während es bei der Sicherung um "Verhinderung oder Verringerung - ggf. vorübergehend - des Schadstoff-Austrags und damit auch des Eintrags in das Schutzgut" geht. In diesem Zusammenhang weist [Schenkel 86] darauf hin, daß "Vorbeugen in der Regel billiger als Sanieren ist" und diskutiert politische und wirtschaftliche Aspekte dieses Problems. Nach Ansicht des Verfassers der vorliegenden Arbeit greift hierbei eine rein kostenorientierte Betrachtung zu kurz.

3.3 Bedeutung der Altlastenproblematik

Viele Altlasten-Fälle gehen auf frühe Phasen der Industrialisierung zurück, in denen keiner oder nur wenige sich damit beschäftigten, was später mit Produktionsrückständen geschieht, woraus leider nicht gefolgert werden kann, daß dies heute immer ausreichend der Fall ist. Bekannte Beispiele für Altlasten (z.B. Hamburg-Georgswerder) wurden erst bei ersten gesundheitlichen Folgen oder Problemen in der Trinkwasserversorgung als Problem erkannt. Inzwischen haben Initiativen der zuständigen Stellen auch zu systematischen Untersuchungen zur Auffindung von Verdachtsflächen geführt.

Die Zahl möglicher mit Gefährdung verbundener Flächen in der Bundesrepublik Deutschland gibt das Umweltbundesamt mit 37000 an, davon 4600 in Baden-Württemberg. Schätzungen über das Volumen voraussichtlicher Sanierungs- und Überwachungskosten liegen für den Bund bei 17 Mrd. DM [Franzius 86].

In [Döppert 88] wird eine Studie zitiert, die das Marktvolumen für Altlastensanierung in der Bundesrepublik bis zum Jahr 2000 sogar mit 29 Mrd. DM angibt. Das weist auch auf den erheblichen Bedarf an sachgerechter Bewertung und Beurteilung hin.

3.4 Sanierung

Bei der Sanierung können je nach Ausgangslage und anwendbaren Verfahren verschiedene sogenannte Sanierungsziele verfolgt werden :

- Wiederherstellung der universellen Verwendbarkeit,
- Abwehr akuter Gefahren,
- Verringerung der Schadstoffbelastung auf ein vorgegebenes Maß,
- Ablösung von Kontrollen,
- Zeitgewinn, bis ein geeignetes Sanierungs- oder Sicherungsverfahren entwickelt oder verfügbar ist, oder
- Ermöglichung einer bestimmten Nutzung des Standortes
(nach [BaWü 87]).

Dementsprechend lassen sich auch die dafür einsetzbaren Techniken einteilen.

Als Maßnahmen zur Sicherung kommen in Frage :

- Oberflächenabdichtung
- Einkapselung
- Senkung des Grundwasserspiegels
- Auskoffnung und Zwischenlagerung

Als Maßnahmen zur Sanierung lassen sich unterscheiden:

- Mikrobiologische Verfahren
- Bodenluftabsaugung
- Extraktionsverfahren (chemisch oder physikalisch)
- Fixierungsverfahren
- Thermische Verfahren
- Auskoffnung und Endlagerung
(nach [Döppert 88])

Abbildung 3-1 zeigt den Einfluß gewählter Prinzipien auf die Verminderung der Gefährlichkeit.

Da es in der vorliegenden Arbeit vorrangig um das Wissen im Bereich Beurteilung und Bewertung von Altlasten geht, sei an dieser Stelle auf Übersichten und interessante Fallstudien aus der Fachliteratur verwiesen ([BaWü 87], [Döppert 88], [Franzius 87], [Freeman 89], [Karczmarzyk et al. 88], [Schenkel 86], [Schuldt 87]).

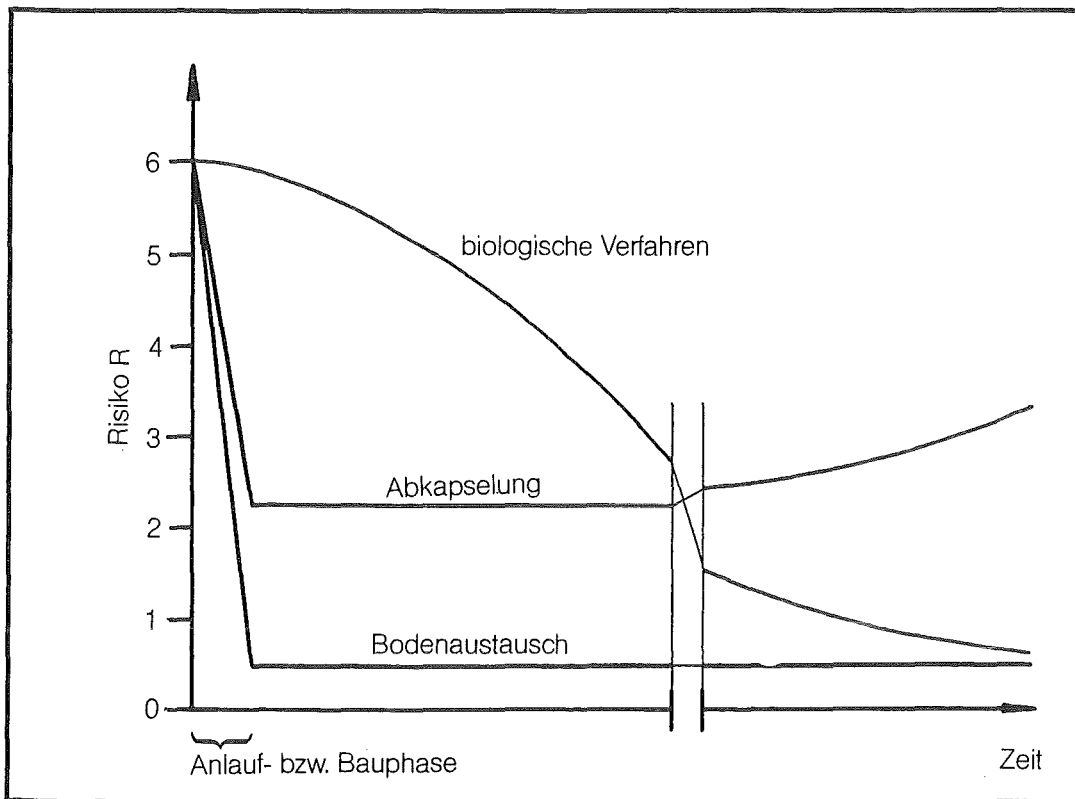


Abbildung 3-1: prinzipielle Sicherungs- und Sanierungsverfahren und bewirkte Gefahrenminderung (nach [BaWü 87])

3.5 Beurteilung und Bewertung von Altlasten

Vorrangiges Ziel der Beurteilung und Bewertung von Altlasten ist es,

- mögliche akute Gefahren zu erkennen, um ihnen begegnen zu können,
- Erkenntnisse über angezeigte bzw. nicht angezeigte Nutzungen von Flächen zu gewinnen und
- Prioritäten bei der weiteren Erkundung und ggf. Sanierung unter einer Anzahl von Flächen setzen zu können.

Fachbeiträge weisen immer wieder darauf hin, daß der Stand der relevanten Wissenschaften (Teilgebiete der Toxikologie, Chemie, Physik, Biologie, Geologie u.a.) zwar Aussagen über einzelne betrachtete Stoffe oder Stoffverbindungen erlaubt, aber die Ermittlung absoluter Kenngrößen für die Bewertung von Altlasten einen nicht gegebenen Objektivitätsgrad andeutet. Nötig ist jedoch die vergleichende Beurteilung nach dem jeweilig aktuellen Wissensstand. Diese Beurteilung sollte aufgrund ihrer wirtschaftlichen und rechtlichen Relevanz nach möglichst einheitlichem Vorgehen erfolgen und auf die oben genannten Ziele

hin ausgerichtet sein. Die Identifizierung von Verdachtsflächen erfolgt oft durch Hinweise von Betroffenen, inzwischen auch durch systematische Suche. Hinsichtlich der Luftekundung zu dieser Frage sei auf [Kinner et al. 86] verwiesen.

Vorarbeiten zur Beurteilung und Bewertung einer Altlast sind das Zusammenstellen von Informationen über die Fläche, bei einem Altstandort etwa Informationen über die Branche oder sogar über die konkreten Produktionsanlagen und ihre Bestandteile. Diese Informationen können z.B. dazu dienen, möglichst geeignete Analysenparameter für die chemisch-physikalischen Untersuchungen auszuwählen und so die vorhandenen Mittel zugunsten einer optimalen Beurteilungsgrundlage einzusetzen. Neben Datensammlungen über branchentypische Verunreinigungen stützt sich der Fachexperte dabei stark auf Erfahrungswissen aus früheren vergleichbaren Fällen.

Wird die Bewertung ausschließlich auf Daten dieser Art aufgebaut und nicht auf Ergebnissen einer Beprobung der Verdachtsfläche, so spricht man von dem, in [Kinner et al. 86] systematisch aufgearbeiteten, "historisch-deskriptiven Ansatz".

Für eine umfassende Beurteilung und Bewertung sind allerdings chemisch-physikalische Analysen von auf der Verdachtsfläche genommenen Proben unverzichtbare Grundlage. Für diese Analysen ist die Erstellung eines spezifischen Analysenplanes notwendig. Aus Kostengründen sollen dabei nicht alle eventuell interessanten Analysen von vorneherein erhoben werden, sondern es werden iterativ weitere Analysen durchgeführt, wenn die Beurteilung früherer Analysen desselben Falles dies nahelegen. Die zusammengestellten Vorinformationen sollen auf diese Weise helfen, die zur Verfügung stehenden Mittel optimal einzusetzen. Die Bearbeitung eines Altlastenfalles führt dann über die Berücksichtigung der örtlichen Verhältnisse und der Nutzung zur Entscheidung über eventuell nötige Sicherungs- und Sanierungsmaßnahmen.

Zur Vereinheitlichung und besseren Transparenz des gesamten Vorgehens bei der Altlastenbewertung wird in [BaWü 87] eine Gliederung in definierte Stufen vorgeschlagen, die mit ihren jeweiligen Ergebnissen im folgenden aufgeführt sind :

Die erste Stufe - Stoffgefährlichkeit in Vergleichslage - liefert den r_0 -Wert, eine Zahl zwischen 0 und 6 oder bei Unklarheiten auch ein Intervall aus diesem Bereich. In dieser Stufe wird von Besonderheiten der Umgebung und der Nutzung zunächst abstrahiert. Die "Vergleichslage" ist dabei eine nach Einrichtung und Unterhaltung dem Deponiemerkblatt der Länderarbeitsgemeinschaft-Abfall (LAGA) vom November 1978 entsprechende Hausmülldeponie. In

einer zweiten Stufe - Berücksichtigung der örtlichen Verhältnisse - werden in drei Verfahrensschritten die Wert r_1 bis r_4 ermittelt. Dabei werden der Austrag gefährlicher Stoffe von der Verdachtsfläche, der Eintrag in ein Schutzgut und der Transport und die Wirkung innerhalb des Schutzgutes berücksichtigt.

Durch Zu- bzw. Abschläge zum Risiko in Vergleichslage ergibt sich das als "tatsächliche Risiko" bezeichnete Ergebnis der zweiten Stufe. Hier liegt auch der Vorteil dieses Vorgehens, denn die Abschätzung des Risikos von Schadstoffaustrag, -eintrag und -transport erfolgt relativ zur Vergleichslage und nicht absolut und ist damit in der Regel einfacher durchführbar.

Eine Gewichtung entsprechend der Bedeutung des Schutzgutes fließt in einer dritten Stufe mit ein und führt schließlich zum "gewichteten Risiko".

Zu dem Ergebnis einer Stufe ist jeweils ein Beweisniveau - eine Zahl zwischen 0 und 4 - anzugeben, je höher das Beweisniveau, desto eingehender die in dieser Stufe durchgeführten Erkundungen.

3.6 Wissen im Altlastenbereich

Wissen im Bereich Altlasten läßt sich - neben der Erkundung und den rechtlichen und politischen Fragen - den nicht ganz voneinander trennbaren Bereichen Bewertung und Sanierung zuordnen. Dabei gibt es u.a. physikalische, chemische, toxikologische und geologische Aspekte. Das Wissen besteht dabei zu einem großen Teil aus Erfahrungsregeln, z.B. unter welchen Umständen welche Stoffkonzentrationen wie zu beurteilen sind. Es gibt die schon genannten Erfahrungswerte über den Zusammenhang von Produktionsanlagen und typischen Verunreinigungen. Aus Erfahrungen heraus können Farbe, Geruch, pH-Wert und andere Parameter bezüglich verschiedener Untersuchungsobjekte wie Wasser, Boden, Eluat interpretiert werden.

Aufgrund der Vielfalt möglicher Altlasten-Fälle wäre eine einfache gesetzliche Regelung etwa in Form einer Grenzwerte-Liste nicht angemessen. Stattdessen ist es Aufgabe der Bewertung in Abhängigkeit von Untersuchungsobjekt und auftretenden Stoffen geeignete der vielen Grenzwerttabellen aus dem Umweltbereich miteinzubeziehen. Bekannte und wichtige Tabellen sind etwa die Kloke-Liste, die Trinkwasserverordnung oder der Niederländische Leitfaden zur Bodensanierung. Aber auch Grenzwerte anderer Tabellen werden betrachtet. Das ist insbesondere erforderlich, wenn es sich um Spezialfälle (selten vorkommende Stoffe) handelt. Bestimmte Größen - der organische Anteil einer Probe beispielsweise - werden aus Analysenergebnissen abgeleitet, was je nach

Ausgangsdaten durch einfache Näherungsformeln oder nur durch Erfahrungswerte geschieht.

Wissen im Bereich Altlasten umfaßt auch Aussagen zur Repräsentativität einer Beurteilung, Hinweise auf eine mögliche Veränderung von Proben seit der Probennahme und eine Vielzahl weiterer Anmerkungen, die bestimmte Ausgangssituationen interpretieren. Nach den Erfahrungen im bisherigen Projektverlauf ist es möglich, einen großen Teil auch dieses Wissens als Regeln zu formulieren.

Die Expertise zu einem Altlasten-Fall, die dann als Grundlage für weitere Entscheidungen dient, besteht zunächst aus einer groben Einstufung der einzelnen Proben und des gesamten Falles entsprechend ihrer Umweltgefährlichkeit. Auf die in Baden-Württemberg eingeführten r_i -Werte wurde schon hingewiesen. Doch diese Grobeinstufung wird mit relevanten Details ergänzt, die die Expertise erst wirklich aussagekräftig machen. Dies kann der weitere Erkundungsbedarf bezüglich ganz bestimmter Stoffe, Hinweise auf relevante Abweichungen zwischen verschiedenen Analysen eines Falles, die genaue Spezifikation, wovon die mögliche Gefährdung abhängt oder die Einstufung der Repräsentativität der Analysen sein. Insgesamt geht es um relevante Details oder Anmerkungen, die die Interpretation eines Falles vervollständigen.

4. Das Altlasten-System XUMA

4.1 Ziele

XUMA (Expertensystem zur Beurteilung der Umweltgefährlichkeit von Altlasten) ist das System, das im Rahmen des eingangs erwähnten gemeinsamen Projektes des Instituts für Datenverarbeitung in der Technik (IDT) des Kernforschungszentrums Karlsruhe und des Instituts für Altlastensanierung der Landesanstalt für Umweltschutz (LfU) erstellt wird. Ziel ist es, das Wissen zur Altlastenbeurteilung, das es bei den wenigen Fachexperten der LfU gibt, mit Hilfe eines wissensbasierten Systems für Sachbearbeiter, zunächst in der LfU und in einem späteren Schritt auch in den Wasserwirtschaftsämtern, zugänglich zu machen. Damit werden die Experten von Routinetätigkeiten entlastet, für die Sachbearbeiter bietet das System eine wichtige Assistentenfunktion, und das Vorgehen der Altlasten-Beurteilung wird insgesamt vereinheitlicht.

4.2 Aufbau

Das System unterstützt die folgenden vier Phasen bei der Bearbeitung von Altlastenfällen :

- Erstellen eines Analysenplans
- Erfassen von Analysen
- Beurteilung
- Bewertung

Hierbei wird unter der Beurteilung eine Stellungnahme in der Art eines Gutachtens - vergleiche letztes Kapitel - verstanden, während es bei der Bewertung um die Ermittlung eines die Umweltgefährlichkeit repräsentierenden Zahlenwertes geht. Außerdem gibt es die sogenannten 'Expertenfunktionen'. Diese geben einmal autorisierten Fachexperten die Möglichkeit die Wissensbasis zu pflegen und zu erweitern. Auch sollen sich Benutzer über das im System vorhandene Wissen unabhängig von der Bearbeitung eines Falles erkundigen können, wenn dies für ihre tägliche Arbeit Vorteile bietet. Eine Funktion zur Stofffassung ist dort bereits realisiert. Weitere Funktionen, auch die in der vorliegenden Arbeit realisierten, werden diese Schnittstelle vervollständigen.

Die Phasen der Fallbearbeitung sind menüorientiert aufgebaut. Dabei können einzelne Phasen ausgewählt werden und müssen bei der konkreten Fallbearbeitung auch nicht streng sequenziell abgearbeitet werden. Bei neuen

Erkenntnissen oder geänderten Randbedingungen gibt es immer wieder einen Rückfluß von Informationen in vorhergehende Phasen, woraufhin bestimmte Schritte wiederholt werden können und die Beurteilung ggf. geändert wird. Das Ergebnis kann in unterschiedlich detaillierter Form ausgegeben werden. Neben einer Beurteilung des gesamten Falles, der die Interpretation zusammenfaßt, gibt es auch die Möglichkeit die Beurteilung einzelner Proben oder Analysen anzufordern. Außerdem wird eine Fallstatistik aufgestellt, die die Einstufung von Meßwerten bezüglich einzelner Analysenparameter in Qualitätsklassen (als Maß für die Umweltgefährlichkeit) übersichtlich darstellt.

4.2.1 Analysenplanerstellung

In der Phase Analysenplanerstellung kann sich der Benutzer zunächst über drei Zugänge einen Analysenplan erstellen lassen. Dieser Plan besagt dann, bei welcher Art von Probe welche Analysenparameter zu bestimmen sind. Die drei Zugänge sind :

- Herkunft (Branche oder Anlage)

Dieser Zugang wird gewählt, wenn aus der Vorerkundung der Altlast bereits Informationen über die frühere Nutzung der Verdachtsfläche vorliegen. Sowohl sehr allgemeine Anhaltspunkte wie die Branche des betreffenden Produktionsbetriebes als auch genauere Angaben etwa aus Lageplänen, auf denen bestimmte Anlagen oder Anlageteile bezeichnet sind, können hier verwertet werden.

- Stoff (Stoffe, Stoffgruppen, Abfallarten)

Gibt es Anhaltspunkte für bestimmte Stoffe, so werden danach die geeignetsten Analysenparameter ausgewählt. Durch den Begriff Stoffe werden im System so verschiedene Substanzen wie einzelne Schwermetalle, komplexe Verbindungen und auch Abfallarten bezeichnet. Die bezüglich bestimmter Stoffe geeigneten Analysenparameter werden über ein semantisches Netz abgeleitet. Dabei helfen geeignete Vererbungsmechanismen bei der Auswertung der Zusammenhänge von Stoffen und ihren Komponenten zu Analysenparametern.

- Standard

Sind weder Vorinformationen über die Herkunft noch über Inhaltsstoffe zu einer Verdachtsfläche bekannt und deshalb keiner der beiden genannten

Zugänge möglich, so kann auf einen in zwei Detaillierungsstufen definierten Standardanalysenplan zurückgegriffen werden.

4.2.2 Erfassung von Analysen

Als mögliche Analysen zu einem Altlastenfall sieht XUMA vor :

- Brutto-Analysen von Abfall- oder Bodenproben,
- Analysen des wäßrigen Eluats einer Abfall- oder Bodenprobe,
- Wasser-Analysen von Grund-, Sicker- und Oberflächenwasserproben und/oder
- Gas-Analysen von Deponiegas, Bodenluft/-gas und atmosphärischer Luft.

Eine typische Analyse umfaßt 5 bis 20 Werte. Von den über 300 dem System bekannten Analysenparametern sind die meisten quantitativer Art, erfordern also numerische Eingaben. Daneben gibt es eine Reihe qualitativer Parameter wie Trübung, Geruch, Farbe, Konsistenz u.a. Durch entsprechende Formulare unterstützt das System die Erfassung der zugehörigen Werte. Die erfaßten Werte werden einer Plausibilitätskontrolle unterzogen, die regelbasiert abläuft.

4.2.3 Beurteilung

Diese Phase umfaßt zunächst die Beurteilung der Stoffgefährlichkeit in Vergleichslage. Dazu werden die Analysenergebnisse zunächst auf Qualitätsklassen - ein qualitatives Maß für die Gefährlichkeit - abgebildet. Sowohl die Auswahl geeigneter Tabellen als auch die Anwendung der Grenzwerte wird durch Regeln gesteuert. Die Schutzgüter Boden, Luft, Oberflächen- und Grundwasser werden getrennt betrachtet. Für die einzelnen Proben, zu denen es in der Regel mehrere Analysen gibt, und schließlich für den gesamten Fall, werden die Ergebnisse durch Anwendung entsprechender Regeln zusammengefaßt. Die Analysendaten werden außerdem nach individuellen, oft auf Erfahrungen mit früheren Fällen zurückgehenden Regeln interpretiert.

In einer, in der aktuellen Version noch nicht eingeschlossenen Phase 'Berücksichtigung der örtliche Verhältnisse' werden die Lage, geologische und hydrologische Eigenschaften und weitere Besonderheiten in die Beurteilung miteinbezogen.

4.2.4 Bewertung

In dieser Phase wird die Maßzahl r_0 für das Risiko in Vergleichslage sowie weitere Kennzahlen r_1, r_2 etc., welche die entsprechenden Risiko- und abschlüsse bei Berücksichtigung der örtlichen Verhältnisse mit einbeziehen. Die Phase ist im Rahmen der aktuellen Version von XUMA nur insofern realisiert, als ein Bereich für r_0 ermittelt wird.

4.3 Implementierung

XUMA wird auf einem Explorer II, einer Lisp-Maschine der Firma Texas Instruments, implementiert. Dabei sind die Entwicklungsumgebung ART und das relationale Datenbanksystem RTMS eingesetzt.

Bestimmte Funktionen etwa im Bereich der Analysenplanerstellung, zur Darstellung der Ergebnisse und für die Erklärungskomponente wurden direkt in Common Lisp implementiert.

4.4 Das Werkzeug ART

4.4.1 Überblick

ART steht für 'Automated Reasoning Tool' und ist eine hybride Entwicklungsumgebung der Firma INFERENCE. Nach einer ersten Version, die 1982 auf den Markt kam, gibt es inzwischen nach Erweiterungen und Verbesserungen die Version 3.0.

ART ist nach der Blackboard-Architektur konzipiert, die durch HEARSAY-II, einem System zum Verstehen natürlicher Sprache, bekannt wurde. Wichtiges Prinzip dieser Architektur ist die Verwendung einer Agenda, mit der die Anwendung von Regeln organisiert wird. Die Architektur ist im wesentlichen für Vorwärtsregeln ausgelegt. Rückwärtsregeln, die ART auch kennt, werden intern auf Vorwärtsregeln abgebildet. Die Agenda enthält die Regeln, deren Prämissen zum jeweiligen Zeitpunkt in der Wissensbasis erfüllt sind. Die Regeln werden in der Reihenfolge ihrer Priorität abgearbeitet. Die Anwendung führt zu neuen Fakten, die wiederum die Prämissen anderer Regeln erfüllen können, die dadurch anwendbar werden.

Zur Darstellung von Objekten und Strukturen können in ART Schemata verwendet werden, die in der einschlägigen Literatur und in vielen anderen Systemen als Frames bekannt sind.

4.4.2 Regeln in ART

Regeln in ART bestehen aus Namen, Kommentar, Prämissen- und Aktionsteil. Der Prämissenteil besteht aus Mustern ('pattern'). Der Regelinterpreter vergleicht diese mit den Fakten der Wissensbasis ('match'). Matchen bedeutet dabei auch die Instanziierung der in den Mustern auftretenden Variablen mit den konkreten Werten, welche die entsprechenden Fakten enthalten. Bei größeren Wissensbasen enthält der Prämissenteil auch Muster, mit denen die Regel einer Phase zugeordnet wird. Damit ist die Bildung von Regelpaketen möglich. Innerhalb einer Phase steuert dann die Priorität ('saliency') einer Regel nach dem oben erwähnten Agendamechanismus die Abarbeitung.

Der Aktionsteil kann das Einfügen von neuen Fakten in die dynamische Wissensbasis enthalten, die Zurücknahme von Fakten, oder auch den Aufruf von Lisp-Funktionen.

Bei der Übersetzung der Wissensbasis wird eine interne Struktur der Regeln erzeugt, das sogenannte Pattern-Join-Netz, mit dem der Regelinterpreter zur Laufzeit effizient die jeweils für die Agenda in Frage kommenden Regeln untersuchen kann.

4.4.3 Schemata in ART

Muster in Regeln können sich auf Schemata beziehen, die - wie oben erwähnt - den Frames anderer Systeme entsprechen. Jedes Schema besteht aus einem identifizierenden Namen und einer Anzahl von Slots. Das gesamte Schema-System in ART basiert auf wenigen Kern-Schemata, die sehr allgemein definiert sind. Schemata und Slots, die vom Benutzer definiert werden, sind jeweils Instanzen dieser Kern-Schemata.

Der Slot eines Schemas hat verschiedene Eigenschaften ('Facets'), die vom ART-Benutzer definiert werden können. Solche Eigenschaften sind z.B. die Anzahl möglicher Slotwerte ('single-value' oder 'multiple-value') oder sagen etwas über die Vererbung von Werte aus. Sogenannte semantische Netze entstehen, wenn Schemata durch Relationen verbunden werden. Diese Relationen sind frei definierbar und werden als Slots in Schemata sichtbar. Grundlegende Relationen sind die zur Definition von Spezialisierung und von Zugehörigkeit zu allge-

meineren Konzepten: 'has-instances' und 'is-a'. ART stellt fünf grundlegende Relationen zur Verfügung, mit denen Klasse-Instanz- oder die Menge-Untermenge-Beziehungen repräsentiert werden können. Zusätzlich kann es frei definierte Relationen geben. Es können Abhängigkeiten zwischen Relationen definiert werden. Beim Einfügen einer Beziehung werden dann automatisch die logisch abhängigen Beziehungen eingetragen.

Es gibt nicht nur Relationen, die zwischen Schemata definiert werden. Auch allgemeine Relationen zwischen beliebigen Werten (Symbolen, Zahlen usw.) können in ART definiert werden. Solche Relationen können als transitiv, reflexiv oder symmetrisch sein. ART sorgt dann für die Einhaltung dieser Eigenschaften, indem es automatisch Rückwärts-Regeln dazu generiert.

Der Vererbungsmechanismus von ART stellt eine Verbesserung zu dem historisch wichtigen - von Smalltalk bekannten - allgemeinen Ansatz dar. Es kann definiert werden, welcher Slotwert übernommen werden soll, wenn dieser Slot von mehreren Schemata vererbt wird ('multiple inheritance').

In ART kann das 'active-value'-Konzept realisiert werden. Nach dem Einlesen, Auslesen oder Verändern von Slotwerten folgt jedesmal der automatische Aufruf einer Lisp-Funktion. So können Trigger-Mechanismen implementiert werden, wie sie etwa für die Umrechnung von Slotwerten zwischen verschiedenen Bezugsgrößen hilfreich sind. Beliebte Beispiel für 'active-values' sind Grafiken, wie Tachometer oder Schalter, welche den jeweils aktuellen Wert des Slots anzeigen. Damit kann in bestimmten Anwendungen eine komfortable Oberfläche realisiert werden.

Eine weitere Möglichkeit ist die Definition von Methoden - in ART Aktionen genannt - zu einem Schema. Der Mechanismus ist allerdings sehr einfach und bietet beispielsweise nicht die Möglichkeiten zur Methodenkombination, wie sie im Common-Lisp-Flavor-System möglich sind.

Insgesamt stellt das Schema-System ein flexibles und mächtiges Hilfsmittel zur Repräsentation objektorientierten Wissens dar, das sich mit den Regeln gut kombinieren läßt.

4.4.4 Bewertung von ART

ART stellt die Formalismen zur Verfügung, die von den anderen gängigen Entwicklungsumgebungen in vergleichbarer Form bekannt sind. Die Stärke von ART liegt in der Mächtigkeit des Pattern Matching und im Laufzeitverhalten. Dies kommt insbesondere bei Systemen zum Tragen, die aus einer breiten

Menge von Fakten in jeweils relativ flachen Ableitungsbäumen Ergebnisse ableiten.

Lisp-Funktionen und damit auch andere von Lisp zugängliche Werkzeuge, wie etwa das relationale Datenbanksystem RTMS, lassen sich gut in Wissensbasen integrieren.

Einige Funktionen von ART, wie etwa Zugriffe auf das Schema-System, stehen direkt von Lisp aus zur Verfügung. Gerade diese Funktionenschnittstelle ist jedoch an vielen Stellen beschränkt, was sich bei der Implementierung einer Systemerweiterung, wie sie Gegenstand der vorliegenden Arbeit ist, sehr nachteilig bemerkbar macht. Viele Funktionen im Zusammenhang mit der Verarbeitung von Regeln, insbesondere solche, die den Zugang zum leistungsfähigen Pattern-Matcher, zum intern aufgebauten Join-Netz oder zu über Regeln gespeicherten Status- und Lokalisierungsinformationen gestatten würden, sind an dieser Funktionenschnittstelle nicht zugänglich.

4.5 Untersuchungen der Wissensbasis in XUMA

Das für die Analysenplanerstellung notwendige Wissen ist fast ausschließlich im Schema-System realisiert. Abbildung 4-1 zeigt einen Ausschnitt des semantischen

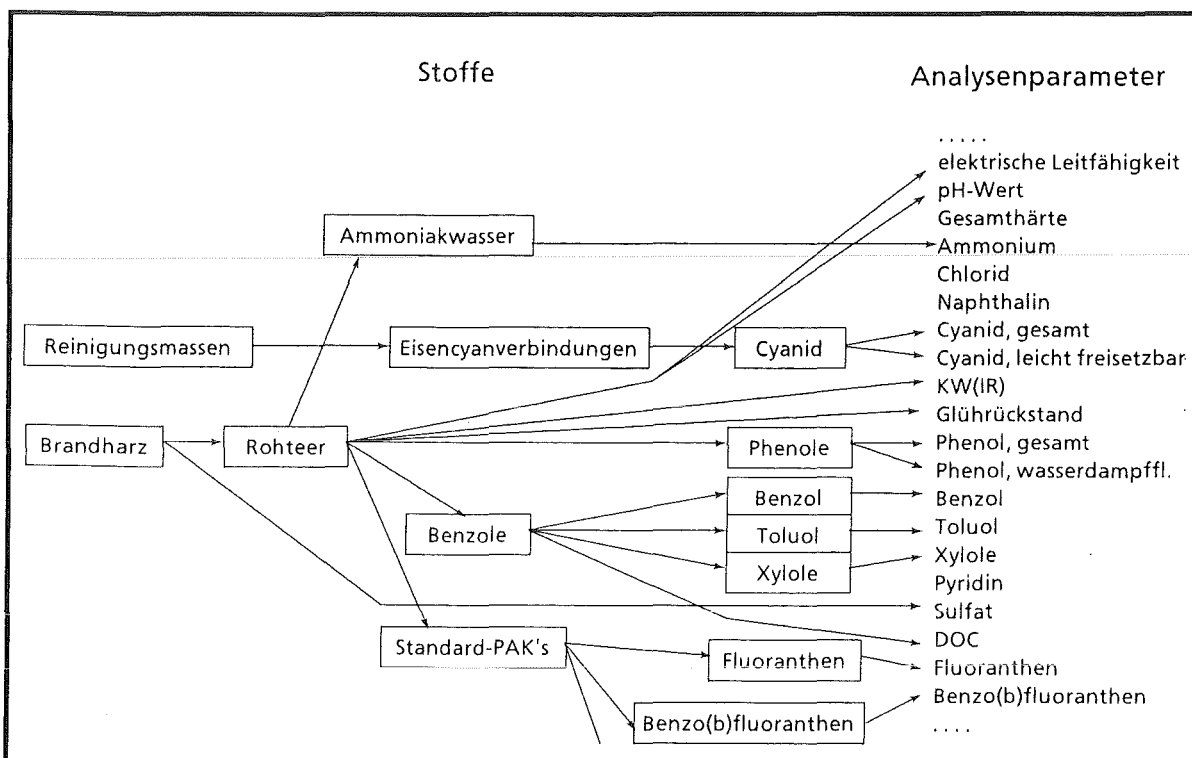


Abb. 4-1 Semantisches Netz der Stoffe und Parameter in XUMA im Ausschnitt (nach [Geiger et al. 89])

Netzes, das die Zusammenhänge zwischen Stoffen und Analysenparametern darstellt.

Zu den einzelnen Parametern existiert weiteres in Slots repräsentiertes Wissen über ihre Eigenschaften wie z.B. ihr Typ und Wertebereich. Zu Branchen, Anlagen und ihren Komponenten existiert ebenfalls ein semantisches Netz, das mit dem gezeigten durch Verweise auf die jeweils typischen Stoffe verbunden ist. Regeln im Bereich der Analysenplanerstellung dienen im wesentlichen der Dialogschnittstelle, der Ablaufsteuerung und der Auswertung der Slots im Schema-System entsprechend einiger wesentlicher Prinzipien. Für den Wissenserwerb ist im Bezug auf Stoffdaten bereits eine Schnittstelle realisiert, weitere werden ergänzt.

Das Wissen im Bereich 'Beurteilung' liegt demgegenüber im wesentlichen in Regelform vor. Es läßt sich den Bereichen:

- Regeln zur Anwendung von Grenzwerttabellen,
- Beurteilungsregeln zu Analysenparametern und
- Regeln zur Zusammenfassung von Beurteilungen auf Proben- und Fallebene zuordnen.

Wissen, ebenfalls in Regelform, das diesen Bereichen aber nicht direkt zugeordnet werden kann, gibt es zur

- Spezifikation unscharfer Vergleichsoperatoren und zur
- Plausibilitätskontrolle bei erfaßten Analysendaten.

Für den Wissenserwerb besonders interessant, sind die beiden Bereiche Anwendung von Grenzwerttabellen und Beurteilung ausgehend von Analysenparametern, denn hier besteht am ehesten Erweiterungs- und Änderungsbedarf.

In beiden Bereichen lassen sich Regelmäßigkeiten feststellen, die zwar nicht auf alle, aber auf eine große Zahl der Regeln zutreffen.

- Regeln zur Anwendung von Grenzwerttabellen (Abbildung 4-2 zeigt dazu zwei typische Beispiele) beziehen sich jeweils auf einen oder mehrere zu einer Tabelle gehörende Grenzwerte. Der zu einem Parameter in der aktuell untersuchten Analyse vorhandene Wert wird nun mit diesem Wert oder mit diesen Werten verglichen. Ein Vergleich erfolgt mit den üblichen arithmetischen oder mit den speziell im System definierten unscharfen Vergleichsoperatoren. Er bezieht sich auf die Spalten der Grenzwerttabelle für einen konkret betrachteten Parameter und den Wert dieses Parameteres in einer untersuchten Analyse. Ist der direkte Vergleich mit dem Wert der Tabelle nicht aussagekräftig, so können Produkte oder Quotienten dieses Wertes mit Zahlen gebildet werden. Außerdem kann die Spalte für den

Wenn ein Messwert mit der Tabelle W151 verglichen wird
und größer als der entsprechende A-Wert ist
und kleiner als der B-Wert
dann
ist er in Qualitätsklasse III einzustufen.

Wenn ein Messwert mit der Tabelle W151 verglichen wird
und der entsprechende A-Wert ist nicht angegeben
und der Messwert ist kleiner als der B-Wert
dann
ist er in Qualitätsklassen-Intervall I bis III einzustufen.

Abbildung 4-2 Regeln zur Grenzwerttabellenanwendung

betrachteten Parameter in der Grenzwerttabelle nicht angegeben sein. Eine Aussage dieser Art kann ebenfalls in Grenzwerttabellen-Regeln als Prämisse auftreten. Neben diesen Bedingungen und ihren Kombinationen enthält jede Regel bestimmte Standardmuster, wie z.B. zur Einbeziehung der zentral als Fakten organisierten Priorisierung zwischen mehreren anwendbaren Tabellen. Das von einer solchen Regel abgeleitete Wissen ist die Einstufung eines Parameters in eine der sechs Qualitätsklassen, die der aus dem zugehörigen Meßwert ersichtlichen Gefährlichkeit entspricht. Ist die Angabe einer Klasse nicht angemessen, so kann stattdessen ein Intervall angegeben werden. Die Einstufung nach Qualitätsklassen erfolgt in Regeln dieses Typs immer. Sie kann jedoch um weitere Anmerkungen ergänzt werden.

Die Realisierung der Beispielregeln in der Wissensbasis von XUMA in der ART-Notation entspricht dann dem in Abbildung 4-3 gezeigten Quellcode.

Die Prämissen der zweiten Gruppe 'Beurteilungsregeln zu Analysenparametern' greifen auf verschiedene Parameter der aktuell untersuchten Analyse, d.h. auf die Slots des entsprechenden Analyse-Schema, zu. Dabei kann es sich um numerische Parameter oder solche mit festem Wertebereich handeln. Auch ob es sich um Boden-, Luft- oder Wasserproben handelt bzw. ob das Untersuchungsobjekt das wäßrige Eluat einer Probe ist, hat Einfluß auf die Beurteilung. Aussagen über numerische Parameter können - ähnlich wie es für die Grenzwerte gilt - Vergleiche zu bestimmten Zahlenwerten, zwischen mehreren anderen Parametern und zu daraus gebildeten arithmetischen Ausdrücken sein. Aussagen über Parameter mit festem Wertebereich sind oft Disjunktionen der Form, daß einer von mehreren Werten zutrifft, oder einfache Attribut-Wert-Aussagen sowie deren Negation. Dies gilt sinngemäß auch für die Art des

```
(defrule glw-W151-2
  ":comment Wenn ein Messwert x mit der W151-Tabelle verglichen wird
    und A < x <= B,
    dann liegt x in Qualitaetsklasse III.
  :evidence 3"
  (DECLARE (saliency ?*qk*))
  (current-phase stoffgef (analyse $?))
  (aktuell-ist analyse ?analyse)
  (parameter ?analyse (?parameter "=" ?wert ?einheit $?))
  (grenzwert ?tabelle&W151 (A ?parameter ?grenze1&:numberp ?einheit))
  (grenzwert ?tabelle&W151 (B ?parameter ?grenze2&:numberp ?einheit))
  (messwert-typ ?analyse ?parameter ?typ)
  (gw-prioritaeten ?typ ($? (?tabelle ?prio) $?))
  (test (> ?wert ?grenze1))
  (test (<= ?wert ?grenze2))
  =>
  (ASSERT (schema ?analyse
    (wertung (parameter ?parameter QK-III ?tabelle ?prio))))))

(defrule glw-W151-3
  ":comment Wenn ein Messwert x mit der W151-Tabelle verglichen wird,
    A ist nicht angegeben und x <= B,
    dann liegt x in den Qualitaetsklassen I bis III.
  :evidence 3"
  (DECLARE (saliency ?*qk*))
  (current-phase stoffgef (analyse $?))
  (aktuell-ist analyse ?analyse)
  (parameter ?analyse (?parameter "="|"<" ?wert ?einheit $?))
  (NOT (grenzwert ?tabelle&W151 (A ?parameter ? ?)))
  (grenzwert ?tabelle&W151 (B ?parameter ?grenze&:numberp ?einheit))
  (messwert-typ ?analyse ?parameter ?typ)
  (gw-prioritaeten ?typ ($? (?tabelle ?prio) $?))
  (test (<= ?wert ?grenze))
  =>
  (ASSERT (schema ?analyse
    (wertung (parameter ?parameter (QK-I QK-III) ?tabelle ?prio))))))
```

Abbildung 4-3 Regeln zur Grenzwerttabellenanwendung in ART-Notation

Untersuchungsobjektes, das einer Analyse zugrundeliegt. Die gesamte Prämisse wird im allgemeinen durch mehrere Bedingungen dieser Form beschrieben. Ihr wird eine Interpretation zugeordnet. Dazu werden in einem Slot 'wertung' der untersuchten Analyse - in der Regel - mehrere Aussagen ergänzt. Die Form dieser Aussagen läßt sich einigen wenigen - etwa 10 - typischen Konzepten zuordnen, die jeweils verschiedene Bezugswerte und teilweise weiteren erläuternden Text enthalten.

Zwei Beispiele für Regeln dieses Typs sind in Abbildung 4-4 aufgeführt. In ART sind die Regeln dann in der von Abbildung 4-5 gezeigten Form implementiert.

4.6 Einordnung des Systems

XUMA vollzieht das Vorgehen der Experten im Bereich der Altlasten nach, so wie es von Experten der LfU im Detail erfragt wurde und wie es teilweise in den

Wenn der Glühverlust viel größer als die Konzentration an petrolextrahierbaren Stoffen ist, dann ist dies ein Hinweis auf in organischem Lösemittel schwerlösliche Bestandteile.

Wenn die Elektrische Leitfähigkeit bezogen auf 25⁰C größer als 2000 uS/cm ist und es werden Grundwasserproben analysiert, dann ist die elektrische Leitfähigkeit signifikant hoch und 'Abdampfdruckstand' und 'Salz'-Konzentrationen sind zu analysieren.

Abb 4-4 Beurteilungsregeln zu Parametern

```
(defrule g2-el-leit1
  ":comment Wenn die elektrische Leitfaehigkeit groesser ist als
    2000 uS/cm und es wird Grundwasser analysiert,
    dann ist die elektrische Leitfaehigkeit signifikant hoch,
    weshalb der Abdampfdruckstand und Salz-Konzentrationen
    untersucht werden sollten.
  :evidence 3"
  (current-phase stoffgef (analyse $?))
  (aktuell-ist analyse ?analyse)
  (schema ?analyse
    (untersuchungsobjekt grundwasser)
    (parameter (#L|P-el-Leitfaehigkeit| "=" ?wert&:
      (> ?wert 2000) "uS/cm" $?)))
  =>
  (assert (schema ?analyse
    (wertung (parameter #L|P-el-Leitfaehigkeit| hoch "" 100))
    (wertung (plan-ergaenzen parameter #L|P-Abdampfdruckstand|))
    (wertung (plan-ergaenzen stoff #L|S-Salze|))))))

(defrule g2-petrol-glueh
  ":comment Wenn der Wert des Gluehverlusts viel groesser ist als der
    Wert fuer die petrolextrahierbaren Stoffe,
    dann ist dies ein Hinweis auf in organischem Loesemittel
    schwerloesliche Bestandteile.
  :evidence 3"
  (current-phase stoffgef (analyse $?))
  (aktuell-ist analyse ?analyse)
  (schema ?analyse
    (parameter (#L|P-Gluehverlust| "=" ?wert1 ?einheit $?))
    (parameter (#L|P-Petrolextrahierbare Stoffe| "=" ?wert2 ?einheit $?)))
  (viel-groesser ?wert1 ?wert2)
  =>
  (assert (schema ?analyse
    (wertung (anmerkungen allgemein
      "Da Gluehverlust >> Petrolextrahierbare Stoffe, koennte die Probe in
      organischem Loesemittel schwerloesliche Bestandteile enthalten."))))))
```

Abb. 4-5 Beurteilungsregeln zu Parametern in der ART-Notation

Unterlagen des zuständigen Ministeriums in Baden-Württemberg skizziert ist. Das Schema-System des Werkzeuges ART, mit teilweiser Auslagerung von Daten

in ein relationales Datenbank-System, und der besonders für die überwiegend auftretenden Vorwärtsregeln geeignete Agenda-Mechanismus bieten eine gute Systemunterstützung.

Das System ist entsprechend der bekannten Klassifikation aus [Hayes-Roth et al. 83] als Interpretationssystem anzusehen. Die - teilweise durch die Definition von Phasen explizit gemachte - Problemlösungsstrategie im Bereich 'Beurteilung und Bewertung' ist das mehrstufige Ableiten und Zusammenfassen von Bewertungsaussagen, ausgehend von Daten der chemikalisch-physikalischen Analysen eines Altlastenfalles. Klassische Diagnostik-Architekturen - wie das auf Clancey zurückgehende 'Hufeisen der heuristischen Klassifikation' [Clancey 85] - wären nicht für das Problem anwendbar, da es keine gut abgrenzbaren Diagnosen gibt. Statt solchen Diagnosen gibt es bezüglich jeder der drei Stufen Analyse, Probe und Fall :

- Gefährlichkeitseinstufungen zu jedem Parameter,
- individuelle Beurteilungsaussagen, mit denen das System - so wie sonst ein Fachexperte auch - auf bestimmte von ihm erkannte Situationsmerkmale hinweist und
- eine Grobeinstufung des Falles (bzw. der Probe, der Analyse).

Daraus ergibt sich auch, daß ein Zurückgreifen auf die in der Forschung schon besser erkundeten Methoden diagnostischen Wissenserwerbs (siehe zweites Kapitel der vorliegenden Arbeit) nicht erfolgversprechend erscheint. Stattdessen ist es Aufgabe des im folgenden Kapitel dargestellten Systementwurfes, mit Blick auf die Untersuchung der Wissensbasis eine spezifische Lösung zu konzipieren.

5. Anforderungen und Systementwurf der Wissenserwerbskomponente

In diesem Kapitel wird der Systementwurf für die in dieser Arbeit realisierte Wissenserwerbskomponente des Altlasten-Systems XUMA erläutert. Zunächst werden die Anforderungen aufgeführt, die sich aus der Analyse im zweiten Kapitel und unter Berücksichtigung der konkreten Systemumgebung stellen. Danach werden die wesentlichen ausgewählten Konzepte und Verfahren erläutert und Struktur und Ablauf modellhaft dargestellt.

5.1 Anforderungen

- Einfache Benutzung der Schnittstelle

Der Interaktionsaufwand für den Benutzer soll durch Verwendung von Menüs und maussensitiven Darstellungen minimiert werden. Die Benutzung soll schnell erlernbar sein und nervierende Dialoge vermeiden.

- Angemessenheit der Darstellung aus der Perspektive des Anwendungsgebietes

Wissen soll in der Terminologie des Anwendungsgebietes in weitgehend natürlichsprachlicher Form dargestellt werden.

- Konsistenz der Wissensbasis unterstützen

Die Eingabe von Regeln soll im Kontext zu anderem - etwa im semantischen Netz repräsentierten - Wissen des Systems geschehen. Dies betrifft etwa die Auswahl aus einer Wertemenge bei Attributen, d.h. Zugriff auf die slots der jeweiligen Schemata.

- Integration der Oberfläche in das System

Die Oberfläche soll sich nach optischen und funktionalen Aspekten in das Gesamtsystem einfügen. Das bezieht sich auf die Bedeutung einzelner Maustasten, den Bildschirmaufbau und die Begriffswelt.

- Komfortable Navigation

Das Identifizieren von vorhandenem Wissen bzw. Aufdecken seines Fehlens soll durch eine übersichtliche, der Denkweise des Fachexperten angemessenen Zugangsstruktur möglich sein. Insbesondere soll sich der Benutzer keine Dateinamen oder Namen von Suchfunktionen merken müssen. Die Struktur soll vom jeweiligen Stand der Wissensbasis generisch abhängen, also insbesondere wachsen können, wenn sich die Wissensbasis verändert.

- Testmöglichkeit von eingegebenem Wissen

Bei Neueingabe bzw. Modifikation der Wissensbasis müssen die Änderungen in einfacher Weise gültig gemacht werden und in der üblichen Systemumgebung auf Testfälle anwendbar sein.

- Rücksetzfunktionen

Ein einfaches Zurücksetzen auf den Stand der letzten Version nach einer Reihe von Änderungen soll möglich sein.

- Verbergen der Kontroll- und Ablaufinformation bei der Regeleingabe

Der Benutzer soll nur die wesentlichen Aussagen eingeben müssen. Die Klauseln von Regeln zur Ablaufsteuerung etwa werden kontextabhängig hinzugefügt bzw. interpretiert.

- Aufsetzen auf der gegebenen Wissensbasis

Die Komponente soll auf die gegebene Wissensbasis in der für das Werkzeug ART geeigneten Notation aufsetzen. Sie muß auch umgekehrt erworbenes Wissen in dieser Notation darstellen können. Das mit der Komponente eingegebene Wissen soll in gleicher Weise wieder änderbar sein.

Bei an der Komponente vorbei manipulierten Regeln muß die neue Fassung analysiert und gegebenenfalls als unverständlich abgefangen werden. Teile der Wissensbasis, die in diesem Sinne unverständlich sind, sind für den Benutzer der Wissenserwerbskomponente nicht zugänglich. Dies gilt ebenfalls für Regeln, die entweder die Ablaufsteuerung oder die grundsätzliche Struktur des Systems betreffen oder zu komplex aufgebaut sind.

- Protokoll und Sicherungsfunktionen

Bei wesentlichen Änderungen und Eingaben des Benutzer werden entsprechende Hinweise in einem im Hintergrund angelegten Protokoll zu Dokumentationszwecken abgelegt.

Bei Verlassen des Systems wird dem Benutzer ein Menü angeboten, mit dem die Änderungen und Eingaben als dauerhaft gültig bestätigt bzw. zurückgenommen werden können.

5.2 Grundlegende Konzepte

Bei der Auswahl der Konzepte für die Wissenserwerbskomponente lagen die als interne Berichte dokumentierten Erfahrungen aus ersten Phasen des Wissenserwerbs - also ohne interaktive Komponente - und die aktuelle Wissensbasis des Prototyps vor. Während die Stoffdaten, die Analysenparameter und das Wissen über den Zusammenhang von Branchen, Anlagenteilen und den dafür typischen Stoffen, wie in Kapitel 4 dargelegt, mit Hilfe eines semantischen Netzes und einer angeschlossenen relationalen Datenbank realisiert sind, liegt das Wissen zur Beurteilung und Bewertung überwiegend in Regelform vor.

Bei ersten Analysen wurde deutlich, daß der Entwurf einer Schnittstelle, die alle Ausdrucksmöglichkeiten von Regeln in ART berücksichtigt, den Verzicht auf einige im vorangehenden Abschnitt genannte Anforderungen, wie etwa leichte Handhabbarkeit, Darstellung in Fachterminologie etc. bedeutet hätte. Andererseits wäre eine Beschränkung auf sehr einfache Ausdrucksmöglichkeiten wie z.B. die Konjunktion von Objekt-Attribut-Wert-Tripeln nicht möglich gewesen, ohne wiederum wesentliche Anforderungen wie etwa die Mächtigkeit der Komponente zu beeinträchtigen. Die schließlich gefundene Lösung stellt nach der Analyse der Wissensbasis und des Anwendungsgebietes einen sinnvollen Kompromiß dar. Dazu wurde neben der Notation für das Werkzeug ART eine objektorientierte Repräsentation für Regeln gewählt, mit deren Hilfe sich die Anforderungen an die Wissenserwerbskomponente besonders gut realisieren ließen.

Diese Lösung geht davon aus, daß Regeln eines Interpretationssystems oft eine unter bestimmten Gesichtspunkten regelmäßige Struktur haben. In dieser Struktur wird einer Situation, die durch eine Konjunktion von Aussagen beschrieben wird, eine oder mehrere interpretierende Bewertungen zugeordnet.

Die Aussagen selbst haben dabei jeweils eine von verschiedenen möglichen Formen. Typen von Aussageformen sind Disjunktion, Vergleich oder einfache Objekt-Attribut-Wert-Aussage und verschiedene anwendungsspezifische Spezialisierungen davon, die im Detailentwurf erläutert werden. Disjunktion bedeutet in diesem Zusammenhang die Disjunktion verschiedener einem Attribut zugeordnete Wert-Aussagen. Vergleiche sind mit arithmetischen oder mit in der Wissensbasis enthaltenen anwendungsspezifischen unscharfen Vergleichsoperatoren zur Repräsentation von vagem Wissen möglich. Die Ausdrücke, auf die sie sich beziehen, können sowohl aus absoluten Werten als auch aus Variablen bzw. ihren Kombinationen mit üblichen arithmetischen Operatoren bestehen. Dabei sind für den Benutzer die Variablennamen, die bei Bedarf mit einfachen Heuristiken erzeugt werden, durch im Kontext verständliche, aussagekräftigere Fachbegriffe verdeckt.

Jeder dieser Aussageformen müssen bestimmte Methoden zugeordnet werden, um sie in fachterminologischer und werkzeugspezifischer Notation darzustellen, um sie einzugeben bzw. zu modifizieren, sie zu erklären u. a. mehr. Da sich die Aussageformen bezüglich der Art, wie diese Methoden anzuwenden sind, in Klassen einteilen lassen, die teilweise ähnlich sind, aber jeweils nicht so viele Elemente haben, daß ein Datenbankansatz gerechtfertigt wäre, bietet sich der objektorientierte Realisierungsansatz an. Der Detailentwurf wird zeigen, daß dabei Mechanismen der Vererbung und Methodenkombination gut nutzbar sind.

5.3 Klassen von Objekten

Regeln und die sie bildenden Aussageformen können jeweils als Objekte aufgefaßt werden. Sie entsprechen nichttrivialen Datenstrukturen, auf die Methoden zur Darstellung, Eingabe bzw. Modifikation und Abbildung in die ART-Notation anzuwenden sind. Dabei bietet der objektorientierte Ansatz die Möglichkeit, konkrete Aussagen in Regeln als Instanzen zu repräsentieren mit den oben erwähnten Vorteilen bei der Realisierung der zugehörigen Methoden entsprechend gerade der für ihren Typ spezifischen Struktur. Dies macht die Entwicklung insofern besonders flexibel, als für neue Typen von Aussageformen nicht in viele bestehende Funktionen eingegriffen werden muß, sondern die gegebenenfalls zusätzlich erforderlichen spezifischen Methoden einfach zusätzlich definiert werden können. Regeln werden in Menüs in natürlicher, fachterminologisch angemessener Form dargestellt. Dazu werden Methoden verwandt, welche für die in den Regeln auftretenden Aussagen deren natürlich-

sprachliche Form erzeugen und dann mit den konstanten Anteilen vervollständigen. Die Aussagen selbst werden als Instanzen, d.h. konkrete Ausprägungen jeweils einer der genannten Typen von Aussageformen dargestellt. Zu den Instanzen existieren Modifikations-, Ergänzungs- und Erklärungsmethoden. Instanzen repräsentieren Aussagen durch eine konkrete Belegung ihrer Variablen. Die Texte dieser Aussagen werden maussensitiv dargestellt. Mit Mausclick kann der Benutzer die Aussage über die entsprechende Methode eingeben oder modifizieren.

Diese Methoden werden zusammengesetzt aus eventuell mehreren der folgenden Kernfunktionen :

- einfache Auswahl
- Multiple Choice Auswahl
- Texteingabe
- bestimmte häufig auftretende und deshalb vordefinierte Auswahlfunktionen
- Eingabe eines Ausdrucks durch Kombination von Zahlen, Variablen und arithmetischen Operatoren.
- Eingabe von Vergleichen mit arithmetischen oder unscharfen Vergleichsoperatoren

Neben den Methoden der Aussageformen zur Modifikation ihrer Werte, gibt es auch Methoden der Regel zum Löschen, Ersetzen und Ergänzen von Aussageformen. Eine Reihe von Methoden wie etwa zur automatischen Erzeugung von Variablennamen und zur Plausibilitätsprüfung vervollständigen den Entwurf.

Die Regeln selbst werden heterarchisch zu Gruppen zusammengefaßt, die ihre Identifikation entsprechend der Denkweise des Benutzers unterstützen. Die Zugehörigkeit zu Gruppen ist dabei nicht statisch. Sie kann sich bei Änderung der Regel ebenfalls ändern.

5.4 Struktur und Ablaufmodell

Aus Benutzersicht können die Funktionen der Wissenserwerbskomponente den Tätigkeiten Auswahl, Eingabe, Test und Bestätigung zugeordnet werden.

5.4.1 Auswahl

Der Auswahl liegt eine Zugangsheterarchie von Gruppen zugrunde. Jede Gruppe enthält Untergruppen oder Regeln. Für den Benutzer erscheinen Gruppen als aussagefähige Begriffe zur Strukturierung der Wissensbasis (Beispiele für solche

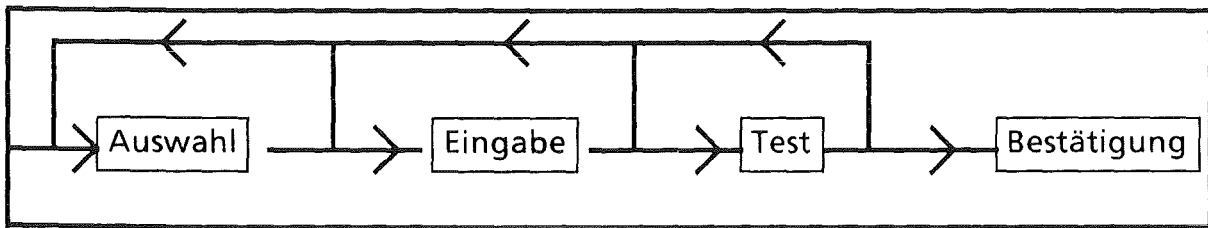


Abb. 5-1 grobes Tätigkeitsmodell für die Wissenserwerbskomponente

den entsprechenden Menüs Oberbegriffe aus, so wird er nach typischerweise zwei bis vier Zwischenschritten eine vorhandene Regel identifizieren oder sich durch Auswahl einer entsprechenden Menü-Zeile das Gerüst für eine neue Regel erzeugen lassen, das dann noch zu vervollständigen ist.

Diese Lösung erlaubt einerseits eine hinreichende Strukturierung der Wissensbasis, die den Fachexperten bei der Verwendung der Komponente unterstützt, und erfordert andererseits relativ wenig Interaktionsaufwand.

Die Abbildung 5-2 zeigt die Verfeinerung des Entwurfes zur Auswahl. Diese erfolgt durch mehrfaches Auswählen von Untergruppen.

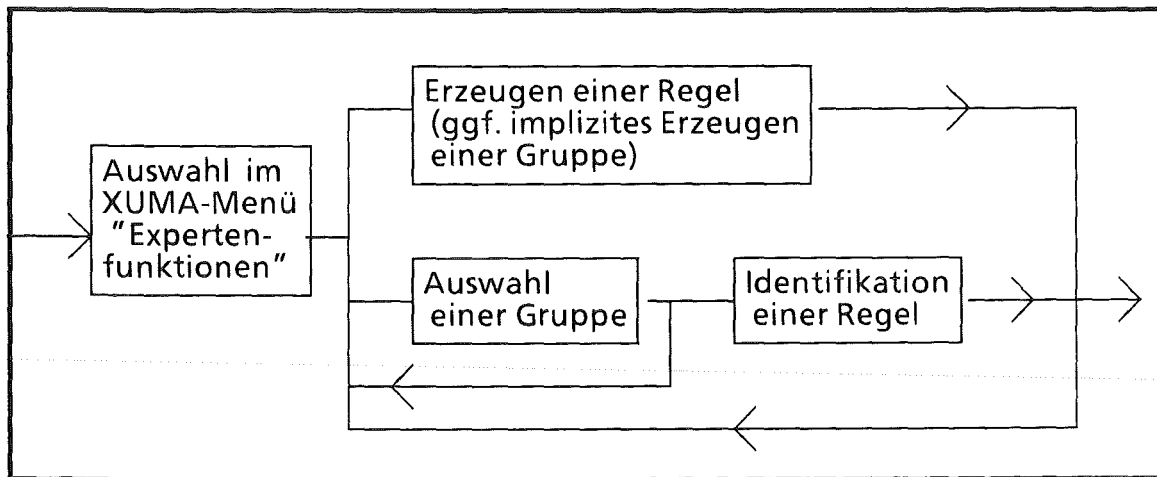


Abb. 5-2 Ablaufmodell für die Auswahl

Die Struktur ist über die Angabe der Eigenschaften von Untergruppen und nicht durch deren Aufzählung definiert und in diesem Sinne generisch. Sie ist eine Heterarchie, da eine Regel nach den Kriterien der entsprechenden Gruppen zu verschiedenen dazugehören kann. Der Weg über verschiedene Auswahl-Pfade führt in diesem Fall zur gleichen Regel. Eine neu erzeugte Regel wird eingefügt, indem in jede geeignete Gruppe ein Verweis eingetragen wird. Eine Gruppe ist 'geeignet' in diesem Sinne, wenn ihre funktionale Instanzvariable zur Prüfung von Regeln ein positives Ergebnis liefert. Ist die neue Regel in einer Gruppe, die

von Regeln ein positives Ergebnis liefert. Ist die neue Regel in einer Gruppe, die selbst keine Regeln erzeugt, geeignet, aber nicht aber in einer ihrer aktuellen Untergruppen, so wird durch die Erzeugung dieser Regel die Struktur erweitert und implizit eine neue Gruppe erzeugt. Die Instanz einer Regel - neu oder aus den vorhandenen identifiziert - steht für die Methoden, die im folgenden Abschnitt zusammengefaßt und erläutert sind, zur Verfügung. Der Benutzer sieht an dieser Stelle die Regel in Menü-Darstellung.

5.4.2 Eingabe

Eingabe ist funktional zu verstehen. Die Methoden in diesem Bereich bestehen überwiegend auch aus inhaltlich zusammenhängenden aufeinanderfolgenden Auswahlmenüs und nur in wenigen Fällen aus Tastatureingabe.

Zunächst wird die Regel - oder das Gerüst einer neuen Regel - als maussensitives Menü dargestellt. Voraussetzung für den Aufbau dieses Menüs ist die Erzeugung der natürlichsprachlichen Darstellung. Dazu existieren bei den Instanzen zur Repräsentation von Aussageformen innerhalb der Regel jeweils Methoden. Ihre Ergebnisse (Texte) werden von Methoden der Regel zusammengesetzt und ergänzt. Durch Anklicken können die eigentlichen Eingabe-Methoden - die auf wenigen schon erwähnten Mechanismen kontextabhängig aufbauen - ausgelöst werden.

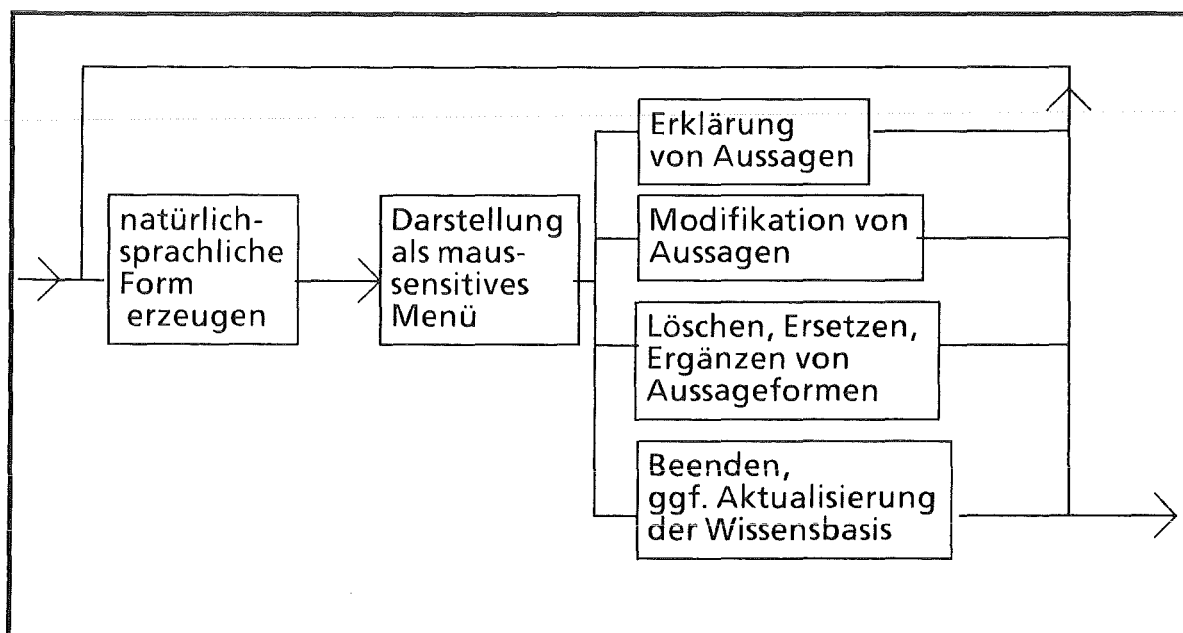


Abb. 5-3 Ablaufmodell für den Bereich Eingabe

Beim Verlassen der Eingabe kann der Benutzer diese gültig machen, d.h. sie wird zur Laufzeit mit dem Werkzeug ART übersetzt und gehört dann zur dynamischen Wissensbasis. Dazu wird sie vorher von entsprechenden Methoden auf Vollständigkeit getestet. Anschließend kann sie der Benutzer in der Anwendung im Zusammenhang testen. Der Benutzer hat daneben die Möglichkeit das Regelmanü ohne die Übersetzung zu verlassen und kann im zuletzt gewählten Menü im Bereich Auswahl fortfahren. Weitere Optionen sind das Zurücksetzen auf den Stand der letzten Version, sowie das Löschen der Regel.

5.4.3 Test

Der Test erfolgt, indem ein Altlastenfall mit den schon bestehenden Funktionen in XUMA bearbeitet wird und die Ergebnisse überprüft werden. Systemgeführte Tests oder Truth Maintenance Funktionen wurden im Rahmen dieser Arbeit nicht realisiert.

5.4.3 Bestätigung

Das System ist so konzipiert, daß der Benutzer zwar nur im Rahmen definierter Konstrukte Wissen formalisieren kann, aber Freiheit bezüglich seines Vorgehens hat. Um aber bei Verlassen des Systems nach einer möglichen Reihe von Änderungen, Tests und Erweiterungen einen definierten Zustand der Wissensbasis herzustellen, werden die im Verlauf der Sitzung vom Benutzer durchgeführten Eingaben aufbereitet und ihm in einem Menü präsentiert. Hier können neue oder geänderte Regeln für die Wissensbasis für gültig bzw. ungültig erklärt werden und auf Wunsch einzelne Eingaben noch einmal genau inspiziert werden.

5.5 Weitere Konzepte

- Unterstützung der Konsistenz der Wissensbasis

Syntaktische Konsistenz wird durch Restriktion der Eingabemöglichkeiten auf als Objektklassen definierte und deshalb interpretierbare Aussageformen erreicht. Semantische Konsistenz wird nur insofern unterstützt, als die Eingabe bezüglich des semantischen Netzes konsistent ist, d.h. nur Aussagen über Werte zur Auswahl anbietet, die in Slots der entsprechenden Schemata definiert sind.

- Übersichtlichkeit, einfache Handhabbarkeit

Die im Kontext irrelevanten Kontrollinformationen einer Regel werden für den Benutzer ausgeblendet. Für Regeln, die verschiedenen Phasen oder Prioritäten zugeordnet werden, sind jeweils andere Objektklassen definiert, so daß die Kontrollinformationen implizit vorhanden sind. Die Handhabbarkeit wird unterstützt, indem der Benutzer keine syntaktischen, werkzeugspezifischen Feinheiten wissen muß. Diese stecken in den Methoden zur Abbildung zwischen der ART-Notation und den entsprechenden Objekten. Von der Klasse des Objektes ist es auch abhängig, welche Alternativen dem Benutzer im jeweiligen Kontext zum Ändern bzw. zur Eingabe angeboten werden.

- Aufsetzen auf der gegebenen Wissensbasis

Um jederzeit auf der gegebenen Wissensbasis aufsetzen zu können, wird bei Bedarf die objektorientierte Darstellung aus den vorliegenden Dateien generiert. Auch für die Zugangsstruktur stehen nur wesentliche Konstruktionsmethoden fest, während die konkreten Regeln aus den in der Wissensbasis von XUMA relevanten Dateien gelesen werden. Von Aussagen der Regeln hängt dabei ab, welche Untergruppen in der Zugangsstruktur erzeugt werden. Beim Parsen der Regeln wird von vorneherein berücksichtigt, daß auch Konstrukte in Regeln auftreten können, zu denen es bei den Objektklassen keine Entsprechung gibt. Diese können dann nur vom Systementwickler geändert werden und sind besonders markiert.

5.6 Zusammenfassung

Der Entwurf sieht eine anwendungsspezifische Lösung vor, die sich auf die Analysen der vorangehenden Kapitel stützt. Dabei wird Wissen in Regelform erworben und dargestellt. Der Interaktionsaufwand wird durch weitgehende Menü-Funktionalität minimiert. Der Benutzer bekommt nur die für ihn relevanten Informationen dargestellt. Die Realisierung der Komponente erfolgt mit Mittel der objektorientierten Programmierung.

5.7 Umfang

Wesentlich für den Wissenserwerb bei XUMA sind neben dem in Schemata abgelegten Wissen über Stoffe, Parameter und Branchen, für die parallel Eingabeschnittstellen in einfacher maskenorientierter Form realisiert sind, bzw. realisiert werden, die Beurteilungsregeln zu Parametern und Regeln zur Anwendung von Grenzwerttabellen.

In der Komponente werden bezüglich der beiden letztgenannten Bereiche Klassen und Methoden für Regeln, Bedingungs- und Wertungsmuster definiert, sowie ein Parser für ART-Regeln und die erforderlichen Menü- und Hilfsfunktionen implementiert.

Damit umfaßt die Implementierung der Komponente :

Lisp-Funktionen :	137
Flavor :	33
Methoden :	217

Daraus ergeben sich insgesamt etwa 8850 LoC (Lines of Code) für den Quellcode der Komponente.

5.8 Erweiterungsmöglichkeiten

Die Einbeziehung bestimmter Regeln zur ausgabetechnischen Aufbereitung der Ergebnisse und zur Dialogsteuerung ist im Rahmen einer Wissenserwerbskomponente nicht sinnvoll. Als Erweiterung denkbar sind hingegen die Definition unscharfer Vergleichsoperatoren, das Wissen über die zusammenfassende Interpretation auf Proben- oder Fallebene oder die 'Berücksichtigung der örtlichen Verhältnisse'.

6. Detailentwurf und Realisierung der Wissenserwerbskomponente

Dieses Kapitel erläutert die Details des Entwurfes und seiner Realisierung in Common Lisp bzw. mit dem Common Lisp Flavor-System von Texas Instruments. Dazu wird zunächst das Flavor-System genauer vorgestellt. Im zweiten Abschnitt werden dann die wichtigsten in der Implementierung definierten Flavor und Methoden erläutert. Weitere Abschnitte behandeln die Konventionen bei Schnittstellen und Datenstrukturen, die Abbildung von bzw. in die Notation des Werkzeuges ART, verwendete Menüfunktionen und bestimmte Besonderheiten der Implementierung. Das Kapitel schließt mit einigen Anwendungsbeispielen.

6.1 Flavor-System

Zur Darstellung von Wissen - Regeln , Regelgruppen, Aussageformen - durch Objekte mit konkreten Attributwerten und zugehörigen Methoden werden Flavor instanziiert, d.h. ein konkretes Objekt wird angelegt. Ein Flavor beschreibt eine Klasse konkreter Objekte - sogenannter Instanzen - durch Angabe von Attributen, die im folgenden als Instanzvariablen bezeichnet werden, und durch Definition von Methoden. Das Flavor-System ist eine Erweiterung zu Common Lisp. Methoden sind durch ein Defmethod- Makroaufruf zu definieren und können den Sprachschatz von Common Lisp nutzen sowie insbesondere eine Argumentenliste nach denselben Konventionen wie Funktionen haben.

Im Gegensatz zu Funktionen, die jeweils einen in einem Namensraum eindeutigen Namen besitzen, kann es jedoch mehrere Methoden gleichen Namens geben. Daraus ergibt sich gerade der Vorzug des objektorientierten Programmierparadigmas, nach dem die Interpretation einer Methode von der entsprechenden Klasse eines Objektes abhängt. Über die Zugehörigkeit eines Objektes zu einer Klasse - gleichbedeutend mit Flavor - wird dann die konkret auszuführende Methode identifiziert. Instanzvariablen können beliebige Lisp-Objekte als Werte zugeordnet werden. Sie können insbesondere Default-Werte haben, die dann für alle Instanzen des Flavors gelten, sofern sie nicht explizit überschrieben werden.

Neben der Angabe der Instanzvariablen und Methoden eines Flavors sind Vererbung und Methodenkombination die wesentlichen Mechanismen zu seiner Definition. Dazu kann die Definition sogenannte Komponentenflavor enthalten.

Der neu definierte Flavor enthält dann neben den in seiner Definition angegebenen Instanzvariablen und Methoden auch alle, die von seinen Komponentenflavor vererbt werden. Die gültige Methode eines Flavors ist, falls keine für ihn explizit definierte existiert, die Methode des in der Liste der Komponentenflavor erstgenannten, für den die Methode definiert ist.

Als Alternative zu diesem einfachen Mechanismus des Überschreibens existieren einige Möglichkeiten vererbte und neudefinierte Methoden zu kombinieren. Stellen die Methoden Prädikate dar, so bieten sich logische Kombinationsmechanismen wie 'und' bzw. 'oder' an. Eine andere Form der Methodenkombination - sie spielt in dieser Arbeit eine Rolle - ist die Spezialisierung von Methoden durch `:before` und `:after`, d.h. vor bzw. nach der ererbten Methode gleichen Namens wird die neudefinierte Teilmethode ausgeführt. Daraus ergibt sich eine gute Möglichkeit die Aufteilung von Abläufen abzubilden. Die Teile, die für mehrere Klassen von Objekten gleich sind, werden von der Methode eines gemeinsamen Komponentenflavors vererbt. Spezielle Teile werden zusätzlich spezifiziert.

6.2 Methoden der Wissenserwerbskomponente

Die Methoden der Wissenserwerbskomponente lassen sich den drei Bereichen Menü-Darstellung, Aktualisierung der Wissensbasis sowie Eingabe und Modifikation von Aussagen zuordnen. Die wichtigsten Methoden werden in den folgenden Abschnitten im Detail erläutert.

6.2.1 Menü-Darstellung von Regeln

Das Kernstück der Komponente ist dabei die Darstellung und Modifikation einer Regel bzw. eines zu vervollständigenden Regelgerüsts. Die Darstellung erfolgt in einem Menü selektierbarer und nicht selektierbarer Zeilen. Nicht selektierbar sind dabei die konstanten Zeilen, mit denen die einzelnen Aussageformen zu einer lesbaren Darstellung vervollständigt werden. Die Zusammenstellung dieser verschiedenen Zeilen geschieht in einer der Regel zugeordneten Methode (`:edit`). Für alle auftretenden Aussageformen im Prämissen- oder Folgerungsteil einer Regel liefern jeweils deren Methoden (`:menu-item`) die Listen zur Menü-Ausgabe. Die Abbildung 6-1 zeigt die Benutzersicht für ein schon in Kapitel 4 erwähntes Beispiel einer Regel.

Dazu wird ein Text von einer Methode (`:make-edit-string`) erzeugt. Bei Vergleichen, die arithmetische Ausdrücke enthalten, ist diese Methode rekursiv

```
Name der Regel : 02-EL-LEIT1  
  
Wenn  
Untersuchungsobjekt ist GRUNDWASSER  
und  
elektrische Leitfaehigkeit bez. auf 25 Grad-Celsius > 2000  
..(Einheit: uS/cm) .  
  
dann gilt :  
Der Parameter elektrische Leitfaehigkeit bez. auf 25  
..Grad-Celsius wird mit HOCH bewertet. (Grund: )  
und  
Analysenplan um Parameter (bzw. Stoff) Abdampfdruckstand  
..ergaenzen.  
und  
Analysenplan um Parameter (bzw. Stoff) Salze ergaenzen.  
  
< E N D E      oder      O P T I O N E N >
```

Abb. 6-1 Benutzersicht einer Regel

definiert, bei einfacheren Aussagen hängt sie aussagefähige Strings und sogenannte 'externe Namen' aneinander, die in XUMA die weniger gut lesbaren, dafür aber eindeutigen internen Namen ergänzen.

Dabei wird bei längeren Aussagen ein Zeilenumbruch vorgenommen. Die :menu-item-Methode liefert eine Liste, von einem oder mehreren Elementen, die u.a. eine erzeugte Zeichenkette enthalten. Dabei enthält das erste Element dieser Liste im allgemeinen Fall neben dem Text auch die Methoden, die den verschiedenen Maustasten zugeordnet sind und von den Aussageformen entsprechend ihres Typs interpretiert werden. Die Zuordnung kann direkt vom Menü-System interpretiert werden.

Zu diesem allgemeinen Fall gibt es zwei Ausnahmen. Im ersten Fall sind Aussagen für eine bestimmte Klasse von Regeln unveränderbar. Das tritt in der Implementierung im Zusammenhang mit der Anwendung von Grenzwerttabellen auf. In solchen Fällen wird die möglicherweise ererbte Methode für :menu-item überschrieben mit einer, die nur den Text und das Schlüsselwort :no-select mit Option T enthält. Im zweiten Fall wird von mehreren für die eindeutige und korrekte Abbildung auf Muster der ART-Notation nötigen Instanzen nur eine für den Benutzer dargestellt. Prämissen, bei denen das auftreten kann, enthalten eine dies markierende logische Instanzvariable, die dann auf T statt auf den default-Wert NIL gesetzt ist. Dieser Fall tritt im Zusammenhang mit Variablen und Vergleichsausdrücken auf. Die für den Benutzer maßgebliche Aussage wird dann von einer den Vergleichsausdruck repräsentierenden Instanz geliefert. Für darin vorkommende Variablen

bekommt sie von der Regelinstanz, die dazu eine Liste angelegt hat, einen aussagefähigen Begriff. Auch Prämissen, welche diesen Variablennamen und das zugehörige Konzept, etwa einen bestimmten im Schema der aktuell untersuchten Analyse bekannten Parameter, miteinander identifiziert, sind in der Regel durch Instanzen vertreten. Sie sind als Prämissen für die logische Vollständigkeit und die korrekte Abbildung für den Schlußfolgerungsmechanismus auch wesentlich. Für den Benutzer, der ja Variablen in diesen Fällen gar nicht als solche zur Kenntnis nimmt, sind sie jedoch unverständlich und werden deshalb über diese Abfrage ausgeblendet.

Der Ablauf bei einzelnen durch Anklicken der Zeilen im Regel-Menü ausgelösten Methoden wird mit Beispielen in 6.2.3 ausgeführt.

6.2.2 Aktualisierung der Wissensbasis

Beim Anklicken der Ende-Zeile ergeben sich für den Benutzer die Alternativen

- Einfügen der Regel in die dynamische Wissensbasis

Bei dieser Option 'gültig machen' wird durch eine Methode `:vollständig-p` gesichert, daß die Regel selbst und auch die in ihr enthaltenen Instanzen vollständig sind. Vollständigkeit bedeutet dabei, daß die Instanzvariablen Werte enthalten, mit denen sinnvolle Muster oder Bewertungsaussagen erzeugt werden können. Ist dies nicht gegeben, so wird die Ursache erläutert und die Eingabe muß vervollständigt werden. Im erfolgreichen, normalen Fall kann die Regelinstanz mit der `:in-art` Methode eine Liste erzeugen, die der Notation von ART (siehe 6.2.4) entspricht. Die Methode `:defrule` stößt dabei diese Methode an und schreibt die Liste dann in eine temporäre Datei, die für ART geladen werden kann

- Rückkehr in das Regelmanü (ohne Veränderung)
- Rückkehr in das letzte Auswahlmenü

Hier wird die Auswahl in dem Menü fortgesetzt, von dem aus die Regel identifiziert bzw. erzeugt wurde. Für die dynamische Wissensbasis und die Zugangsstruktur bleiben etwaige Änderungen bedeutungslos, die Definition der Regel entspricht ihrem alten Stand. Der Benutzer kann zu einem späteren Zeitpunkt die Regel in die Wissensbasis einfügen. Beim Beenden des Gesamtsystems XUMA wird diese 'offene' Änderung in einem Menü kenntlich gemacht und kann dann zurückgenommen oder für gültig erklärt werden.

- Zurücksetzen auf den Stand der letzten Version

Die Aussage der Regel zum Zeitpunkt des Initialisierens wird gesichert. Bei der Auswahl dieser Option erhält der Benutzer wieder die dem ursprünglichem Zustand entsprechende Belegung von Instanzvariablen.

Methoden wie :defrule und :edit-until-ok sind für alle Regeln gleich. Sie werden dem abstrakten Flavor für Regeln 'regel-flvr' zugeordnet. Andere Methoden sind für alle Regeln für die Beurteilung der Stoffgefährlichkeit gleich, so daß hierfür ein abstrakter Komponenten-Flavor definiert wird. Die schließlich instanziierten Flavor, welche die Instanzvariablen und Methoden der beiden zuvor genannten erben, folgen aus der Analyse des vierten Kapitels und beziehen sich auf die Bereiche

- Anwendung von Grenzwerttabellen
- Beurteilungsregeln zu Analysenparametern

Die Namen orientieren sich an den im System verwendeten Präfixen für diesen Bereich. Die Abbildung 6-2 zeigt die Flavor für Regeln im schematischen Überblick.

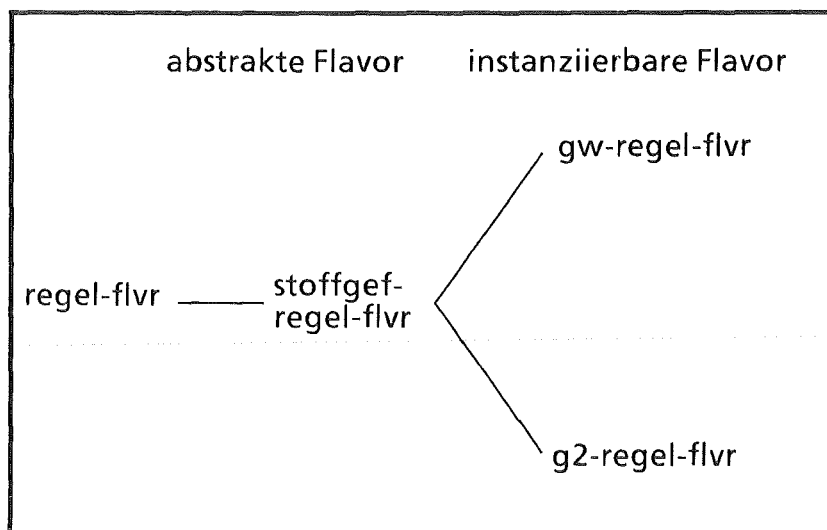


Abb. 6-2 Regel Flavor

Besteht bezüglich einer Aussage in der Regel noch Erklärungsbedarf, so wird durch Klicken mit der rechten Maustaste die Methode :dokumentiere an die entsprechende Instanz geschickt. Dadurch wird die Ausgabe eines Erklärungstextes veranlaßt, der z.B. die Interpretation von Begriffen durch das System genauer erläutert.

6.2.3 Eingabe und Modifikation von Aussagen

Die Eingabe noch unspezifizierter oder die Modifikation vorhandener Werte einer Instanz erfolgt bei Klicken mit der linken Maustaste. Dadurch wird die :choose-Methode der entsprechenden Instanz ausgelöst. Die dahinter stehenden Aktionen sind entsprechend dem Typ der Aussage verschieden und nach den schon im Systementwurf genannten Prinzipien zusammengesetzt.

Dazu einige Beispiele :

- Disjunktive Aussagen zu Parametern mit festem Wertebereich werden über ein Multiple-Choice -Menü dieser festen Werte modifiziert. Abbildung 6-3 zeigt das Beispiel einer Regel mit einer Prämisse dieser Art. Abbildung 6-4 das zugehörige Menü zur Eingabe bzw. Modifikation der Werte.

```
Name der Regel : 02-KONSISTENZ2  
  
Wenn  
Untersuchungsobjekt ist nicht ELURT  
und  
Konsistenz ist oelig oder zaehfluessig  
dann gilt :  
Die allgemeine Ausbreitungstendenz ist HOCH (Grund: Konsistenz)  
..(definitiv)  
  
< E N D E      oder      O P T I O N E N >
```

Abb. 6-3 Beispiel-Regel mit disjunktiver Aussage
(hier zur Konsistenz einer Probe)

- Für Ausdrücke in Vergleichen mit numerischen Parametern gibt es rekursive Eingabefunktionen, die Menüs der möglichen Operanden bzw. Operatoren anbietet. Dabei wird der bisher eingegebene Teilausdruck angezeigt. Die Erweiterung der Operandenmenge ist dynamisch durch entsprechende Auswahl in diesem Menü möglich.

Die Methoden für Bewertungsmuster beinhalten oft Zugriffe auf Auswahl-funktionen für Stoffe oder qualitative Werte und die Eingabe von Text. Zweiwertige Eingaben werden durch die Integration eines 'toggle' so realisiert, daß im Eingabe-Menü zwischen diesen Werten gewechselt werden kann.

Auswahl zu Konsistenz

gasfoermig
fluessig
zaehfluessig
fest
oelig
schlammfoermig
staubfoermig
stichfest
pastoes
granuliert

Ausfuehren Zurueck

Abb. 6-4 Multiple-Choice Menü der Methode zur Eingabe bzw. Modifikation disjunktiver Aussagen

Sollen ganze Aussageformen ergänzt, gelöscht oder ersetzt werden, so kann der Benutzer ein Löschen-Ändern-Menü anfordern. Dieses ist ebenfalls Instanz eines Flavors (delete-or-exchange-flvr), da die Abarbeitung dieser Funktionen jeweils für Bewertungsmuster und Aussageformen des Prämissenteils verschieden erfolgt und sich auch zwischen den Regeltypen unterscheidet.

Die aufgerufene Instanz läßt zunächst die Auswahl zwischen den drei Alternativen Löschen, Ersetzen und Ergänzen zu oder auch die Rückkehr in die letzte Auswahlebene der Zugangsstruktur (vergleiche 6.6). Dann wird - bei Ergänzung oder Ersetzung einer Aussageform - über ein weiteres Menü ein Stichwort erfragt, zu dem die im Kontext passende Instanz einer neuen Aussageform erzeugt und in der Regel ergänzt wird.

Für den Benutzer ist die Schnittstelle bis auf die vorkommenden Oberbegriffe und die ggf. auf seine Auswahl hin neu erzeugten Aussageformen identisch. Die Abbildung 6-5 zeigt die Benutzersicht bei der Auswahl von Aussageformen zur Bewertung. Während bei Bewertungsmustern eine eins-zu-eins Abbildung der Benutzerauswahl auf den zu instanziiierenden Flavor erfolgt, wird bei Auswahl einer Prämisse zur Parameterbeurteilung die Benutzereingabe durch Zugriff auf das Wissen in Slots ausgewertet. Die Erzeugung von Aussageformen (bzw. Instanzen dafür) hängt dann von den Eigenschaften dieser Parameter ab (d.h. für solche mit festem Wertebereich stehen dann a priori andere Methoden zur Verfügung als für numerische Parameter).

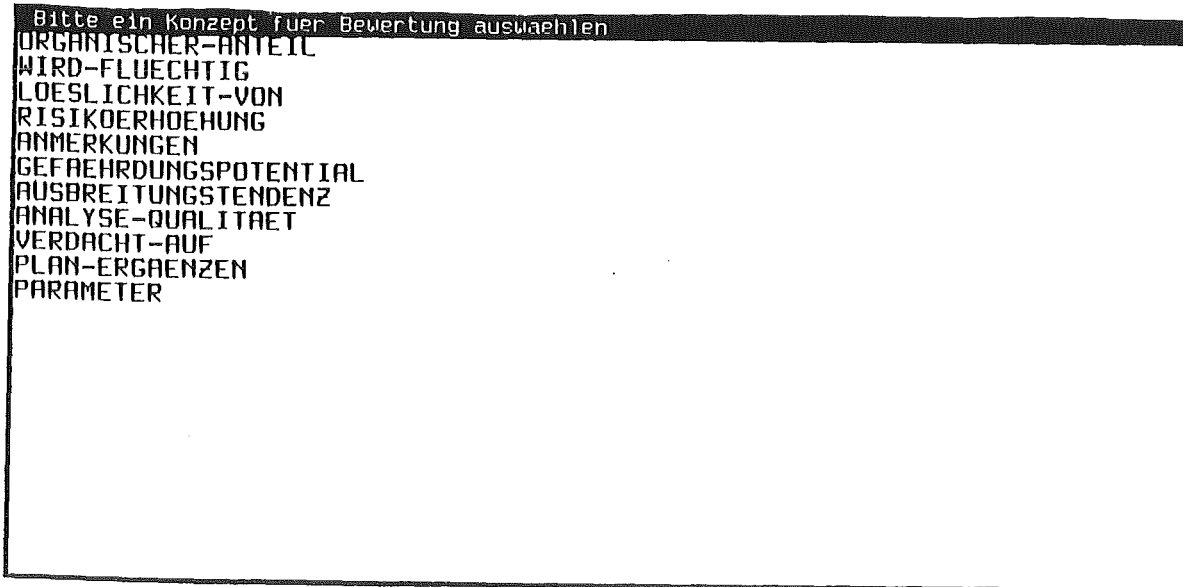


Abb. 6-5 Auswahl eines Musters für Aussageformen
(hier Wertungsaussagen)

6.3 Konventionen bei Namen und Strukturen

Alle auf '-p' endenden Methoden sind Prädikate, d.h. sie liefern T oder NIL.

Die Methoden :in-art und :von-art nehmen die Abbildung zwischen Instanzvariablen und entsprechenden Listen in ART-Notation vor. Dabei soll eine Methode entweder vollständig oder gar nicht wirksam werden. Für die Methode :von-art bedeutet dies, daß nur eine vollständig korrekte und interpretierbare Liste zum Überschreiben der Instanzvariablen führt. Sind nur einige Werte in einer insgesamt inkorrekten Liste interpretierbar, so werden diese nicht übernommen; Resultatwert der Methode ist dann NIL, im Erfolgsfall T. Bei :in-art gilt analog, daß bei auch nur teilweise unspezifizierten Werten der Resultatwert der Methode NIL ist, bei gültigen Werten die Liste, welche die Aussage in ART repräsentiert.

Funktionen, die innerhalb des Parsers eine Regel in die einzelnen Teile zerlegen, zu denen dann ggf. die Instanzen erzeugt werden, sind entsprechend der Konvention implementiert, daß sie nicht nur das Ergebnis des Analysenschrittes, sondern immer auch die restliche noch nicht analysierte Liste sowie einzelne aus der Liste entfernte Konstrukte als Werte liefert. Dabei gibt es reihenfolge-sensitive Funktionen, die jeweils nur den Anfang dieser Liste untersuchen, und Funktionen, welche die gesamte restliche Liste für ein Muster durchsuchen. Im ersten Fall haben die Funktionsnamen dann ein 'get-' Präfix bei lesenden oder 'skip-' Präfix bei überlesenden Funktionen, im zweiten ein 'scan-' Präfix.

Variablen in Regeln werden als dreielementige Liste aus dem internen Variablennamen, einem Text - ein für den Benutzer verständlicher Oberbegriff - und dem Typ der Variable repräsentiert. Diese Konvention gilt ebenfalls für arithmetische und unscharfe Vergleichsoperatoren, so daß die Methoden zur Bildung lesbarer Aussagen hier immer auf das zweite Listenelement zugreifen, was die Anwendung von Umwandlungsfunktionen einspart. Hier wäre die ebenfalls mögliche Repräsentation durch Flavor-Instanzen aufgrund der geringen Anforderungen übertrieben gewesen.

6.4 Abbildung zwischen ART-Notation und Instanzen

Die Abbildung zwischen ART-Notation und Instanzen erfolgt über einen Zwischenschritt, da Sonderzeichen in ART verwendet werden, die das direkte Lesen von ART-Mustern erschweren. Lisp bietet zwar einige Möglichkeiten für die Funktion 'read' die Bedeutung von Zeichen durch veränderten Eintrag in eine sogenannte 'readtable' zu ändern, allerdings stößt dies an Grenzen, wenn die Kategorie eines Zeichen verändert werden soll. In ART gibt es den als Oder-Zeichen fungierenden Längsstrich '|', der gleichzeitig auch, wenn er in einem mit #L beginnenden Ausdruck steht, in seiner normalen Lisp-Semantik als Doppel-Escape-Zeichen auftreten kann. Die zwischen zwei aufeinanderfolgenden Längsstrichen stehenden Zeichen werden dann zu einem Symbol zusammengefaßt.

Diese zwei Bedeutungen werden vom Parser unterschieden, der eine Folge von Strings oder Symbolen, die jeweils durch '|' getrennt sind, für die weitere Verarbeitung in eine Liste dieser Symbole umwandelt. Der Parser erkennt Kommentarzeichen und jeweils zusammengehörende Klammern. Für die Sonderzeichen ':', sonst Lisp-Package-Markierung, sowie '&' und '~' werden reservierte Symbole eingeführt. Für eine Folge zusammengehöriger, durch '|' getrennter Symbole, setzt er das Symbol 'oder-liste' gefolgt von dieser Liste ein. Insgesamt liefert er also Listen, die nur diese Sonder-Symbole, die Symbole und Strings, die in Regeln auftreten, Zahlen sowie Listen, die rekursiv wieder diese Eigenschaften haben, zu. Diese Zwischendarstellung kann durch entsprechende Umkehrfunktionen wieder auf die ursprüngliche ART-Notation abgebildet werden. Die genannten Bedingungen erlauben die Verarbeitung durch die Methoden, indem etwa Listenelemente in Instanzvariablen übertragen werden und umgekehrt Instanzen entsprechende, sie repräsentierende ART-Listen erzeugen.

6.5 Menü-Oberfläche

Bei der Realisierung der Menü-Oberfläche wurde darauf geachtet, die unvermeidlichen Abhängigkeiten vom Menü-System der Implementierungsumgebung über wenige zentrale Mechanismen zu steuern, um den möglichen Übergang zu anderen Window-Systemen nicht durch Entwurfsentscheidungen zu erschweren.

Menü-Items, welche die drei Maustasten und ihre Interpretation durch Methoden des systemspezifischen `w:menu`-Flavor nutzen, werden nur bei wenigen zentralen Methoden der Regeln und Regelgruppen eingesetzt. Andere Menü-Eingaben werden über zentral definierte Funktionen wie `we-einfach-auswahl` und `we-mehrfach-auswahl` sowie über besondere ihrerseits darauf zugreifende Funktionen wie `we-stoff-auswahl` u.a. realisiert.

Die Schnittstelle dazu ist neben der Text-Überschrift für das Menü eine Liste von Items. Dabei werden unter Items entweder Strings, Symbole oder zweielementige Listen verstanden, welche die folgende, auf vielen Window-Systemen realisierbare Semantik haben :

- Symbole werden als Strings dargestellt und beim Anklicken als Werte des Funktionsaufrufes zurückgeliefert,
- Strings werden dargestellt, sind aber Kommentar und nicht auswählbar
- Listen enthalten ein erstes im Menü darzustellendes Element (ein String oder Symbol) und eine zweite Liste, die im Falle der Auswahl zu evaluieren ist.

6.6 Zugangsstruktur

Für die Zugangsstruktur wurde der '`regel-gruppe-flvr`' definiert. Instanzen dieses Flavors haben

- einen Namen
- eine Liste der Instanzen der Untergruppen
- eine Liste der Instanzen der Regeln
- eine Funktion zum Prüfen, ob eine gegebene Regel nach dem Ordnungskriterium der Gruppe dazupasst.
- eine Funktion zum Erzeugen einer neuen Regel
- ggf. eine Funktion zur Auswahl eines Oberbegriffes für neue Regeln oder Untergruppen
- eine logische Variable, die anzeigt, ob die Gruppe neue Regeln bei sich selbst einordnet oder ggf. zunächst Untergruppen erzeugt.

Die Methoden dieser Gruppe erlauben die Aktualisierung der Zugehörigkeit von Regeln zu Gruppen, falls sich die Aussage von Regeln ändert. Für den Benutzer

stellen die Gruppen einen Menüzugang dar, der die Wissensbasis strukturiert und nach verschiedenen Kriterien die Auswahl vorhandener Regeln oder die Erzeugung neuer Regeln erlaubt. Die Gruppen werden dabei nur über ihre Eigenschaften definiert. Dadurch kann die Struktur wachsen und erfüllt auch bei Veränderung der Wissensbasis, sofern die Veränderung innerhalb zur Verfügung stehender Prüffunktionen geschieht, ohne späteren Eingriff des Systementwicklers ihre Aufgabe. In der aktuellen Version stehen als Ordnungskriterien zur Verfügung, ob und welche verdächtigen Substanzen in Beurteilungsregeln abgeleitet werden, auf welche Parameter sich Regeln beziehen, wo etwas über den organischen Anteil ausgesagt wird u.a. Die Abbildungen 6-6 und 6-7 zeigen an



Abb. 6-6 Regelgruppen im Bereich Beurteilungsregeln aus Benutzersicht

zwei Beispielen wie der Benutzer Regelgruppen sieht. Das Einfügen neuer Ordnungskriterien ist dabei einfach möglich, indem eine entsprechende auf Instanzen anzuwendende Funktionen definiert wird und als Wert der Instanzvariable einer Gruppe eingefügt wird.

Die Verweise auf Regeln sind mittelbar, d.h. sie erfolgen zunächst auf eine Instanz des 'identify-flvr' und erst durch dessen Instanzvariable 'rule-ref' auf einen instanziierten Regel-flvr. Die Instanz des 'identify-flvr' weist dabei nur die wesentlichen Statusinformationen sowie Name, Datei und weitere Hinweise aus. Hintergrund dieser Abstufung ist, daß somit auch Regeln der Wissenserwerbskomponente, zu denen bisher keine editierbaren Instanzen erzeugt wurden, zu Regelgruppen gehören können. Dies kann aus Effizienzgründen geschehen oder weil für die Konstrukte dieser Regeln keine Entsprechung bei den Flavorn vorhanden sind.

Auswahl in der Gruppe : Parameter	
Parameter	Petroleextrahierbare Stoffe
Parameter	Gluehverlust bei 550 Grad-Celsius
Parameter	Konsistenz
Parameter	Bor
Parameter	Geloeester organischer Kohlenstoff (DOC)
Parameter	elektrische Leitfaehigkeit bez. auf 25 Grad-Celsius
Parameter	Gluehrueckstand bei 550 Grad-Celsius
Parameter	Abdampfueckstand
Parameter	Geruch
Parameter	Homogenitaet (Boden)
Parameter	Farbe, qualitativ
Neue Regel	

Abb. 6-7 Regelgruppen im Bereich Beurteilungsregeln aus Benutzersicht hier : Untergruppen nach dem Kriterium 'beurteilter Parameter'

6.7 Besonderheiten

- Protokollfunktionen

Von den von einem Benutzer durchgefuehrten Aktionen wird ein Kurzprotokoll mit Name und Zeitstempel angefertigt. Dies kann archiviert werden und bei Bedarf dazu dienen durchgefuehrte Aenderung zu rekonstruieren. Entsprechende Eintraege erfolgen durch die Methode, die auch die Aktualisierung der Wissensbasis nach Abschluß einer Regeleingabe oder Modifikation durchfuehrt.

- Verwendung nichtlokaler, dynamischer Ausgaenge

Dieses Sprachkonstrukt von Common Lisp entspricht in etwa dem Goto-Befehl in prozeduralen Programmiersprachen: Dabei wird auch aus aufgerufenen Funktionen bei einem 'throw'-Befehl an das Ende eines mit 'catch' markierten Blockes gesprungen. Da es sich um eine Ausnahme von Lisp's funktionalen Programmierparadigma handelt und die Verwendung aus Gruenden der Software-Qualitaet sparsam zu handhaben ist, wird diese hier besonders erwaehnt.

Sie tritt in der Implementierung in genau einem Zusammenhang auf. Bei der Analyse von Regeln in ihrer ART-Notation ergeben sich eine Reihe von Abfragen, syntaktische und semantische Tests, bei deren Fehlschlag jeweils

die Analyse beendet wird. Hier werden einige 'throw'-Befehle eingesetzt um weitere Analyseschritte zu überspringen; in diesem Zusammenhang eine übersichtliche und vertretbare Lösung.

- Verwendung globaler Variablen

Globale Variablen werden im Bezug auf einzelne Objekte nicht verwendet, da hier Instanzvariablen eine saubere und sinnvolle Form der Datenhaltung darstellen. Globale Variablen werden verwendet, wenn es um an verschiedenen Stellen benötigte Wertebereiche geht, die nicht im Schema-System abgelegt sind. Dies trifft für die Menü-Namen und Flavorbezeichnungen von möglichen Wertungsmustern, für die Qualitätsklassen und für einige andere ähnliche Daten zu.

- Verwendung von Funktionen, die nicht innerhalb des Standards von Common Lisp nach [Steele 84] liegen

Dies betrifft nur die zwei Funktionen 'LIST*\$' zur Umwandlung einer ART-Sequence in eine Liste und 'SUBSTRING' um einen in einem anderen String enthaltenen Teilstring zu finden, die aus Wartungs- und Portierungsgesichtspunkten hier aufgeführt seien.

Das Flavor System ist ebenfalls nicht Teil des genannten Common Lisp Standards. Viele Lisp-Umgebungen kennen diese Form der Klassen und Methoden-Implementierung, ein offizieller Standard existiert aber bislang noch nicht.

6.8 Anwendungsbeispiele

Anhand weiterer Beispiele wird in diesem Abschnitt die Funktionsweise der Wissenserwerbskomponente erläutert. Abbildung 6-8 zeigt den Einstieg in die Komponente über das XUMA-Menü 'Expertenfunktionen'. Das Menü umfaßt die Namen der obersten Regelgruppen. Diese entsprechen gerade den im vierten Kapitel genannten Typen von Regeln. Weiterhin wurde eine Systementwickler-Schnittstelle mit zusätzlichen Funktionen implementiert, die nach Abschluß des Testbetriebs ausgeblendet wird.

Nach Auswahl einer Tabelle in der entsprechenden Gruppe (Abbildung 6-9) erreicht der Benutzer das Menü der auf diese Tabelle bezogenen Regeln.

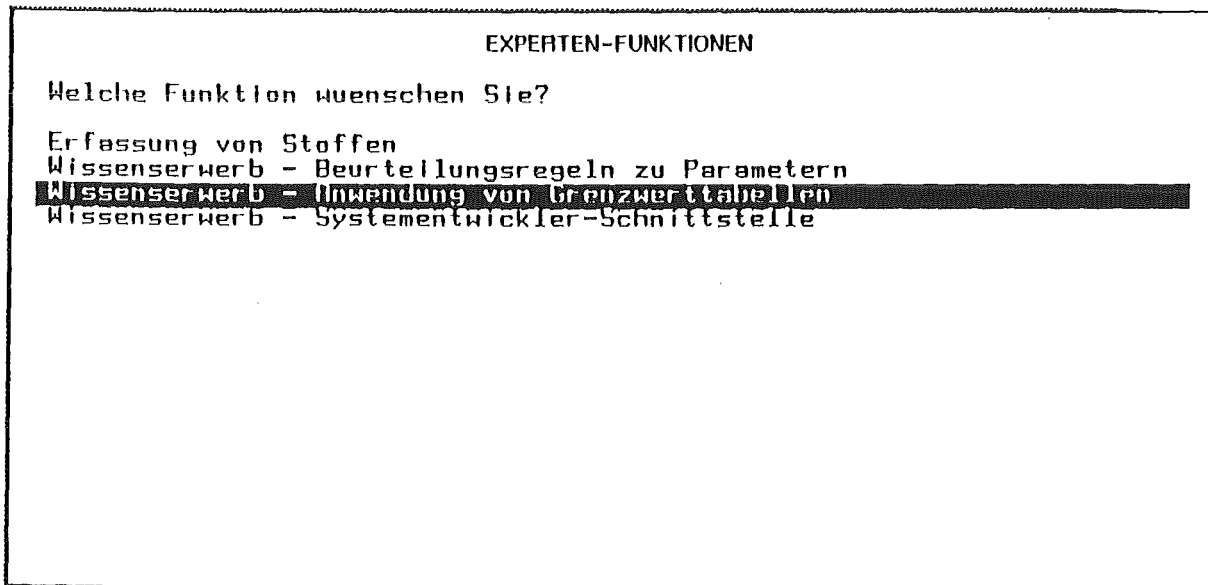


Abb. 6-8 Einstieg in die Wissenserwerbskomponente

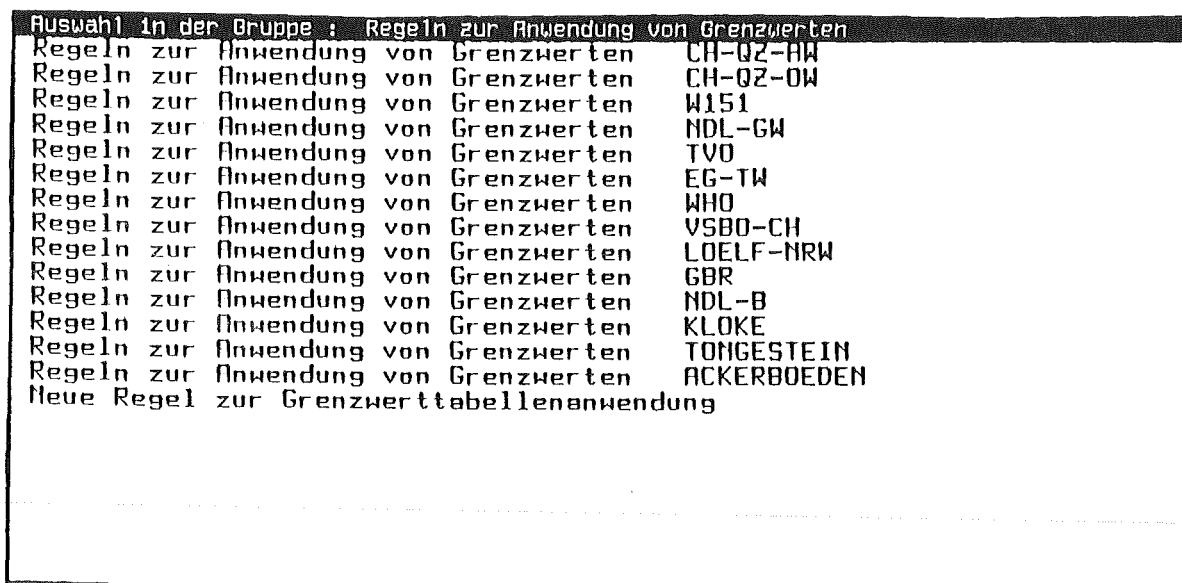


Abb. 6-9 Auswahl einer Regelgruppe nach dem Kriterium Tabelle

Die Kurzformen für Tabellen sind Referenzen auf im Altlastenbereich relevante Tabellen. Auf Klicken mit der rechten Maustaste erfolgt die Ausgabe genauerer Hinweise. Die Auswahl einer Regel in einer entsprechenden Gruppe führt direkt auf deren Menü-Darstellung (Abbildung 6-10). Die Generierung der natürlichsprachlichen Form erfolgt bei den Ausdrücken rekursiv über die Teilausdrücke. Zu Variablen und Operanden existieren jeweils geeignete Strings. Bei anderen Aussagen - etwa bei 'eine Spalte ist nicht angegeben' - existieren individuelle Methoden zur Repräsentation.


```
Name der Regel : G1W-W151-3

Wenn
ein Messwert X mit der Tabelle W151 verglichen wird
und
die Spalte A ist in der Tabelle nicht angegeben.
und
Messwert X    <=  B

dann gilt :
der Messwert X wird eingestuft in das
..Qualitaetsklassen-Intervall (QK-I bis QK-III)

< E N D E      oder      O P T I O N E N >
```

Abb. 6-10 Darstellung einer Regel zur Anwendung von Grenzwerttabellen

Jeweils die erste zu einer Aussage gehörende Zeile ist maussensitiv. Als Beispiel für die Modifikations- bzw. Eingabe-Methoden sei hier diejenige einer Bedingung in einer Regel obigen Typs angegeben ('gw-bedingung-flvr'). Aus

```
Eingabe einer Bedingung zur Anwendung der Tabelle W151
Messwert X = .. <Ausdruck>
Messwert X ≠ .. <Ausdruck>
Messwert X > .. <Ausdruck>
Messwert X < .. <Ausdruck>
Messwert X >= .. <Ausdruck>
Messwert X <= .. <Ausdruck>
Messwert X ist kleiner oder ungefaehr gleich .. <Ausdruck>
Messwert X ist groesser oder ungefaehr gleich .. <Ausdruck>
Messwert X ist viel groesser als .. <Ausdruck>
Messwert X ist liegt in der selben Groessenordnung wie .. <Ausdruck>
Messwert X ist ungefaehr gleich .. <Ausdruck>
Messwert X ist nicht kleiner oder ungefaehr gleich .. <Ausdruck>
Messwert X ist nicht groesser oder ungefaehr gleich .. <Ausdruck>
Messwert X ist nicht viel groesser als .. <Ausdruck>
Messwert X ist nicht liegt in der selben Groessenordnung wie .. <Ausdr
Messwert X ist nicht ungefaehr gleich .. <Ausdruck>
A ist nicht angegeben
B ist nicht angegeben
```

Abb. 6-11 Eingabe einer Bedingung

dem Kontext ergibt sich, daß eine solche Bedingung sich auf einen beliebigen Meßwert und Spalten der Grenzwerttabelle beziehen kann. Dabei können durch aufeinanderfolgende einfache Menü-Auswahl beliebig verschachtelte arithmetische Ausdrücke gebildet werden. Dabei wird durch entsprechende Anpassung der Menü-Optionen sichergestellt, daß jeweils die Anzahl öffnender

und schließender Klammern übereinstimmen. Zur Unterstützung der schnellen Eingabe werden eine Reihe von typischen Zahlenwerten vorgegeben, aber auch andere Zahlen können eingegeben werden. Für Vergleiche stehen übliche arithmetische sowie systemspezifische, unscharfe Operatoren zur Verfügung. Eine Spalte kann, abhängig vom Typ des Meßwertes, auch nicht angegeben sein. Diese Alternativen zur Formulierung von Bedingungen sieht der Benutzer in der in Abbildung 6-11 gezeigten Weise. Die Abbildung 6-12 zeigt ein Menü, wie es

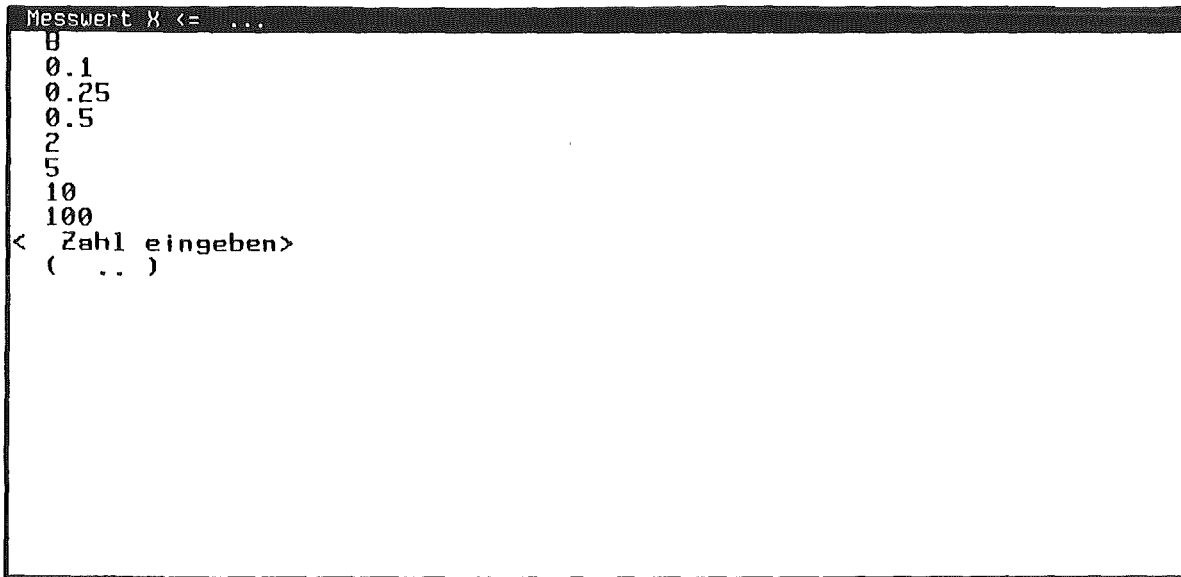


Abb. 6-12 Eingabe eines Ausdruckes

bei der Eingabe von Ausdrücken auftritt. Die Operanden ergeben sich aus dem Kontext. Hier sind es die Spalten der zugehörigen Grenzwerttabelle. Eine von ihnen wurde jedoch als nicht angegeben bezeichnet, so daß nur 'B' als Operand auftritt. Nach der Auswahl eines Operanden folgt ein Menü, das dem Benutzer verschiedene arithmetische Operatoren anbietet bzw. ihn den Ausdruck in der aktuellen Form abschließen läßt.

Hat der Benutzer seine Eingaben oder Änderungen bezüglich einer Regel durchgeführt, so werden ihm die in Abbildung 6-13 gezeigten Optionen angeboten. Der erzeugte ART-Code, durch den Regeln gültig gemacht werden können, entspricht - von Feinheiten der Formatierung abgesehen - dem in den Abbildungen 4-3 und 4-5 vorgestellten. Auf seine Darstellung kann daher hier verzichtet werden.

```
Ende oder Optionen - Menu
< zurueck > ohne Veraenderung
< Beenden > Regel in dieser Form gueltig machen
< Beenden > ohne die Regel in XUMR gueltig zu machen
< Beenden > Regel loeschen
< zurueck > dabei Regel auf den urspuenglichen Zustand zuruecksetzen
```

Abb. 6-13 Optionen bei Beenden der Eingabe bzw. Modifikation einer Regel

7. Zusammenfassung

Bei der Entwicklung und Wartung eines wissensbasierten Systems stellt der Wissenserwerb einen wesentlichen Teil des Gesamtaufwandes dar. Trotz dieser Bedeutung bietet die Forschung bislang nur wenige vielversprechende Ansätze zur Unterstützung des direkten Wissenserwerbs. Die marktgängigen hybriden Entwicklungsumgebungen bieten sehr mächtige und gut kombinierbare Mechanismen zur Darstellung von Wissen an. Zur Eingabe und Pflege von Wissen verlangen sie aber die Beherrschung der Dateisystem- und Editorfunktionalität sowie die Kenntnis von Syntax und Semantik der Repräsentationsformalismen, was sie für den Einsatz unmittelbar durch Fachexperten disqualifiziert.

Die genauere Analyse des Wissenserwerbs zeigt, daß für seine frühen, stark mit klassischen Aufgaben der Systemanalyse verwandten Phasen wie der Problemidentifikation und Konzeptionalisierung eine Automation des üblicherweise erfolgenden Wissenserwerbs mit Wissensingenieur nur schwer vorstellbar ist. In den Phasen, in denen Wissen bezüglich einer schon entworfenen Konzeptstruktur formalisiert wird, ist der Einsatz von unmittelbar durch den Fachexperten nutzbaren Wissenserwerbskomponenten jedoch oft möglich und muß im Interesse der Wirtschaftlichkeit auch erfolgen. Die erfolgversprechendsten Ansätze, welche die Forschung hier hervorgebracht hat, nutzen entweder die Ausdrucksmittel eines klar abgegrenzten Anwendungsgebietes oder sie bauen auf der probabilistisch gewichteten Zuordnung von Diagnosen und Symptomen auf, welche in neueren Arbeiten um Hilfsmittel wie Ausnahmen, Frageklassen u.a. ergänzt wurden. Allgemeine, Problemlösungstyp-übergreifende Wissenserwerbskomponenten scheinen zur Zeit nicht realisierbar, was der Tendenz der gesamten KI-Forschung entspricht, die sich von Ansätzen für universelle Problemlöser zu solchen Verfahren hin orientiert, die Lösungen für Probleme eines spezifischen Typs bieten.

Als Entwurfsgrundlage für die für den Bereich 'Altlasten' zu realisierende Komponente wurden das Anwendungsgebiet und die Wissensbasis des vorhandenen Prototyps des wissensbasierten Systems XUMA untersucht. Ein Teilergebnis dieser Untersuchung ist die Bestätigung, daß das Wissen und die Problemlösungsstrategie nicht geeignet auf Konzepte der im Bereich Wissenserwerb fortgeschrittensten wissensbasierten Diagnostik abbildbar sind. Ein weiteres Ergebnis ist das Aufdecken bestimmter Regelmäßigkeiten in der Struktur des Wissens. Daraufhin wurde eine Wissenserwerbskomponente entworfen, die

Wissenserwerb für bestimmte Klassen von Regeln zuläßt. Die Realisierung erfolgte mit Mitteln der objektorientierten Programmierung. Die Komponente kennt Objekte der Ebenen Regeln, Aussagemuster innerhalb von Regeln sowie Regelgruppen. Sie umfaßt Methoden, mit denen jeweils eine Klasse von Objekten dargestellt, modifiziert und in die für den Regel-Übersetzer des Werkzeuges ART erforderliche Notation abgebildet werden kann. Sie wurde in Common Lisp unter Verwendung der objektorientierten Spracherweiterung Flavor implementiert.

Die Wissenserwerbskomponente ermöglicht eine einfache, menüorientierte Bedienung durch Beschränkung auf einige wesentliche Interaktionsmechanismen, aus deren Kombination sich eine für den Anwendungsbereich ausreichende Mächtigkeit ergibt. Tastatureingabe wird dabei nur selten eingesetzt. Die Auswahlmenüs für die möglichen Werte der einzelnen Attribute verändern sich automatisch bei entsprechenden Änderungen im zugehörigen semantischen Netz. Das System setzt auf einer gegebenen Wissensbasis auf; hierzu wurde ein Parser nach der Methode des rekursiven Abstiegs implementiert, der die ART-Notation von Regeln analysiert. Der Einstieg des Benutzers in die Wissenserwerbskomponente erfolgt durch eine Zugangsheterarchie. Diese stellt ein Hilfsmittel zur Strukturierung der Wissensbasis und damit zur Identifikation von Wissen bzw. dem Erkennen seines Fehlens dar.

Der Entwurf selbst wird für andere Anwendungsgebiete i.a. nicht anwendbar sein, wohl aber die Entwurfsmethodologie. Sie ermöglicht aufgrund der objektorientierten Realisierung die Erweiterung der Ausdrucksmächtigkeit durch Ergänzung weiterer Methoden.

8. Literaturverzeichnis

- [BaWü 87] Ministerium für Ernährung, Landwirtschaft, Umwelt und Forsten, Baden-Württemberg :
Altlasten-Handbuch, Teil I und II
Stuttgart , 1987
- [Boose 84] Boose, J.H. : *"Personal construct theory and the transfer of human expertise"*
Proceedings of the National Conference on Artificial Intelligence, 1984, Austin, S. 27-33
- [Boose 85] Boose, J.H. : *"A knowledge acquisition program for expert systems based on the personal construct theory"*
Internat. Journal of Man-Machine Studies 23(1985), S.495-525
- [Boose 88a] Boose, J.H. : *"Recent progress in AQUINAS : a knowledge acquisition workbench"*
Proceedings of the European Knowledge Acquisition Workshop, 1988, Bonn, GMD-Studie 143, S. 2/1-15
- [Boose 88b] Boose, J.H. : *"A research framework for knowledge acquisition techniques and tools"*
Proceedings of the European Knowledge Acquisition Workshop ,1988, Bonn, GMD-Studie 143, S. 10/1-13
- [Bottaci 88] Bottaci, L. : *"Automatic knowledge base construction for expert systems "*
Internat. Journal of System Research and Information Science 2(1986), S. 57-68
- [Chandrasekaran 87] Chandrasekaran, B. : *"Towards a functional architecture for intelligence based on generic information processing tasks "*
Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI-87, Mailand, S. 1183-1192

- [Clancey 85] Clancey, W. : *"Heuristic Classification"*
AI-Journal, 20(1985), S.215-251
- [Crasemann et al. 88] Crasemann, C. , Krasemann, H. : *"Der Wissensingenieur, ein alter Hut auf neuem Kopf"*
Informatik Spektrum 11(1988), S. 43-48
- [Cullen et al. 88] Cullen, J. , Bryman , A. : *"The knowledge acquisition bottleneck: time for reassessment ?"*
Expert Systems 5(1988)3, S. 216-225
- [Davies et al. 88] Davies, M. , Hakiel , S. : *"Knowledge harvesting : a practical guide to interviewing"*
Expert Systems 5(1988)1, S. 42-50
- [Davis 79] Davis, R. : *"Interactive transfer of expertise: Acquisition of new inference rules"*
Artificial Intelligence 12(1979), S. 121-157
- [Davis 82] Davis, R. : *"Application of Meta-Level Knowledge"*
in : Davis, R. , Lenat D. B. :
"Knowledge-based systems in artificial intelligence"
McGraw-Hill, New York, 1982
- [Döppert 88] Döppert, M. : *"Altlastensanierung - Marktübersicht"*
in : *"Handbuch Umwelttechnik 88/89"*,
wlb-Verlag, 1988, S. 60-78
- [Drastal et al. 89] Drastal, G. , Czako, G. , Raatz, S.
"Induction in an abstraction space"
eingereicht bei: International Joint Conference on Artificial Intelligence, IJCAI-89, Detroit

- [Ecke-Schüth et al. 86] Ecke-Schüth, J. , Mettendorf, G.: *"Systemanalytische Aspekte bei der Entwicklung von Expertensystemen"*
Interner Bericht Nr.23/86
Universität Karlsruhe, Fakultät für Informatik, 1986
- [Eshelman et al. 87] Eshelman, L. , Ehret, D. D. , McDermott, J., Tan M.:
"MOLE : a tenacious knowledge-acquisition tool "
Internat. Journal of Man-Machine Studies 26(1987), S. 41-54
- [Feld 87] Feld, R. , Dittrich, W.: *"Vorgehensweise bei der Bewertung von Altlasten"*
wlb 10(1987), S. 57-61
- [Franzius 87] Franzius, V. : *"Altlasten - Ursachen, Dimension und Problemlösung"*
Umweltmagazin 12(1987), Anhang "Kommunale Entsorgung", S. 12-18
- [Freeman 89] Freeman, H. M. : *"Standard handbook of hazardous waste treatment and disposal"*
McGraw-Hill, New York, 1989
- [Gaines 88] Gaines, B. : *"Second generation knowledge acquisition systems"*
Proceedings of the European Knowledge Acquisition Workshop 1988, Bonn, GMD-Studie 143, S. 17/1-14
- [Gappa 88a] Gappa, U. : *"Classika , A knowledge acquisition system facilitating the formalization of advanced aspects in heuristic classification"*
Proceedings of the European Knowledge Acquisition Workshop 1988, Bonn, GMD-Studie 143, S.19/1-16

- [Gappa 88b] Gappa, U. : *"Über den 2. European Knowledge Acquisition Workshop 1988"*
KI 4(1988), S. 21-22
- [Geiger et al. 89] Geiger, W. ,Weidemann, R. ,Eitel ,W. : *"Konzepte des Expertensystem XUMA für Altlasten"*
eingereicht bei; KfK-Nachrichten, 1989
- [Gross 88] Gross, D. : *"Induction and ID/3 : more powerful than we think"*
Expert Systems 5(1988)4, S. 348-350
- [Hall et al. 85] R.P. Hall, D.F. Kibler : *"Differing methodological perspectives in AI Research"*
AI Magazine 3(1985), S. 166-178
- [Hayes-Roth et al. 83] Hayes-Roth, F. ,Waterman, D.A. , Lenat, D.B. : *"A guide to building expert systems"*
Addison-Wesley, Reading, 1983
- [Jackson 88] Jackson, A.H. : *"Machine Learning"*
Expert Systems 5(1988)2, S. 132-150
- [Kahn et al. 84] Kahn, G. , Nowlan, S. , McDermott, J. :
"A foundation for knowledge acquisition".
Proceedings of the IEEE Workshop on Principles of Knowledge-Based Systems, Denver, 1984, S. 89-114
- [Karczmarzyk et al. 88] Karczmarzyk, W. , Pfeifer, R. :
"Ein Gaswerksgelände wird saniert"
VDI-Umwelt 5(1988), S. 494-496
- [Kidd 87] Kidd, A. L. : *"Knowledge acquisition for expert systems : a practical handbook "*
Plenum Press, New York, 1987

- [Kinner et al. 86] Kinner, U. H. , Kötter, L. , Niclauß , M. :
*"Branchentypische Inventarisierung von
Bodenkontaminationen - ein erster Schritt zur
Gefährdungsabschätzung für ehemalige Betriebsgelände"*
Umweltbundesamt, Berlin, 1986
- [Kitto et al. 87] Kitto, C. M., Boose, J. H. : *"Heuristics for expertise transfer :
an implementation of a dialog manager for knowledge
acquisition"*
Internat. Journal of Man-Machine Studies 26(1987), S.183-202
- [Kornell 87] Kornell, J. : *"Formal thought and narrative thought in
knowledge acquisition"*
Internat. Journal of Man-Machine Studies 26(1987), S.203-212
- [Michalski et al. 84] Michalski, R.S. , Carbonell, J.G. , Mitchell, T.M. (Hrsg.):
"Machine Learning - an Artificial Intelligence approach"
Springer Verlag, Berlin, 1984
- [Musen et al. 86] Musen, M. A. , Combs, D. M. , Walton, J. D. , Shortliffe, E. H.
Fagan, L. M. : *"OPAL : Toward the computer-aided design of
oncology advice systems"*
Proceedings of the Tenth Symposium on Computer
Applications in Medical Care, 1986, Washington, S. 43-52
- [Musen et al. 87] Musen, M. A. , Fagan, L. M., Combs, D. M. , Shortliffe, E. H. :
*"OPAL: Use of a domain model to drive an interactive
knowledge-editing tool"*
Internat. Journal of Man-Machine Studies 26(1987), S.105-121
- [Puppe 88] Puppe, F. : *"Einführung in Expertensysteme"*
Springer Verlag, Berlin, 1988
- [Schenkel 86] Schenkel, W. : *"Altlasten"*
VDI-Umwelt 6(1986), S. 445-451

- [Schuldt 87] Schuldt, M. *"Wie gefährlich sind Altlasten im Boden"*
VDI-Umwelt 5(1987), S. 295-298
- [Schreiber et al. 88] Schreiber, G. , Breuker, J. , Bredeweg , B., Wielenga , B. :
"Modelling in KBS development"
Proceedings of the European Knowledge Acquisition
Workshop 1988, Bonn, GMD-Studie 143, S. 7/1-15
- [Steele 84] Steele, G. L. : *"Common Lisp : The Language"*
Digital Press, Bedford, 1984
- [Stoyan 88] Stoyan, H. : *"Programmiermethoden der Künstlichen
Intelligenz, Band 1"*
Springer-Verlag, Berlin, 1988
- [Trauboth 87] Trauboth, H. : *"Was kann die Informationstechnik für den
Umweltschutz tun ?"*
at-Automatisierungstechnik 35(1987)11, S. 431-442
- [Weidemann et al. 87] Weidemann, R. , Geiger, W. , Eitel, W. : *"Entwurf eines
Expertensystems zur Beurteilung von Abfallstoffen"*
in Jaeschke, A. , Page, B. (Hrsg.) : *"Informatikanwendungen
im Umweltbereich"*, Informatik-Fachberichte 170,
Springer-Verlag, Berlin 1988, S. 116-126
- [Wielinga 86] Wielinga, B.J. , Breuker, J. A. : *"Models of expertise"*
Proceedings of the Seventh European Conference on Artificial
Intelligence, ECAI-86, Brighton, S. 306-318