

ALGORITHMEN ZUR BEANTWORTUNG VON 'WARUM'-FRAGEN IN DIALOGSYSTEMEN

Wolfgang Wahlster

Summary

This paper discusses aspects of the semantics and pragmatics of 'why'-questions within the framework of Artificial Intelligence research. It is shown why 'why'-questions are a particularly important type of question in natural language systems. The explanation capabilities of several running AI systems are surveyed. Examples are given showing the approach of the dialogue system HAM-RPM to the processing of three types of 'why'-question. Several types of rejection of 'why'-questions and methods for their automatic generation are described. Formal definitions of the notions of a logical and a communicatively adequate answer to a 'why'-question are given. New methods for the selection of a communicatively adequate answer from a set of possible explanations on the basis of meta-knowledge are presented within the framework of fuzzy and non-monotonic logic. Finally, it is proposed that the dialogue system itself should be able to ask 'why'-questions to gain a better understanding of the 'why'-questions posed by the user.

1. Einführende Übersicht

Bei der Computersimulation sprachlicher Interaktion werden in dem System HAM-RPM (Hamburger Redepartnermodell, vgl. (12)) Sprechaktsequenzen berücksichtigt, in denen der menschliche Dialogpartner, nachdem eine von ihm gestellte Frage von dem simulierten Redepartner beantwortet wurde, das System durch eine 'Warum'-Frage zu einer Begründung seiner Antwort auffordert. Besonders wenn Antworten des Systems auf vagen Inferenzen und Mutmaßungen beruhen, muss das System seine Antwort durch die Verbalisierung der vollzogenen Inferenzschritte begründen und den Wahrheitsgehalt der gegebenen Information verteidigen können. In bestimmten Situationen begründet das System sogar unaufgefordert seine Antwort.

Die Beantwortung von 'Warum'-Fragen beruht in HAM-RPM auf der Fähigkeit des Systems

- den der Antwort zugrundeliegenden Inferenzprozess in einem hier als Inferenzgedächtnis bezeichneten Teil der Wissensbasis zu speichern,
- das Inferenzgedächtnis nach Information zu durchsuchen, die für eine Begründung relevant ist,
- ausgewählte Teile des Inferenzgedächtnisses zu verbalisieren.

Aufgrund einfacher dialog- und partnertaktischer Strategien versucht das System den für eine kommunikativ adäquate Begründung erforderlichen Detaillierungsgrad und das richtige Argumentationsniveau zu finden.

Triviale Schlüsse werden bei der Generierung einer Begründung zunächst nicht berücksichtigt (vgl. (25)), und erst aufgrund weiterer Nachfragen des Dialogpartners wird die gesamte Inferenzkette ausgegeben.

Als konkreter Ausgangspunkt für die nachfolgenden Untersuchungen dient ein Protokoll (vgl. Fig. 1) eines mit HAM-RPM geführten Dialogs, der typische Reaktionen des Systems auf 'Warum'-Fragen des menschlichen Dialogpartners enthält und damit den Umfang der bei der Beantwortung von 'Warum'-Fragen simulierten Fähigkeiten andeutet.

Der Beispieldialog entstand als Teil einer längeren Interaktion mit einer auf der Rechenanlage DECsystem 10 des Fachbereichs für Informatik der Universität Hamburg implementierten Version des Systems HAM-RPM und wurde bis auf den Wechsel der Schriftart (Ausgaben des Systems in kursiver Schrift) gegenüber dem originalen Computerausdruck nicht verändert.

Die dem Dialog in Fig. 1 zugrundeliegende Diskurswelt ist eine Straßenverkehrsszene, die dem Untersuchungsmaterial von NAGEL [19] entnommen ist.

- (1) **WO PARKT DENN DAS AUTO IM AUGENBLICK?**
WELCHES AUTO MEINST DU? WARUM WILLST DU DAS DENN WISSEN?
WEIL ES 16 DAVON GIBT!
WELCHES DAVON MEINST DU? DEN KOMBI
MEINST DU DEN RECHTEN HINTEREN KOMBI?
WARUM FRAGST DU DENN NUN WIEDER?
WEIL ES ZWEI GIBT!
MEINST DU DEN RECHTEN HINTEREN KOMBI? GENAU!
HINTER DEM GELBEN, PARKENDEN PERSONENWAGEN RECHTS HINTEN.
- (2) **IST DIE PARKZONE HINTEN RECHTS EIGENTLICH GETEERT?**
ICH GLAUBE JA.
- (3) **WIE KOMMST DU DARAUF?**
EINE PARKZONE IST MEIST EIN TEIL VON EINER STRASSE,
UND EINE STRASSE IST OFT GETEERT.
- (4) **WIESO?**
EINE STRASSE IST EIN VERKEHRSWEG, UND EIN VERKEHRSWEG
IST OFT GETEERT.
- (5) **IST DIE HARTUNGSTRASSE EBEN?**
JA.
- (6) **WARUM DENN DAS?**
EINE STRASSE IST EIN VERKEHRSWEG, UND EIN VERKEHRSWEG IST EBEN.
- (7) **KÖNNTEST DU SIE DENN TRINKEN?**
NEIN, DENN EINE STRASSE IST NICHT FLÜSSIG!

Fig. 1

Während der Verarbeitung von Frage (1) in Fig. 1 löst die Mehrdeutigkeit der definiten Nominalphrase *das Auto* einen Klärungsdialog aus (vgl. (26)), in dem sich der Benutzer durch zwei metakommunikative 'Warum'-Fragen nach den Gründen für die Rückfrage des Systems erkundigt. Für die Beantwortung von Frage (2) muss das System in der Wissensbasis gespeicherte Erfahrungswerte benutzen, da es in der zugrundegelegten Situation die Beschaffenheit der Parkzone aufgrund einer Verdeckung nicht direkt feststellen kann. Da durch die lexikalische Analyse von HAM-RPM Ausdrücke wie *Wie kommst Du darauf* und *Wieso* in Frage (3) bzw. (4) auf die kanonische Form *Warum* abgebildet werden, sprechen wir im folgenden in diesem Zusammenhang stets nur von 'Warum'-Fragen. Die Antwort auf Frage (3) enthält die linguistischen Hecken *meist* und *oft*, da die Antwort auf Frage (2) nur durch vage Inferenzregeln abgeleitet werden konnte (vgl. (24)). Mit der 'Warum'-Frage (4) hinterfragt der menschliche Dialogpartner die Antwort des Systems auf Frage (3).

Solche Sequenzen von 'Warum'-Fragen können schließlich zur Ausgabe der gesamten einer Behauptung zugrundeliegenden Inferenzkette führen. Frage (6) wird durch die Verbalisierung der durch Frage (5) ausgelösten Schlussfolgerung beantwortet, die mit Hilfe der als Teil des begrifflichen Semantischen Netzwerks gespeicherten Oberbegriffshierarchie durchgeführt wurde. Frage (7) wird vom System wegen der Verletzung einer Selektionsrestriktion für das Verb *trinken* zurückgewiesen. Das Auftreten von mehr als zwei aufeinanderfolgenden Fragen führt nach einer einfachen Dialogstrategie dazu, dass das System unaufgefordert die Zurückweisung von Frage (7) begründet.

2. Warum 'Warum'-Fragen in Systemen der Künstlichen Intelligenz?

Während von Seiten der Linguistik und Fragelogik im Vergleich zu anderen Fragetypen (vgl. (9), (10)) nur wenige Untersuchungen zu 'Warum'-Fragen vorliegen (vgl. (2), (23), (11), (15), (16), (5)), erweist sich in der Forschung zur Künstlichen Intelligenz (KI) gerade dieser Fragetyp als besonders wichtig. Dies ist darauf zurückzuführen, dass die Integration von Komponenten zur Beantwortung von 'Warum'-Fragen in KI-Systeme einerseits zahlreiche Vorteile bei der praktischen Anwendung solcher Systeme mit sich bringt und andererseits für die KI – verstanden als Teil der Grundlagenwissenschaft 'cognitive science' – auch theoretisch relevant ist.

Zunächst wird die Analyse der praktischen Motivation zur Untersuchung von 'Warum'-Fragen auf den Typ der Frage und den Status der mit dem System interagierenden Person bezogen.

A. Typ der 'Warum'-Frage

Mit Hilfe von Paraphrasen (\Leftrightarrow) sollen zwei für KI-Systeme wichtige Typen von 'Warum'-Fragen unterschieden werden.

- a. Die Frage zielt auf die Erklärung des propositionalen Gehalts P einer Antwort des Systems (vgl. (1), S. 110).
Warum? \Leftrightarrow Was ist eine Erklärung für die Tatsache, dass P?
- b. Die Frage zielt auf eine Begründung eines vorangegangenen Sprechaktes (z.B. eine Frage X? des Systems).
Warum? \Leftrightarrow Was ist eine Begründung für die Tatsache, dass Du die Frage X? stellst?

B. Status der mit dem System interagierenden Person

Der mit einem natürlichsprachlichen System interagierende Personenkreis lässt sich nach v. HAHN (13) unterteilen in:

- a. Benutzer (Menschlicher Dialogpartner)
- b. Bearbeiter (Systemkonstrukteur oder -verwalter)

Für die einzelnen Interaktionsmodi, die sich aus der Kombination der Kategorien aus A und B ergeben, lassen sich die folgenden praktischen Vorteile einer Verarbeitung von 'Warum'-Fragen in KI-Systemen angeben:

- Interaktionsmodus Aa/Ba

- Das System kann den Wahrheitsgehalt der gegebenen Information verteidigen und zeigt sich damit für seine Antworten 'verantwortlich' Für den Benutzer ist die Möglichkeit der kritischen Auswertung der einer Konklusion zugrundeliegenden Prämissen und Inferenzregeln besonders wichtig, falls er
 - mit der Konklusion nicht einverstanden ist. Dies ist zu erwarten, wenn in der Datenbasis des Systems Information über wenig standardisierte Wissensgebiete (z.B. Medizin) gespeichert ist,
 - bei der Weitergabe von Systemauskünften eine besondere Verantwortung gegenüber Dritten hat (z.B. Arzt, Manager),
 - die Konklusion nicht versteht oder wenn die Auskunft für ihn unerwartet ist.
 - Falls das System zu Ausbildungszwecken eingesetzt wird, können die vom System als Reaktion auf 'Warum'-Fragen generierten Erklärungen pädagogisch wertvoll sein.
 - Insgesamt steigt das Vertrauen in die Antworten des Systems, wenn diese hinterfragbar sind.

- Interaktionsmodus Ab/Ba

- In Expertensystemen, die mit Hilfe von abgespeichertem Fachwissen und weiterer vom Benutzer auf Anfrage bereitgestellter Information ein vorgegebenes Problem lösen, kann leicht eine asymmetrische Dialogstruktur entstehen, da der Benutzer vom System über weite Teile des Dialogs wie in einem Verhör recht schematisch abgefragt wird. Wenn dem Benutzer eine Frage des Systems unangemessen zu sein scheint, geben 'Warum'-Fragen diesem mit der Dominanz des Systems unzufriedenen Benutzer die Möglichkeit, sich über die Ziele, die das System mit einer Frage verfolgt, zu informieren (vgl. (14), (21)). Dieser erste Schritt in Richtung auf ein System mit 'gemischter Initiative' führt zu einer psychologischen Aufwertung der Benutzerrolle und damit zu erhöhter Akzeptanz.
- Eine 'Warum'-Frage des Benutzers kann als Verständnisfrage dienen, deren Beantwortung zu einer treffenderen Antwort des Benutzers auf eine vorangegangene Frage des Systems führt.

- Interaktionsmodus Aa/Bb
 - Mit einer Sequenz von 'Warum'-Fragen kann der Bearbeiter zunächst auf der Ebene natürlichsprachlicher Performanz nach Ursachen für 'falsche' Antworten des Systems suchen, ohne direkt in der Wissensbasis oder dem Programmtext nachsehen zu müssen. Dies führt zu einer erleichterten Fehlerdiagnose.
 - Der schrittweise Ausbau und die Korrektur der Wissensbasis sowie damit verbundene Konsistenzprüfungen werden dadurch erleichtert, dass Auswirkungen von Veränderungen in der Wissensbasis auf das Systemverhalten durch 'Warum'-Fragen direkt kontrollierbar werden.

Die Fähigkeit zur Beantwortung von 'Warum'-Fragen in den Interaktionsmodi Aa/Ba und Aa/Bb ist besonders wichtig für KI-Systeme, die folgenden Bedingungen genügen:

Die Antwort des Systems beruht auf

- Operationen über einer großen Wissensbasis mit starker gegenseitiger Abhängigkeit einzelner Wissenseinheiten oder
- sehr langen Inferenzketten oder
- approximativen Inferenzen auf vagem oder unvollständigem Wissen.

Zusammenfassend stellen wir fest: Die durch die Beantwortung von 'Warum'-Fragen erzielte Plastizität des Verhaltens komplexer KI-Systeme führt zu größerer Akzeptanz, zu stärkerem Vertrauen des Benutzers in die Antworten des Systems und zur vereinfachten Handhabung des Systems durch den Bearbeiter, oder wie STALLMAN und SUSSMAN es formulieren "Such programs are more convincing when right, and easier to debug when wrong" ((22), S. 136).

Die Sonderstellung, die 'Warum'-Fragen gegenüber anderen Fragetypen in der KI zukommt, beruht zum großen Teil aber auch darauf, dass sie einen wichtigen Teilaspekt der zur Zeit laufenden theoretischen Untersuchungen zur Darstellung und Verarbeitung von Metawissen sind. Fernziel dieser Forschungsrichtung, an deren Anfang das automatische Planungssystem NOAH (20) stand, ist es, Programmsysteme zu entwickeln, die aufgrund von Metawissen in einem gewissen Umfang ihr eigenes Verhalten auf einer höheren Ebene als der des Programmtextes erklären können. Obwohl erste Fortschritte in der Richtung auf "ein System, das weiß, was es weiß, und versteht, was es tut" (vgl. (20), S. 7) erzielt wurden (vgl. (6), (8), (26)), ist ein allgemeiner Erklärungsmechanismus für beliebige Wissensbasen und beliebige Inferenzmechanismen noch nicht entwickelt worden.

3. Die Beantwortung von 'Warum'-Fragen in bestehenden Systemen

Obwohl die Semantik von 'Warum'-Fragen und die ihrer Beantwortung zugrundeliegenden kognitiven Prozesse erst ansatzweise untersucht sind, existieren bereits mehrere natürlichsprachliche Systeme, die bestimmte Typen

von 'Warum'-Fragen beantworten können (vgl. Fig. 2). Typisch für die Mehrzahl der in Fig. 2 aufgeführten Systeme¹ ist, dass jeweils nur eine Verwendungsweise von 'Warum'-Fragen berücksichtigt wird. So können beispielsweise von SHRDLU nur solche 'Warum'-Fragen beantwortet werden, die auf die Begründung einer simulierten nicht-verbalen Handlung abzielen, z.B. Frage: *Warum hast Du den roten Block von dem grünen Block entfernt?* Antwort: *Damit ich die grüne Pyramide darauf stellen kann.*

System	Experten-System	Diskurswelt	Vage Inferenzen	Simulation von Sprachverstehen	'Warum'-Frage zielt auf
SHRDLU [28]	0	'Blockwelt'	0	1	Simulierte nicht-verbale Handlung
MYCIN [21]	1	Pharmazie	1	0	Sprechakte: Frage
EL/ARS [22]	1	Elektronik	0	0	Inferenzergebnis
HAM-RPM [12]	0	variabel	1	1	Sprechakte: Frage/ Zurückweisung Inferenzergebnis
AMORD [17]	0	variabel	0	0	Inferenzergebnis
PROSPECTOR [14]	1	Mineralogie	1	0	Sprechakte: Frage
SAM [18]	0	variabel	1	1	Inferenzergebnis

Fig. 2

In HAM-RPM wurden erstmals Versuche unternommen, verschiedene Verwendungsweisen von 'Warum'-Fragen in einem System zu erfassen. Dagegen wird es für die Anwendung von 'Warum'-Fragen in Expertensystemen (vgl. Fig. 2), in denen charakteristischerweise oft auch von 'Why' oder 'Explain'-Kommandos gesprochen wird, weiterhin sinnvoll sein, sich auf eine Interpretation von 'Warum'-Fragen zu beschränken, da in diesen Systemen nicht die Simulation von Sprachverstehen angestrebt wird (vgl. Fig. 2).

4. Pragmatische Funktionen von 'Warum'-Fragen

Bevor die verschiedenen Arten von 'Warum'-Fragen, die in HAM-RPM verarbeitet werden, genauer analysiert werden, muss das Thema des Aufsatzes noch stärker eingeschränkt werden.

Da Dialoge mit den hier betrachteten KI-Systemen ausschließlich schriftlich über einen graphischen Kanal (an die Rechenanlage angeschlossenes Sichtgerät) geführt werden, besteht die in natürlichen Dialogen vorhandene

¹ In mehreren anderen KI-Systemen sind die Voraussetzungen für eine Integration von 'Warum'-Fragen gegeben. So könnten einige 'Warum'-Fragen z.B. in einer Version von PLIDIS, in der die Antwortsuche von einem Theorembeweiser durchgeführt wird, auf eine nachträgliche Auswertung des Deduktionsgangs zurückgeführt werden (vgl. (7), S. 46).

Möglichkeit der Beantwortung einer 'Warum'-Frage durch nicht-verbale Reaktionen in diesen KI-Systemen nicht. Sequenzen der Form

A: *Warum soll ich diese Flüssigkeit nicht berühren?*

B: Wirft Holz in die Flüssigkeit; es löst sich auf.

werden daher im folgenden nicht betrachtet.

Gemäß der Sprechakttheorie unterscheiden wir strikt zwischen Fragesätzen und Fragehandlungen, indem wir Fragesätze als besonderes verbales Schema zur Ausführung von Fragehandlungen betrachten (vgl. (30)), und lassen in den folgenden Abschnitten 'Warum'- Fragesätze, deren Äußerung keine 'Warum'-Fragehandlung bedeutet, unberücksichtigt. Die Äußerung eines solchen 'Warum'-Fragesatzes kann z.B. eine Aufforderung darstellen wie in folgender Dialogsequenz

A: *Warum kannst Du nicht ausführlicher antworten?*

B: Beantwortet die folgenden Fragen von A ausführlicher.

oder eine Einladung wie in

A: *Warum kommst Du denn heute abend nicht mal zum Essen vorbei?*

B: *Vielen Dank. Wann soll ich denn kommen?*

Auch Vorschläge, Ratschläge und Kommentare können mit Hilfe von 'Warum'-Fragesätzen kodiert werden (vgl. (11)). Voraussetzung für die Verarbeitung solcher 'Warum'-Fragesätze in einem KI-System ist die Möglichkeit zur automatischen Analyse und Generierung direkter und indirekter Sprechakte. Diese ist aber bisher in der KI erst in rudimentärer Form entwickelt worden.

Rhetorische 'Warum'-Fragen (z.B. *Warum ich Milde Sorte rauche? Weil sie sehr leicht ist und angenehm schmeckt*) und multiple Fragen, die das Fragewort 'Warum' enthalten (z.B. *Wie und warum hast Du das Programm geändert?*) werden im folgenden ebenfalls nicht betrachtet.

Bei der Simulation von Textverstehen, wie sie z.B. durch das in Fig. 2 aufgeführte System SAM angestrebt wird, ist eine wichtige mit 'Warum'-Fragen zusammenhängende Fragestellung die Erkennung von Kausalbeziehungen zwischen konjunktionlos aneinandergereihten Sätzen durch Inferenzen. Solche Systeme können nach der Eingabe eines Textes wie

Peter will ein Eis. Er schüttelt sein Sparschwein und hört nichts. Er muss sich das Geld irgendwie anders besorgen.

z.B. die folgenden Fragen beantworten

Warum schüttelt er das Sparschwein?

Warum muss er das Geld anders besorgen?

Auch auf diese Fragestellung kann ich im folgenden nicht eingehen (vgl. (3), (18)).

5. 'Warum'-Fragen in HAM-RPM

In HAM-RPM wird in Abhängigkeit von der einer 'Warum'-Frage unmittelbar vorausgegangenen Sprechaktsequenz zwischen drei Interpretationen einer 'Warum'-Frage unterschieden (vgl. Fig. 3). Da diese Interpretationsmöglichkeiten nur einen Teil der Bedeutungsvarianten von 'Warum'-Fragen erfassen, werden die entsprechenden Systemkomponenten mit Hilfe der bereits erarbeiteten Verfahren laufend erweitert.

Die häufig vorgeschlagene Unterscheidung zwischen Erklärung und Begründung (vgl. z.B. (16), oder die Unterscheidung zwischen CAUSE-explanation und GOAL-explanation in (27)) in Abhängigkeit davon, ob die 'Warum'-Frage sich auf ein Ereignis oder einen Zustand der physischen Welt bzw. auf eine menschliche Handlung bezieht, braucht derzeit in HAM-RPM nicht berücksichtigt zu werden, da handelnde Personen innerhalb der den simulierten Dialogen zugrundegelegten Diskurswelten nicht vorkommen. Die Einführung der Unterscheidung setzt u.a. die Klassifikation der Prämissen von Inferenzregeln, die in der Wissensbasis von HAM-RPM gespeichert sind, nach 'Ursache' und 'Grund' voraus. Allerdings ist es in den philosophischen Erklärungstheorien immer noch umstritten, ob die angeführte Unterscheidung überhaupt sinnvoll ist (vgl. den Überblick in (29)).

Zu Schema (a) in Fig. 3: Wie Satz (4) in Fig. 1 zeigt, kann dieser Typ von 'Warum'-Frage auch iterativ verwendet werden. Ab dem zweiten Glied in einer Kette von 'Warum'-Fragen ergibt sich folgende Mehrdeutigkeit (vgl. auch (16)): Die Frage kann sich

- (1) auf die Gültigkeit des letzten Konjunks innerhalb der als Explanans angeführten instanziierten Prämissen einer Inferenzregel oder
- (2) auf die Gültigkeit der zugrundeliegenden Inferenzregel beziehen.

- (a) BE: < Entscheidungsfrage > P?
SY: < Positive Antwort (mit abschwächender Hecke) >
BE: *Warum?*

Interpretation: Aufgrund welcher Schlussfolgerungen kommst Du darauf (zu vermuten), dass P?

- (b) BE: < Frage >
SY: < Rückfrage > X?
BE: *Warum?*

Interpretation: Warum hast Du die Frage X? gestellt?

- (c) BE: < Frage > X?
SY: < Zurückweisung >
BE: *Warum?*

Interpretation: Warum hast Du die Frage X? zurückgewiesen?

Fig. 3

In HAM-RPM ist derzeit nur die erste Möglichkeit vorgesehen. Frage (4) in Fig. 1 müsste gemäß Interpretation (2) mit: Wenn x Teil von z ist, und z die Eigenschaft y hat, dann kann man annehmen, dass auch x die Eigenschaft y hat beantwortet werden. Diese Art der Beantwortung entspricht der Verbalisierung des nur eingeschränkt gültigen Schlusses von der Eigenschaft eines Objekts auf die Eigenschaft seiner Teile, der in der Wissensbasis von HAM-RPM als DEDUCE-Prozedur mit Hilfe der KI-Programmiersprache FUZZY dargestellt ist (vgl. (25)). Interpretation (2) wird im allgemeinen als unnatürlich empfunden.

Zu Schema (b) in Fig. 3: In der bestehenden Version von HAM-RPM ist diese Interpretation einer 'Warum'-Frage nur innerhalb von Klärungsdialogen vorgesehen, die vom System initiiert werden, wenn mehrere potentielle Referenzobjekte für eine in der eingegebenen Anfrage vorkommende Nominalphrase gefunden werden. Als Grund für die Rückfrage nennt das System die Anzahl der aufgrund simulierter visueller Suchprozesse und begrifflichen oder situationsabhängigen Wissens gefundenen möglichen Referenzobjekte (vgl. Satz (1) in Fig. 1).

In der derzeitigen Implementation wird dabei nur die erste der beiden folgenden Antwortmöglichkeiten berücksichtigt²:

- es wird der Zustand ausgegeben, der zu der Fragehandlung geführt hat (z.B. *Weil es drei davon gibt!*)
- es wird der Zustand angegeben, der durch die Fragehandlung angestrebt wird (z.B. *Damit ich weiß, worauf Du Dich beziehst!*)

Zu Schema (c) in Fig. 3: Ist in einer Anfrage des Benutzers eine Selektionsbeschränkung des Verbs verletzt, so weist das System die Anfrage zurück. In diesem Fall wird eine auf die Zurückweisung bezogene 'Warum'-Frage gemäß Schema (c) beantwortet. Wie bei der Interpretation einer 'Warum'-Frage nach Schema (b) so wird auch in diesem Fall die Beantwortung der 'Warum'-Frage von einem einfachen Algorithmus geleistet. Das System speichert die Instanziierung der verletzten Selektionsbeschränkung unmittelbar bei ihrer Erkennung und verbalisiert sie bei einer sich möglicherweise anschließenden 'Warum'-Frage.

Aufwendigere Algorithmen sind derzeit in HAM-RPM nur für den durch Schema (a) definierten Typ der 'Warum'-Frage vorgesehen. In den folgenden Abschnitten werden wir daher nur noch auf diesen Typ von 'Warum'-Frage eingehen. Dabei werden wir auch diejenigen 'Warum'-Fragen der Form Warum <Entscheidungsfrage> (z.B. *Warum ist die Schlüterstraße gesperrt?*) berücksichtigen, die eine Schema (a) entsprechende Interpretation haben. Auf solche 'Warum'-Fragen können, nachdem intern die Entscheidungsfrage beantwortet wurde, ausgewählte Teile der dabei aufgebauten Inferenzkette als Antwort ausgegeben werden.

² Diese beiden Möglichkeiten bestehen für alle Sprechaktsequenzen, die Schema (b) genügen. In MYCIN und PROSPECTOR (vgl. Fig. 2) wird stets nur die zweite Antwortmöglichkeit verwendet.

6. Logisch adäquate und kommunikativ adäquate Antworten auf 'Warum'-Fragen

Für die in Abschnitt 5 mit Hilfe der Interpretation in Schema (a) von Fig. 3 eingeführte Klasse von 'Warum'-Fragen soll definiert werden, was unter einer Antwort auf eine solche Frage verstanden wird.

Sei R die Menge aller Inferenzregeln, die einem System zur Verfügung stehen. $SN-R_x$ sei die Menge aller Propositionen, durch die der Zustand der Diskurswelt zum Zeitpunkt x beschrieben ist. $SN-R_x$ entspricht in HAM-RPM das referentielle Semantische Netz, das aus Repräsentationskonstruktionen für Propositionen wie *Das Auto vorne rechts ist grün* besteht. $SN-B$ sei die Menge aller Propositionen, die unabhängig vom jeweiligen Zustand der Diskurswelt für wahr gehalten werden. $SN-B$ entspricht in HAM-RPM das begriffliche Semantische Netz, das aus Repräsentationskonstruktionen für Propositionen wie *Rot, Gelb und Grün sind die Ampelfarben* besteht. $F_x = SN-B \cup SN-R_x$ sei das Faktenwissen zum Zeitpunkt x .

In den folgenden Definitionen wird die Adäquatheit von Antworten auf 'Warum'-Fragen in Relation zur Wissensbasis $W_x = R \cup F_x$ gesetzt. Damit wird der Tatsache Rechnung getragen, dass Antworten auf die gleiche 'Warum'-Frage je nach Wissensstand der befragten Person stark differieren können.

DEFINITION: Sei $I = (I_1, \dots, I_n)$ die Menge aller Ableitungswege für eine Proposition P . Jedes $I_i \in I$ habe die Form

$$I_i = (S_{i1} \xRightarrow{r_{i1}} S_{i2} \dots \xRightarrow{r_{im}} S_{im+1} = P)$$

mit $S_{ij} \in F_x$ und $R_i = \{r_{i1}, \dots, r_{im}\} \subset R$.

Eine Antwort A auf die Frage *Warum P?* ist logisch adäquat $\Leftrightarrow A \in I$.

DEFINITION: Sei $T = \{T_1, \dots, T_m\}$ die Menge aller Teibleitungswege von I . Jedes $T_i \in T$ sei eine Menge von Inferenzschritten und habe die Form:

$$T_i = \{S_{ix} \xRightarrow{r_{ia}} S_{ix+1}, \dots, S_{iy} \xRightarrow{r_{iz}} S_{iy+1}\}$$

wobei jedes $t \in T_i$ eine Teilkette von I_i ist. Damit gilt auch $\{r_{ia}, \dots, r_{iz}\} \subset R_i$. $IT \subset T$ sei die Menge aller Teibleitungswege, die bezüglich des Gesprächspartners für informativ gehalten werden.

Eine Antwort A auf die Frage *Warum P?* ist kommunikativ adäquat $\Leftrightarrow A \in IT$.

IT muss noch operational definiert werden. In HAM-RPM wird dabei vereinfachend von folgendem Modell ausgegangen: Aus der Menge der möglichen Ableitungswege wird nur derjenige berücksichtigt, welcher

innerhalb eines Inferenzprozesses als erster gefunden wird. Mit jeder Inferenzregel in der Wissensbasis ist eine reelle Zahl aus dem Intervall $[0,1] \subset \mathbb{R}$ als Metawissen assoziiert, die als subjektiv empfundene Informativität oder 'Triftigkeit' (vgl. (25)) dieser Schlussregel interpretiert wird. IT enthält dann jeweils die Teilketten aus T, deren Informativität maximal ist. Bei der Verarbeitung iterierter 'Warum'-Fragen wird T um bereits verbalisierte Teilketten vermindert, so dass sich auch IT verändert.

In diesem Modell wird noch nicht berücksichtigt, dass die Informativität einer Antwort vom bisherigen Dialogverlauf und dem Vorwissen des Fragenden abhängt. Um gezielter kommunikativ adäquate Antworten aus der Menge logisch adäquater Antworten bilden zu können, müsste dem System das Vorwissen des menschlichen Dialogpartners in einem Teil der Wissensbasis als Partnermodell (vgl. (4)) zur Verfügung gestellt werden.

7. Zurückweisungen von 'Warum'-Fragen

Verbale Reaktionen auf die in Abschnitt 6 betrachtete Klasse von 'Warum'-Fragen, die nach den dort eingeführten Definitionen keine Antworten sind, bezeichnen wir als 'Erwiderungen'. Da in KI-Systemen meist erfolgsorientiertes Kommunizieren simuliert wird, brauchen Erwiderungen wie *Das verstehen Sie nicht?* oder *Frag mich nicht!*, durch die eine verlangte Begründung verweigert wird, hier nicht berücksichtigt zu werden. Dagegen treten Zurückweisungen von Anfragen, z.B. aufgrund von in einer frühen Verarbeitungsphase erkannten Präsuppositionsverletzungen häufig auf, da einerseits in anwendungsorientierten Systemen unnötige Verarbeitungszeiten für fehlerhafte Eingaben vermieden werden sollen, und es andererseits zur Simulation von intelligentem Sprachverhalten gehört, dass der Dialogpartner sofort auf Schwierigkeiten bei der Verarbeitung seiner Frage hingewiesen wird.

Unter Verwendung der in Abschnitt 6 eingeführten Begriffe können die in Fig. 4 dargestellten Arten von Zurückweisungen der Frage *Warum P?* unterschieden werden (vgl. (2); dort wird eine ähnliche Klassifikation anders hergeleitet).

- | | |
|----------------|--|
| (a) Bedingung: | $\neg P$ ist in der Wissensbasis gespeichert oder ableitbar |
| Beispiel: | BE: <i>Warum ist die Schlüterstraße gesperrt?</i>
SY: <i>Die Schlüterstraße ist nicht gesperrt.</i> |
| (b) Bedingung: | $P \in \text{SN-B}$ |
| Beispiel: | BE: <i>Warum sind Rot, Gelb und Grün die Ampelfarben?</i>
SY: <i>Das ist bekanntlich immer so.</i> |
| (c) Bedingung: | $P \in \text{SN-R}_x$ |
| Beispiel: | BE: <i>Warum ist das Auto vorne rechts grün?</i>
SY: <i>Das ist eben so.</i> |

Fig. 4

Fig. 5 bietet eine zusammenfassende Übersicht zu den unterschiedlichen Formen von Antworten und Er widerungen als Reaktionen auf die Frage *Warum P?*. Dabei wird auch die Möglichkeit berücksichtigt, dass Information über den Inferenzweg, der zu einer Proposition P führte, als Ableitungsverweis in der Wissensbasis mit der Proposition P assoziiert ist. Die Menge aller Ableitungsverweise bildet in den KI-Systemen AMORD (17), ARS (22) und TMS (8) ein Dependenznetz, aus dem alle Inferenzwege rekonstruiert werden können.

P gespeichert oA	P gespeichert mA	P ableitbar	\neg P gespeichert oA	\neg P gespeichert mA	\neg P ableitbar	REAKTIONEN AUF DIE FRAGE 'WARUM P'
1	0	0	0	0	0	<i>Das ist eben (bekanntlich immer) so.</i>
0	1	0	0	0	0	Ableitungsverweise verbalisieren
0	1	1	0	0	0	
0	0	1	0	0	0	Teile der Inferenzkette verbalisieren
1	0	1	0	0	0	
0	0	0	1	0	0	Zurückweisung ohne Begründung
0	0	0	0	1	1	Zurückweisung mit Begründung
0	0	0	0	0	1	
0	0	0	1	0	1	
0	0	0	0	0	0	<i>Darüber kann ich nichts sagen</i>
1	0	0	1	0	0	Inkonsistenz der Wissensbasis
<input type="checkbox"/>						
<input type="checkbox"/>						

$\text{mA} \hat{=}$ mit Ableitungsverweis

$\text{oA} \hat{=}$ ohne Ableitungsverweis

Fig.5

8. Die Auswahl einer kommunikativ adäquaten Antwort bei Mehrfachableitungen

Bisher haben wir das Problem der Auswahl einer kommunikativ adäquaten Antwort aus einer Menge möglicher Antworten auf eine 'Warum'-Frage dadurch vereinfacht, dass wir im Gegensatz zur allgemeinen Definition in Abschnitt 6 stets nur einen der möglichen Ableitungswege für eine bestimmte Konklusion betrachtet haben.

Antworten auf 'Warum'-Fragen können dann durch einen Algorithmus mit folgenden drei Hauptkomponenten generiert werden:

- 1) der einer Antwort zugrundeliegende Inferenzprozess wird in Form eines linearisierten Zielbaums in einem als Inferenzgedächtnis bezeichneten Teil der Wissensbasis gespeichert. Information über den Wahrheitswert der Prämissen und der abgeleiteten Konklusion und über den Informativitätswert der Inferenzregel sowie eine Angabe zur eindeutigen Zuordnung des Zielbaums zu der Frage, die den Inferenzprozess auslöste, wird jeweils zusammen mit den einzelnen Teilzielbäumen im Inferenzgedächtnis abgelegt.
- 2) Das Fragewort *Warum* löst einen Suchprozess aus, der aus der Menge der während eines Dialogs gespeicherten Zielbäume denjenigen bestimmt, auf den sich die Frage bezieht. In diesem Zielbaum wird dann nach einer Teilinferenzkette mit maximalem Informativitätswert gesucht, wobei im Fall von iterierten 'Warum'-Fragen bereits verbalisierte Teilketten übergangen werden.
- 3) Der in 2) ausgewählte Teilzielbaum wird in eine natürlichsprachliche Oberflächenstruktur überführt.

Im folgenden werden zwei Typen von Inferenzen diskutiert, bei denen es sinnvoll wird, mehrere alternative Ableitungswege für eine Konklusion zu betrachten. Es werden heuristische Kriterien angegeben, aufgrund derer aus einer Menge von alternativen Ableitungswegen derjenige bestimmt werden kann, der am besten als Ausgangspunkt für die weitere Suche nach einer kommunikativ adäquaten Antwort auf eine 'Warum'-Frage geeignet ist.

Approximative Inferenzen bei vagem Wissen

In Alltagssituationen werden häufig Inferenzregeln verwendet, die nur abgeschwächte Implikationen darstellen. Bei der Anwendung solcher approximativer Inferenzen auf meist nur partiell erfüllte Prämissen muss der subjektive Wahrheitswert des Inferenzergebnisses entsprechend eingeschränkt werden (vgl. (24)). Wenn sich ein Inferenzergebnis zusätzlich auf weiteren Inferenzwegen unter Verwendung anderer Prämissen ableiten lässt, so erhöht sich der subjektive Wahrheitswert der Konklusion entsprechend. Dieses Inferenzverhalten kann mit Hilfe der F-Logik (engl.: fuzzy logic) modelliert und durch Und/Oder-Graphen dargestellt werden (vgl. Fig. 6). Bei der Generierung einer kommunikativ adäquaten Antwort auf die Frage *Warum P?* muss bei dem in Fig. 6 dargestellten Schema eines Inferenzprozesses eine der Inferenzketten

$A \overset{.7}{\Rightarrow} E \overset{.8}{\Rightarrow} P$ und $B \nexists C \overset{.3}{\Rightarrow} D \overset{.2}{\Rightarrow} P$ für die weitere Suche ausgewählt werden. Eine plausible Heuristik ist dabei die Wahl derjenigen Inferenzkette, welche den höchsten Beitrag zu dem durch Kombination der einzelnen Inferenzergebnisse entstandenen Wahrheitswert der Konklusion leistet. Falls die Prämissen A, B und C in Fig. 6 uneingeschränkt gültig sind, ergibt sich aufgrund dieser Heuristik $A \overset{.7}{\Rightarrow} E \overset{.8}{\Rightarrow} P$ als Basis für eine Antwort. Gemäß dem oben beschriebenen Schritt 2) wird dann innerhalb dieser Inferenzkette nach einer Teilkette mit maximaler Informativität gesucht.

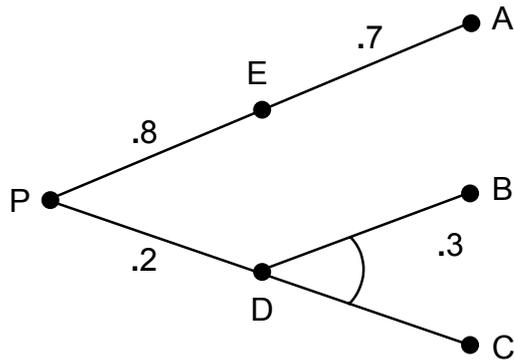


Fig. 6

Hypothetische Inferenzen bei unvollständigem Wissen

In Alltagssituationen verfügt eine befragte Person oft nur über unvollständiges Wissen über einen bestimmten Sachverhalt. Um bestimmte Fragen trotzdem beantworten zu können, muss der Befragte häufig hypothetische Inferenzen durchführen. Dabei wird oft von Annahmen ausgegangen, die sich lediglich darauf stützen, dass es aufgrund mangelnden Wissens nicht möglich ist, die Verneinung der einer Annahme zugrundeliegenden Proposition zu beweisen. Solche Annahmen sind eine typische Form von nicht-monotonem Wissen (vgl. (8)), da es jederzeit möglich ist, dass sie aufgrund neu hinzukommenden Wissens oder aufgrund eines während des Inferenzprozesses erkannten Widerspruchs zurückgenommen werden müssen. Hypothetische Inferenzen können im Rahmen einer nicht-monotonen Logik modelliert werden (vgl. (8)). Wenn Inferenzergebnisse in starkem Maße auf unsicheren Annahmen beruhen, ist es zweckmäßig, sich alternative Ableitungswege (vgl. Fig. 7), die während des Inferenzprozesses gefunden werden, zu merken, da sich einige der Ableitungswege später als nicht valide erweisen können. Bei der Generierung einer kommunikativ adäquaten Antwort auf eine 'Warum'-Frage ergibt sich dann allerdings wieder das Problem der Auswahl einer Inferenzkette.

Bei dieser Auswahl ist es eine plausible Heuristik, sich stets für die kürzeste Inferenzkette zu entscheiden (z.B. in Fig. 7: $A \Rightarrow P$).

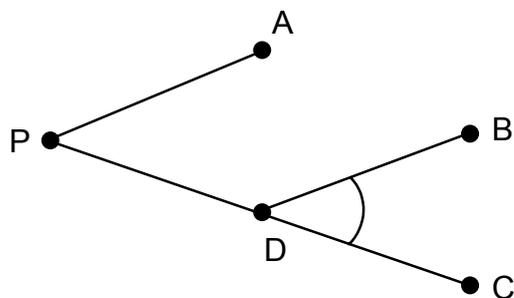


Fig. 7

An dieser Stelle soll betont werden, dass die skizzierten Algorithmen zur Beantwortung von 'Warum'-Fragen — auch wenn sie, wie zuletzt diskutiert, im Kontext der F-Logik oder der nicht- monotonen Logik arbeiten — nur erste, noch sehr stark eingeschränkte Ansätze zur Formalisierung von Alltagswissen und komplexem menschlichem Problemlöseverhalten sind, die mit dem Ziel verfolgt werden, Aspekte intelligenten Verhaltens auf Rechenanlagen zu simulieren und für praktische Anwendungen verfügbar zu machen.

9. Ausblick: Das Dialogsystem stellt 'Warum'-Fragen

In den vorangegangenen Abschnitten wurden ausschließlich Dialogsequenzen betrachtet, in denen ein KI-System 'Warum'-Fragen eines Benutzers/Bearbeiters beantwortet. In diesem abschließenden Abschnitt soll darauf hingewiesen werden, dass in KI-Systemen auch die Simulation von Dialogsequenzen wie:

BE: < Frage >

SY: *Warum willst Du das wissen?*

zweckmäßig sein kann. Antworten auf solche 'Warum'-Fragen enthalten oft wichtige Hinweise über die Intention des Fragestellers. Die Information kann dann die vom System bei der Antwortsuche verfolgte Strategie beeinflussen oder als Entscheidungshilfe bei der in Abschnitt 8 dargestellten Auswahl einer kommunikativ adäquaten Antwort dienen. Antworten auf 'Warum'-Fragen des Systems werden m.W. bisher in keinem KI-System dazu benutzt, treffendere Antworten auf Anfragen des Benutzers zu finden. Eine Dialogsequenz wie:

BE: *Warum P?*

SY: *Warum willst Du das wissen?*

zeigt aber, dass die Auswertung von Antworten auf 'Warum'-Fragen des Systems ein erster Schritt zur Beseitigung des wesentlichen Defekts aller für die Beantwortung von 'Warum'-Fragen in der KI bisher vorgeschlagenen Verfahren ist. Der Defekt besteht darin, dass die Intention des Fragestellers beim Prozess der Antwortsuche kaum berücksichtigt wird.

LITERATUR

- (1) **Aqvist, L.** (1975): A new approach to the logical theory of interrogatives: Analysis and formalization, 2. erw. Aufl. Tübingen.
- (2) **Bromberger, S.** (1966): Why-questions, in: Colodny, R. (Ed.): Mind and Cosmos, Pittsburg, pp. 86-111.
- (3) **Charniak, E.** (1976): Inference and knowledge I, II, in: Charniak, E. / Wilks, Y. (Eds.): Computational Semantics, Amsterdam, pp. 1-21, 129-154.
- (4) **Cohen, P.R.** (1978): On knowing what to say: Planning speech acts, Toronto. (= Univ. of Toronto, Dept. of Computer Science, Technical Report No. 118.)
- (5) **Conrad, R.** (1978): Studien zur Syntax und Semantik von Frage und Antwort, Berlin. (= Studia Grammatica XIX.)
- (6) **Davis, R.** (1976): Applications of meta level knowledge to the construction, maintenance and use of large knowledge bases, Stanford. (= Stanford Univ., Dept. of Computer Science, Report STAN-CS-76-562.)
- (7) **Dilger, W.** (1976): Ein Frage-Antwort-System auf der Basis einer prädikatenlogischen Sprache, in: Lausch, J./Schneider, H.-J. (Eds.): Dialoge in natürlicher Sprache und Darstellung von Wissen, Freudenstadt, S. 31-50.
- (8) **Doyle, J.** (1978): Truth maintenance systems for problem solving, Cambridge, Mass. (= MIT AI-Lab., Report AI-TR-419.)
- (9) **Egli, U. / Schleicher, H.** (1976): A bibliography on the theory of questions and answers, in: Linguistische Berichte 41, S. 105-128.
- (10) **Ficht, H.** (1978): Supplement to a bibliography on the theory of questions and answers, in: Linguistische Berichte 55, S. 92-114.
- (11) **Freeman, C.** (1976): A pragmatic analysis of tenseless why-questions, in: Mufwene S. S. / Walker, C. A. / Steeuer, S. B. (Eds.): Papers of the twelfth regional meeting of the Chicago Linguistic Society, Chicago, pp. 208-219.
- (12) **Hahn v., W. / Jameson, A. / Hoepfner, W. / Wahlster, W.** (1978): HAM-RPM: Natural dialogues with an artificial partner, in: Proc. of the AISB/GI Conference on Artificial Intelligence, Hamburg, pp. 122-131 (auch erschienen als: Bericht Nr. 3 der Projektgruppe 'Simulation von Sprachverstehen', Univ. Hamburg, Germanisches Seminar, Feb. 1978).
- (13) **Hahn v., W.** (1979): Probleme der Simulationstheorie und Fragepragmatik bei der Simulation natürlichsprachlicher Dialoge, in: Ueckert, H. / Rhenius, D. (Eds.): Komplexe menschliche Informationsverarbeitung. Beiträge zur Tagung 'Kognitive Psychologie' in Hamburg. Bern, S. 260- 269 (auch erschienen als: Bericht Nr. 6 der Projektgruppe 'Simulation von Sprachverstehen', Univ. Hamburg, Germanisches Seminar, Mai 1978).
- (14) **Hart, P.E. / Duda, R.O.** (1977): PROSPECTOR - A computer-based consultation system for mineral exploration, Stanford. (= Stanford Research International, AI Center, Technical Note 155.)
- (15) **Heringer, H.J.** (1974): Praktische Semantik, Stuttgart.
- (16) **Heringer, H.J. / Öhlschläger, G. / Strecker, R. / Wimmer, R.** (1977): Einführung in die Praktische Semantik, Heidelberg.
- (17) **Kleer de, J. / Doyle, J. / Steele, G.L. / Sussman, G.J.** (1977): AMORD - Explicit control of reasoning, in: Proc. of the Symposium on Artificial Intelligence and Programming Languages, S. 116-125. (= SIGART Newsletter 64.)

- (18) **Lehnert, W.** (1977): Human and computational question answering, in: *Cognitive Science*, Vol. 1, No. 1, pp. 47-73.
- (19) **Nagel, H.-H.** (1977): Analyzing sequences of TV-frames: System design considerations, Hamburg. (= Univ. Hamburg, Fachbereich für Informatik, Bericht IFI-HH-B-33/77.)
- (20) **Sacerdoti, E.D.** (1977). A structure for plans and behavior, New York, N. Y.
- (21) **Scott, C.A. / Clancey, A. / Davis, R. / Shortliffe, E.H.** (1977): Explanation capabilities of production-based consultation systems, in: *American Journal of Computational Linguistics*, Microfiche 62.
- (22) **Stallman, R.M. / Sussman, G.J.** (1977): Forward reasoning and dependency directed backtracking in a system for computer-aided circuit analysis, in: *Artificial Intelligence* 9, pp. 135-196.
- (23) **Teller, P.** (1974): On why-questions, in: *Nous* 8, pp. 371-380.
- (24) **Wahlster, W.** (1977): Die Repräsentation von vagem Wissen in natürlichsprachlichen Systemen der Künstlichen Intelligenz, Hamburg. (= Univ. Hamburg, Fachbereich für Informatik, Bericht IFI-HH-B-38/77.)
- (25) **Wahlster, W.** (1979). Die Simulation vager Inferenzen auf unscharfem Wissen: Eine Anwendung der mehrwertigen Programmiersprache FUZZY, in: Ueckert, H. / Rhenius, D. (Eds.): *Komplexe menschliche Informationsverarbeitung. Beiträge zur Tagung 'kognitive Psychologie' in Hamburg.* Bern, S. 249-259 (auch erschienen als: Bericht Nr. 5 der Projektgruppe 'Simulation von Sprachverstehen', Univ. Hamburg, Germanisches Seminar, Mai 1978).
- (26) **Wahlster, W./Jameson, A./Hoepfner, W.** (1978): Glancing, referring and explaining in the dialogue system HAM-RPM, Hamburg. (= Bericht Nr. 7 der Projektgruppe 'Simulation von Sprachverstehen', Univ. Hamburg, Germanisches Seminar, Juli 1978.)
- (27) **Wilks, Y.** (1977): What sort of taxonomy of causation do we need for language understanding?, in: *Cognitive Science*, Vol. 1, No. 3, pp. 235-264.
- (28) **Winograd, T.** (1972): *Understanding natural language*, New York, N.Y.
- (29) **Wright v., G.H.** (1974): *Erklären und Verstehen*, Frankfurt/M.
- (30) **Wunderlich, D.** (1976): Fragesätze und Fragen, in: Wunderlich, D.: *Studien zur Sprechakttheorie*, Frankfurt/M., S. 181ff.