BERICHT
aus dem
PSYCHOLOGISCHEN INSTITUT
DER UNIVERSITÄT HEIDELBERG

Theo Held, Martin Schrepp und Stefan Fries

Vergleich von Methoden zur Ableitung von Wissensstrukturen

Dezember 1994

Bericht Nr. 81

69117 HEIDELBERG, HAUPTSTR. 47-51, TEL. 06221/54 73 46

Inhaltsverzeichnis

1 Einführung **3**

2 Aufgabenmaterial und Daten **5**

3 Methoden zur Ableitung von Wissensstrukturen **7**
3.1 Datenanalyse: Item Tree Analyse (Van Leeuwe, 1974) 8
3.1.1 Theoretischer Hintergrund 8
3.1.2 Anwendung des Verfahrens 10
3.1.3 Ergebnisse 10
3.1.4 Diskussion des Verfahrens 11
3.2 Expertenbefragung: Der Algorithmus von Dowling (1993a) . . . 13
3.2.1 Theoretischer Hintergrund 13
3.2.2 Anwendung des Verfahrens 14
3.2.3 Ergebnisse 16
3.2.4 Diskussion des Verfahrens 18
3.3 Aufgabenkonstruktion und Aufgabenanalyse: Sequenzeninklusion (Albert, Schrepp & Held, 1994) 19
3.3.1 Theoretischer Hintergrund 19
3.3.2 Anwendung des Verfahrens 21
3.3.3 Ergebnisse 21
3.3.4 Diskussion des Verfahrens 21

4 Vergleich der Ergebnisse **23**

5 Diskussion einer Integration der Verfahren **26**

Anhang **31**

Die Arbeit entstand innerhalb des Projektes „Wissensstrukturen“ (Lu-385/2) der *Deutschen Forschungsgemeinschaft* unter der Leitung von PD Dr. J. Lukas und Prof. Dr. D. Albert.

Zusammenfassung

Die Theorie der Wissensräume Doignon & Falmagne (1985) bildet eine Voraussetzung für die Entwicklung und Anwendung einer Klasse ökonomischer Prozeduren zur Wissensdiagnose. Diese Prozeduren basieren auf der Existenz eines „Wissensraumes“ auf einer Item-Menge. Ein Problem stellt die angemessene Konstruktion solcher Wissensräume für gegebene Item-Mengen dar. Drei Methoden zur Konstruktion von Wissensräumen werden in dieser Arbeit hinsichtlich ihrer praktischen Anwendbarkeit und ihrer Ergebnisse verglichen: (1) Inhaltliche Aufgabenanalyse, (2) Expertenbefragung und (3) Analyse von Antwortmustern. Diese drei Verfahren wurden zur Konstruktion von Wissensräumen auf einer Menge von 16 Schachaufgaben angewandt. Zum Vergleich der Verfahren werden die resultierenden Wissensräume den in zwei experimentellen Untersuchungen beobachteten Antwortmustern gegenübergestellt.

1 Einführung

Die vorliegende Untersuchung stellt drei Verfahren gegenüber, die dazu geeignet sind, Aufgabenstrukturen im Sinne der Theorie der Wissensräume nach Doignon & Falmagne (1985) abzuleiten. Im Rahmen dieser Theorie ist die Erzeugung wohlfundierter Aufgabenstrukturen ein zentrales Problem. Um wissensdiagnostische Prozeduren (vgl. Falmagne & Doignon, 1988a; Falmagne & Doignon, 1988b) anwenden zu können, müssen solche (empirisch validierten) Strukturen existieren.

Zunächst werden einige Grundbegriffe der Wissensraumtheorie dargestellt, die für das Verständnis der folgenden Ausführungen unerläßlich sind (für eine genauere Darstellung sei auf Doignon & Falmagne, 1985 und Falmagne, Koppen, Villano, Doignon & Johannesen, 1990 verwiesen).

Annahmen bzgl. struktureller Eigenschaften von Aufgabenmengen werden nach Doignon & Falmagne (1985) in Form einer „Surmise-Relation" oder einer „Surmise-Funktion" formuliert. Eine Surmise-Relation $\preceq$ ist eine reflexive und transitive binäre Relation auf einer Aufgabenmenge Q. Für zwei Aufgaben $p, q \in Q$ interpretiert man $p \preceq q$ als „Jede Person, die q lösen kann, kann auch p lösen." Ein „Wissenszustand" $K \subseteq Q$ bezüglich einer gegebenen Surmise-Relation $\preceq$ wird wie folgt definiert:

$$K \subseteq Q \text{ ist ein Wissenszustand } :\Leftrightarrow \forall q, t \in Q\ (q \preceq t \wedge t \in K \Rightarrow q \in K).$$

Die Familie $\mathcal{F}$ aller Wissenszustände, die mit einer Surmise-Relation korrespondieren wird als „quasi-ordinaler Wissensraum" (Doignon & Falmagne, 1985) bezeichnet.[1] Quasi-ordinale Wissenräume sind abgeschlossen bzgl. Vereinigung und Durchschnitt. Als Verallgemeinerung und 'realistischere' Variante dieser Struktur führen Doignon & Falmagne (1985) das Konzept der Surmise-Funktion ein. Eine solche Surmise-Funktion ist eine Abbildung

$$\sigma : Q \to 2^{2^Q} \setminus \{\emptyset\},$$

die jedem Element $q \in Q$ eine Menge von Teilmengen von Q zuweist. Doignon & Falmagne (1985) interpretieren $\sigma(q) = \{K_1, \ldots, K_n\}$ für eine Aufgabe $q \in Q$ als „Jede Person, die q lösen kann, muß auch alle Aufgaben aus mindestens einer der Mengen $K_1, \ldots, K_n$ lösen können." Die Mengen $K_1, \ldots, K_n$ werden als *Klauseln* von q bezeichnet. Aufgrund dieser Interpretation wird vorausgesetzt, daß eine Surmise-Funktion drei Bedingungen erfüllt. Erstens muß eine Aufgabe Element jeder ihrer Klauseln sein. Ist zweitens eine Aufgabe p Element einer Klausel K einer Aufgabe q, so muß eine in K enthaltene Klausel von p existieren. Drittens

[1] In der Terminologie von Doignon & Falmagne werden Familien von Wissenszuständen generell als „Wissensstrukturen" bezeichnet, wenn nicht genauer Bezug auf die Eigenschaften der Familie genommen wird.

kann keine Klausel von q echte Teilmenge einer anderen Klausel von q sein, d. h. die Klauseln sind in diesem Sinne minimal bezüglich der Teilmengenbeziehung.

Durch Surmise-Funktionen sind die Fälle von Wissensstrukturen berücksichtigt, in denen Aufgaben alternative Mengen von Vorausetzungs-Aufgaben zugeordnet werden sollen. Dies ist z. B. dann sinnvoll, wenn Aufgaben über unterschiedliche Lösungswege gelöst werden können. Eine Surmise-Funktion σ bestimmt in eindeutiger Weise einen Wissensraum $\mathcal{W}$, d. h. eine bezüglich Vereinigungsbildung abgeschlossene Menge von Wissenszuständen, durch

$$W \in \mathcal{W} :\Leftrightarrow \forall q \in W \exists S \in \sigma(q)(S \subseteq W).$$

Die eins-zu-eins Korrespondenz zwischen quasi-ordinalen Räumen und Quasi-Ordnungen ist seit Birkhoff (1937) bekannt. Die Leistung der Wissensraum-Theorie besteht insbesondere im Nachweis der eins-zu-eins Beziehung zwischen Surmise-Funktionen und Wissensräumen (siehe Corollary 3.8 in Doignon & Falmagne, 1985).

Ein (quasi-ordinaler) Wissensraum wird im allgemeinen nur einen Teil der $2^{|Q|}$ Teilmengen einer Aufgabenmenge Q enthalten. Dies ermöglicht eine ökonomische Diagnose des Wissenszustandes einer Person (siehe Falmagne & Doignon, 1988a und Falmagne & Doignon, 1988b). Man gibt dabei — ausgehend vom gesamten Wissensraum $\mathcal{W}$ — Fragen aus Q vor und eliminiert nach jeder Antwort einer Person alle nicht mit dieser Antwort kompatiblen Elemente von $\mathcal{W}$. Dies wird solange fortgesetzt, bis nur noch ein Element verbleibt, das dem diagnostizierten Wissenszustand der Person entspricht. Dieses Diagnose-Verfahren benötigt bei geeigneter Fragenwahl („half-split"-Prozedur nach Falmagne & Doignon, 1988a) nur etwa $\log_2(|\mathcal{W}|)$ Fragen für die Diagnose des Wissenszustandes einer Person, d. h. bei kleinem $\mathcal{W}$ müssen nicht alle Fragen aus Q bearbeitet werden. Die zentrale Voraussetzung zur Anwendung der beschriebenen wissensdiagnostischen Prozedur ist also die Bestimmung einer möglichst *kleinen* Wissensstruktur.

Die untersuchten Methoden zur Bestimmung einer solchen Struktur unterscheiden sich bezüglich ihres theoretischen Konzepts, der Durchführung und der erforderlichen Voraussetzungen für die Anwendung. Gemeinsam ist den Verfahren, daß mit ihrer Hilfe zumindest Surmise-Relationen bzw. quasi-ordinale Wissensräume bestimmt werden können. Folgende Verfahren sollen im folgenden besprochen und angewandt werden:

- *Datenanalyse:* Es existieren zahlreiche Verfahren, mit denen auf der Basis von existierendem Datenmaterial Wissensstrukturen erzeugt werden können. Deterministische Verfahren werden z. B. von Airasian & Bart (1973), Bart & Krus (1973), Bart & Airasian (1974), Van Leeuwe (1974) und Van Buggenhaut & Degreef (1987) vorgeschlagen. Probabilistische

Ansätze stellen Falmagne (1989a, 1989b) und Villano (1991) vor. Die Anwendung dieser probabilistischen Ansätze setzt jedoch die Existenz sehr großer Datensätze voraus, um die erforderlichen Parameterschätzungen durchführen zu können. Aus diesem Grund werden die probablistischen Verfahren hier nicht weiter berücksichtigt. Innerhalb der vorliegenden Arbeit soll die Methode der „Item Tree Analysis" nach Van Leeuwe (1974) angewandt werden (vgl. auch Held & Korossy, in prep.).

- *Expertenbefragung:* Von Bedeutung sind in diesem Rahmen die Arbeiten von Müller (1989), Koppen & Doignon (1990), Dowling (1993a) und Koppen (1993). Diese Verfahren zeichnen sich dadurch aus, daß sie eine relativ ökonomische Befragung von Experten ermöglichen. Die Verfahren sind konzipiert, um Surmise-Systeme bzw. Wissensräume abzuleiten. In dieser Arbeit wird der Algorithmus von Dowling (1993a) angewandt.

- *Systematische Aufgabenkonstruktion und -analyse:* Verfahren werden von Albert (1989), Albert et al. (1994), Albert & Held (1994), Held (1993), Schrepp (1993) und Korossy (1993) vorgeschlagen. Im Rahmen dieser Untersuchung wird das Prinzip der „Sequenzen-Inklusion" (vgl. Albert et al., 1994) angewandt. Dieses Prinzip ermöglicht die Ableitung von Surmise-Relationen bzw. quasi-ordinalen Wissensräumen anhand inhaltlicher Annahmen bezüglich der zur Problemlösung notwendigen kognitiven Fähigkeiten.

Ziel der Untersuchung ist es, die Anwendung dieser Verfahren am Beispiel einer Aufgabenmenge zu demonstrieren und die Verfahren bezüglich der Validität der erzeugten Wissensstrukturen zu vergleichen. Außerdem sollen die spezifischen Vor- und Nachteile der Verfahren diskutiert werden. Als Konsequenz daraus wird dargestellt, wie jede der Methoden innerhalb des Prozesses der Bestimmung einer validen Wissensstruktur eingesetzt werden kann.

Im folgenden Abschnitt soll auf die Art der in dieser Vergleichsstudie verwendeten Aufgaben und das Zustandekommen der Daten eingegangen werden. Insbesondere wird dargestellt, wie diese Aufgaben inhaltlich analysiert werden können, d.h. welche Wissenskomponenten bei der Lösung eine Rolle spielen. Dies ist ausschließlich für die Konstruktion von Wissensräumen durch inhaltliche Analyse bedeutsam, d. h. die beiden anderen in dieser Arbeit geschilderten Verfahren Datenanalyse und Expertenbefragung kommen ohne eine solche Analyse aus.

2 Aufgabenmaterial und Daten

Die verwendete Aufgabenmenge besteht aus 16 taktischen Schachproblemen (siehe Tabelle 7 im Anhang). Bei solchen Problemen wird eine Schachposition, d.h.

eine Konfiguration von Figuren auf dem Schachbrett, vorgegeben und die Aufgabe besteht darin, den besten Zug für die am Zug befindliche Seite zu finden.

Zwei solche Schachprobleme sind in Abbildung 1 dargestellt. Das Ziel in beiden Problemen besteht für die schwächere Seite (Weiß) darin, durch taktische Kombinationen, d. h. erzwungene Zugfolgen, ein Remis zu erreichen. In beiden Positionen ist jeweils Weiß am Zuge, d.h. es ist der beste Zug von Weiß zu ermitteln. Die Lösung für Problem (a) ist in der üblichen Notation gegeben

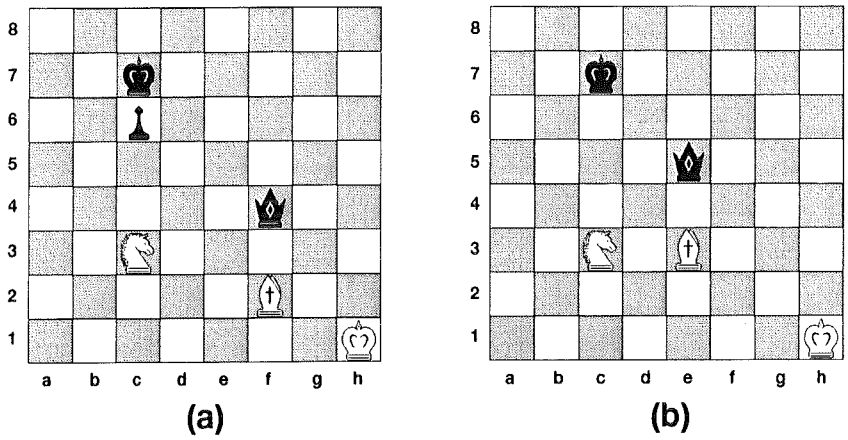


Abbildung 1. Beispiele für den in der Untersuchung verwendeten Typ von Aufgaben.

durch 1. Sd5+ cd5: 2. Lg3 Dg3: Patt. Die Lösung für Problem (b) ist 1. Lf4 Df4: 2. Sd5+ K $\sim$ 3. Sf4:.

Um solche Probleme lösen zu können, muß also eine solche Zugfolge ausgearbeitet werden, die zum gewünschten Ziel führt. Für die Ausarbeitung einer solchen Zugfolge ist Wissen über taktische Standardideen, sogenannte *Motive*, eine wesentliche Vorraussetzung.

Die 16 Aufgaben wurden in Hinblick auf eine Kodierung durch Motive konstruiert. Die vorkommenden Motive sind „Hinlenkung“ (Kodierung: G), „Gabel“ (F), „Beseitigung“ (E), „Räumung“ (C), „Patt“ (S) und „Unterverwandlung“ (T).[2]

Wir wollen nun das Motivkonzept an unseren beiden Beispielaufgaben kurz erläutern. Das Ziel für Weiß in Problem (a) ist, ein Remis durch Erzwingen eines Patts zu erreichen. Hierfür muß zuerst der eigene Springer auf c3 beseitigt werden. Der Zug 1. Sd5+ erzwingt diese Beseitigung (cd5:). Der Zug 2. Lg3 lenkt die schwarze Dame auf das Feld g3 und erreicht ein Remis durch Patt. Also sind die Motive „Beseitigung“, „Hinlenkung“ und „Patt“ in dieser Reihenfolge, für die Lösung von Aufgabe (a) zu erkennen. Das Ziel für Weiß in Problem (b) ist durch Gewinn der schwarzen Dame, ein Remis zu erreichen. Der Zug 1. Lf4

[2]Die Kodierungssymbole wurden aus den englischsprachigen Artikeln von Albert et al. (1994) und Schrepp, Held & Albert (1994) übernommen. Deshalb stehen die Symbole in keinem Zusammenhang zu den deutschsprachigen Motivnamen.

lenkt die schwarze Dame auf das Feld f4 und 2. Sd5+ gewinnt die schwarze Dame durch eine Gabel. Damit sind die Motive „Hinlenkung" und „Gabel" in dieser Reihenfolge für die Lösung von Problem (b) zu erkennen.

Wie man aus diesen beiden Beispielen ersieht, beruht der Zug Sd5+ in sehr ähnlichen Positionen auf unterschiedlichen Motiven (Beseitigung bei (a) gegenüber Gabel bei (b)). Motive sind also Teilziele im Lösungsprozeß. Detailierte Informationen zum Motivkonzept finden sich in Albert et al. (1994).

Hinsichtlich des Schwierigkeitsgrades der Aufgaben muß bemerkt werden, daß diese für Hobbyspieler konzipiert waren, d.h. Spieler, die keinem Schachklub angehören und sich nie systematisch mit Schach beschäftigten.

Im folgenden wollen wir noch kurz auf das Zustandekommen des in dieser Arbeit verwendeten Datenmaterials eingehen. Die 16 Schachaufgaben wurden in zwei Untersuchungen jeweils 46 Versuchspersonen (Vpn) zur Bearbeitung vorgelegt. Es liegen also insgesamt Datensätze von 92 Vpn vor. Es traten 69 unterschiedliche Antwortvektoren auf, wobei acht Vpn alle Aufgaben lösen konnten und sechs Vpn keine der Aufgaben lösten (siehe Tabelle 8 im Anhang).

Die Datenerhebung erfolgte rechnergesteuert. Die Vpn der ersten Untersuchung wurden durch Zeitungsinserate angeworben und für ihre Teilnahme mit 12,– DM honoriert. Diese Untersuchung fand am Psychologischen Institut der Universität Heidelberg statt. Die zweite Untersuchung wurde am Psychologischen Institut der Universität Graz (Österreich) durchgeführt.[3] Bis auf wenige Ausnahmen waren alle Vpn reine Hobbyspieler, gehörten also keinem Schachklub an.

3 Methoden zur Ableitung von Wissensstrukturen

In diesem Abschnitt werden die drei untersuchten Verfahren im Detail dargestellt. Für jedes Verfahren wird zunächst der theoretische Hintergrund erläutert, dann wird das Verfahren auf die beschriebenen Aufgaben bzw. Daten angewandt. Zum Abschluß der Darstellung werden die spezifischen Vor- und Nachteile des jeweiligen Verfahrens und die Ergebnisse der Untersuchungen diskutiert.

[3]Wir danken Herrn Prof. Dr. D. Albert (Universität Graz, Österreich), unter dessen Anleitung die Untersuchung durchgeführt wurde, dafür, daß er uns die Daten zur Verfügung gestellt hat. Außerdem danken wir den Herren P. Mörtl und H. Mariacher (beide Universität Graz) für die Durchführung der Untersuchung.

3.1 Datenanalyse: Item Tree Analyse (Van Leeuwe, 1974)

3.1.1 Theoretischer Hintergrund

Die Anwendung der hier beschriebenen Methode basiert — ebenso wie die oben erwähnten verwandten Methoden — auf der Existenz empirisch gewonnener Daten (Antwortmuster). Aus diesen Daten werden quasi-ordinale Wissensräume abgeleitet. Die Item Tree Analyse (ITA) kann nur auf dichotome Antwortmuster angewandt werden, also z. B. Antworten innerhalb eines Wissenstests, die entweder als falsch oder richtig kodiert werden.

Die Quasi-Ordnung wird mit Hilfe von 2×2-Matrizen abgeleitet, die für jedes Paar von Items aufgestellt werden. Abbildung 2 zeigt zwei solche Tabellen.

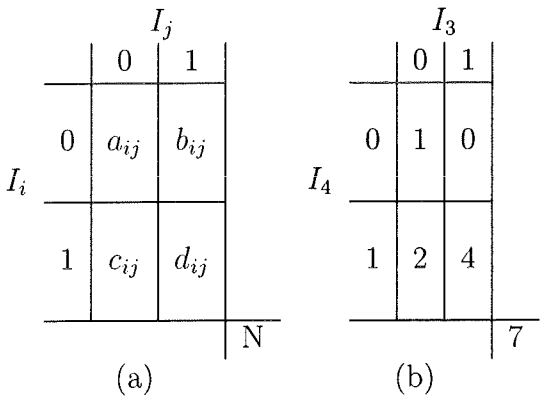


Abbildung 2. ITA: Item-Matrizen

Jede dieser Matrizen gibt an, wie häufig für zwei Items I_j und I_i die folgenden Fälle aufgetreten sind: (1) beide Items wurden falsch beantwortet (a_{ij} in Abbildung 2), (2) Item I_j wurde richtig und Item I_i falsch beantwortet (b_{ij}), (3) Item I_j wurde falsch und Item I_i richtig beantwortet (c_{ij}) und (4) beide Items wurden richtig beantwortet (d_{ij}). Aus Matrix (b) in Abbildung 2 geht hervor, daß z. B. Items I_4 und I_3 von einer der (fiktiven) Versuchspersonen gemeinsam falsch beantwortet wurden. Vier Vpn haben beide Items richtig beantwortet.

Für die Anwendung der Item-Tree-Analyse ist der Wert b_{ij} von besonderer Bedeutung. Wenn b_{ij} gleich null oder kleiner gleich einem festgelegten Toleranzlevel L ist, wird angenommen, daß das jeweilige Paar von Items ein Element der abzuleitenden Relation $\leq^L$ ist. Die Relation $\leq^L$ ist somit durch die folgende Bedingung definiert:

$$I_i \leq^L I_j :\Leftrightarrow b_{ij} \leq L.$$

Van Leeuwe (1974) zeigt, daß die Relation $\leq^L$ für $L = 0$ stets transitiv ist. Wenn L größer Null ist, muß die Transitivität der resultierenden Relation überprüft werden.

Zur Beurteilung der Passungsgüte der abgeleiteten Relation $\leq^L$ in Bezug auf den gegebenen Datensatz schlägt Van Leeuwe (1974) die Berechnug eines

„Correlational-Agreement-Koeffizienten“ (CA) vor. Im Prinzip werden für die Bestimmung dieses Koeffizienten die beobachteten Korrelationen zwischen Items mit den erwarteten Korrelationen unter der Annahme, daß $\leq^L$ die „richtige“ Relation ist, verglichen. Der CA-Koeffizient ist folgendermaßen definiert:

$$\mathrm{CA} := 1 - \frac{2}{n(n-1)} \sum_{i<j} (r_{ij} - r^*_{ij})^2,$$

wobei n die Anzahl der Items und r_{ij} die Pearson-Korrelation (Phi-Koeffizient; vgl. z. B. Bortz, 1993, S. 210) für Items i und j bezeichnet. Der Wert r^*_{ij} wird definiert durch

$$r^*_{ij} = \begin{cases} 1 & \text{wenn } i \leq^L j \text{ und } j \leq^L i \\ \sqrt{\frac{(1-p_i)p_j}{p_i(1-p_j)}} & \text{wenn } i \leq^L j \text{ und } j \not\leq^L i \\ \sqrt{\frac{(1-p_j)p_i}{p_j(1-p_i)}} & \text{wenn } j \leq^L i \text{ und } i \not\leq^L j \\ 0 & \text{sonst.} \end{cases}$$

Die relative Häufigkeit von Probanden, die Item i richtig beantwortet haben wird als p_i bezeichnet, p_j gibt die relative Häufigkeit von Probanden an, die Item j richtig beantwortet haben. Wir werden diesen Wert lediglich dazu verwenden, die „beste“ Lösung der ITA zu ermitteln (s. u.).

Als Kriterium für die Passung von Wissensräumen auf gegebene Daten werden wir im folgenden den von Schrepp (1993) vorgeschlagenen Quotienten aus der durchschnittlichen minimalen symmetrischen Mengendifferenz[4] zwischen Daten und Wissensstruktur („ddat“) und der durchschnittlichen minimalen Mengendifferenz zwischen der Potenzmenge der Aufgabenmenge und der Wissensstruktur („dpot“) verwenden.[5]

Dieser Wert wird als „Distance-Agreement-Coefficient“ (DA) bezeichnet. Es gilt also:

$$\mathrm{DA} := \frac{\mathrm{ddat}}{\mathrm{dpot}}.$$

Dieser Wert muß ebenso wie CA als ad-hoc Kriterium angesehen werden. Mit Hilfe von DA wird ebenso wie bei CA der Trade-Off zwischen Größe und Passung einer Wissensstruktur berücksichtigt. Im Gegensatz zu CA wird eine umso bessere Passung der Struktur angenommen, je kleiner DA ist. Für eine genauere Erläuterung dieses Kriteriums sei auf Schrepp (1993) und Schrepp et al. (1994) verwiesen.

[4] Die symmetrische Mengendifferenz d zwischen zwei Mengen A und B wird definiert als $d(A,B) = \mid A \Delta B \mid$, wobei $A \Delta B = (A \setminus B) \cup (B \setminus A)$. Die Wissenszustände, die am 'nächsten' an einem Antwortvektor liegen, sind die Zustände, die sich in den wenigsten Einträgen vom Antwortvektor unterscheiden.

[5] Formal gilt $\mathrm{ddat} := \sum_{P \in D} \min(\{d(P,W) \mid W \in \mathcal{W}\}) / \mid D \mid$, wobei D die Menge aller beobachteten Antwortvektoren beschreibt, und $\mathrm{dpot} := \sum_{P \in \mathcal{P}} \min(\{d(P,W) \mid W \in \mathcal{W}\}) / \mid \mathcal{P} \mid$, wobei $\mathcal{P}$ die Potenzmenge der Aufgabenmenge Q bezeichnet.

3.1.2 Anwendung des Verfahrens

Nach einem Vorschlag von Van Leeuwe (1974) wird die 'beste' Struktur diejenige, die (1) transitiv ist und (2) einen möglichst hohen CA-Wert ergibt. Die konkrete Vorgehensweise bei der Ermittlung der Struktur beginnt mit der Suche nach dem Toleranzlevel L, bei dem CA maximal ist. Die resultierende Struktur wird auf Transitivität geprüft. Falls Intransitivitäten auftreten, muß das Toleranzlevel gesenkt werden bis eine transitive Relation gefunden ist.

Im Falle des hier untersuchten Datensatzes ist CA bei $L = 8$ maximal. Die Relation für $L = 8$ ist jedoch nicht transitiv. Erst für $L = 6$ ($CA = 0.95$) kann eine transitive Lösung gefunden werden (DAT_1).

Um den Einfluß der verwendeten Daten auf die Ergebnisse der ITA zu verdeutlichen, wurden die 92 Antwortpatterns gemischt und in zwei Hälften zu je 46 Patterns aufgeteilt. Mit den Patterns der ersten Hälfte wurde eine Relation per ITA bestimmt (einzige transitive Lösung bei $L = 0$). Die Passung auf diese Relation wurde sowohl für den kompletten Datensatz bestimmt (DAT_2), als auch für die zweite Hälfte des (gemischten) Datensatzes (DAT_2^*).

Schließlich wurden die beiden Hälften des Datensatzes, die aus unterschiedlichen Untersuchungen stammen getrennt analysiert. Für die erste Hälfte (Heidelberger Untersuchung) ergibt sich das beste Ergebnis bei einem Toleranzlevel von 1 (CA = 0.90). Der korrespondierende quasi-odinale Raum DAT_3 enthält 232 Zustände. Das Ergebnis der zweiten Hälfte (Grazer Untersuchung) wurde bei $L = 2$ (CA = 0.92) gefunden. Die Kardinalität des entsprechenden Raumes DAT_4 ist 121.

3.1.3 Ergebnisse

In Tabelle 1 sind die die symmetrischen Mengendifferenzen (Distanzen), die durchschnittlichen Distanzen ($\overline{d}$), die CA-Werte und die DA-Werte für die erzeugten quasi-ordinalem Wissensräume angegeben. In Abbildung 3 ist das

Tabelle 1. Ergebnisse Datenanalyse

Struktur	# Zustände	Distanzen								$\overline{d}$	CA	DA
		0	1	2	3	4	5	6	7			
DAT_1	57	43	30	11	7	1	–	–	–	0.84	0.95	0.18
DAT_2	496	72	16	3	1	–	–	–	–	0.27	0.87	0.07
DAT_3	232	56	27	6	3	–	–	–	–	0.52	0.90	0.13
DAT_4	121	45	35	10	2	–	–	–	–	0.66	0.92	0.15
DAT_2^*	496	26	16	3	1	–	–	–	–	0.54	0.88	0.16

* Ergebnis für den halben gemischten Datensatz.

Hasse-Diagramme der Surmise-Relation für DAT_1 dargestellt. Abbildung 4 zeigt die Hasse-Diagramme für DAT_3 (links) und DAT_4 (rechts). Der quasi-ordinale

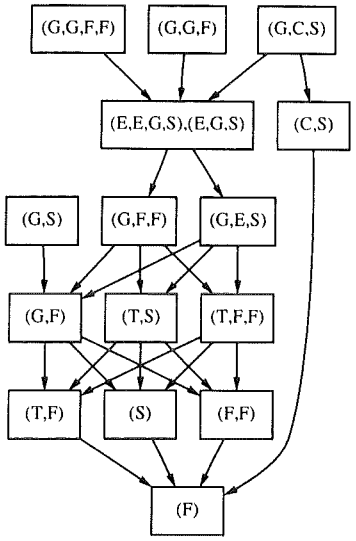


Abbildung 3. Hasse-Diagram zur Struktur DAT_1

Raum DAT_1 besteht aus 57 Zuständen. 43 der in den Daten enthaltenen Antwortvektoren stimmen mit einem der Zustände überein (Distanz 0). Genau 30 der Vektoren unterscheiden sich in einem Eintrag vom nächstliegenden Zustand (Distanz 1) etc. Die durchschnittliche Distanz für DAT_2 ist deutlich geringer als die für DAT_1. In diesem Zusammenhang ist jedoch zu beachten, daß DAT_2 mehr als die achtfache Anzahl von Zuständen enthält. Dieses Verhältnis wird auch deutlich, wenn man die CA-Werte betrachtet. Trotz deutlich niedrigerem $\overline{d}$ ist der CA-Wert für DAT_2 niedriger als der für DAT_1. Die Passungsgüte von DAT_1 wird gemäß dem Kriterium CA als höher eingestuft als die von DAT_2.

Beim Vergleich der Strukturen DAT_3 und DAT_4 ist zu bemerken, daß für die beiden Datensätze bezüglich der allgemeinen Kennwerte relativ unterschiedliche Ergebnisse zu beobachten sind. DAT_3 enthält nahezu zweimal so viele Zustände wie DAT_4. Die Passung von DAT_3 kann aufgrund des DA-Wertes dennoch als besser angesehen werden. Die beiden Strukturen stimmen in lediglich 54 Zuständen überein (s. u., Tab. 5).

3.1.4 Diskussion des Verfahrens

Datenanalytische Verfahren wie die ITA sind universell anwendbar auf beliebige Sätze von dichotomen Daten. Unabhängig von Überlegungen zur Validität der erzeugten Strukturen, stellen datenanalytische Verfahren ein leicht handhabbares „Werkzeug“ zur übersichtlichen Visualisierung von Daten dar. Falls man annehmen kann, daß die erhobenen Daten für einen untersuchten Bereich repräsentativ sind (Voraussetzung dafür ist meist eine sehr große Stichprobe und eine kleine Itemmenge), bieten datenanalytische Verfahren die Möglichkeit, in einfacher Form Wissensstrukturen zu erzeugen, die dann auch in wissensdiagno-

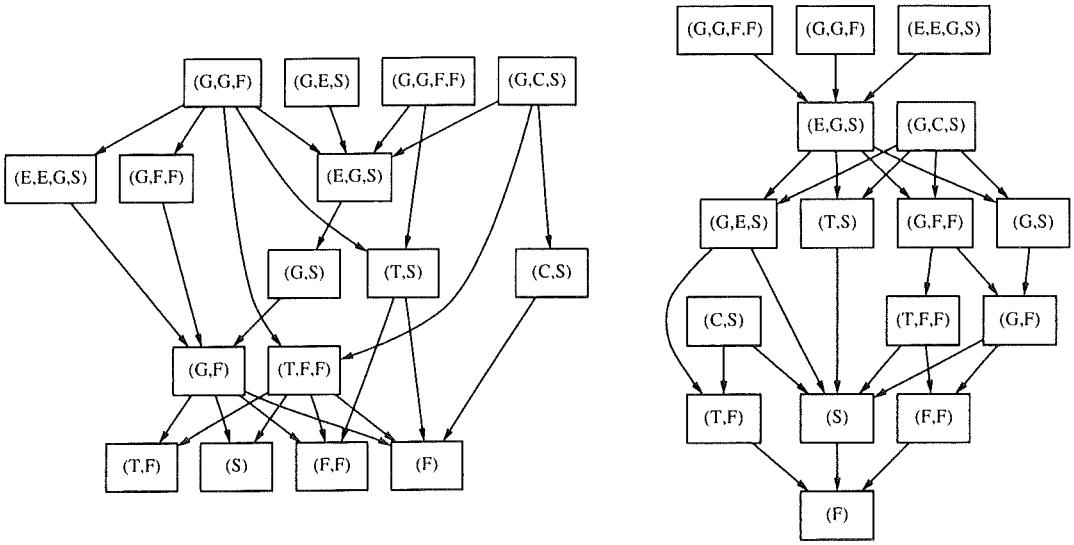


Abbildung 4. Hasse-Diagram zu den Strukturen DAT_3 (links) und DAT_4 (rechts)

stischen Verfahren eingesetzt werden können.

Die ITA in der vorgestellten Form ist auf die Ableitung quasi-ordinaler Wissensräume beschränkt. Für Wissensdomänen in denen alternative Voraussetzungen für Wissensinhalte bestehen können, ist die ITA somit nicht geignet. Generell (und trivialerweise) sind die Ergebnisse datenanalytischer Verfahren von der Qualität der analysierten Daten abhängig. Wenn deutliche Decken- oder Bodeneffekte auftreten, kann eine Datenanalyse natürlicherweise kaum zu verwertbaren Ergebnissen führen. In solchen Fällen liefert die ITA große Äquivalenzklassen von Aufgaben, da wegen der Häufigkeit der Fälle, in denen bei einem Paar von Items entweder beide oder kein Item gelöst wurde, die b_{ij} Werte sehr niedrig (bzw. gleich 0) sind und sich somit sehr viele Paare von Items in der abgeleiteten Relation befinden.

Der Einsatz der ITA empfielt sich also nur bei sehr großen Datensätzen und kleinen Itemmengen. Außerdem muß die Bandbreite der Probanden-Fähigkeiten innerhalb der untersuchten Stichprobe möglichst hoch sein. Die Vereinigung der Ergebnisse beider Untersuchungen führt im Vergleich zu DAT_3 und DAT_4 zu einer wesentlich kompakteren Struktur (DAT_1). Dies kann als als Hinweis darauf gesehen werden, daß eine große Anzahl von Daten eine Voraussetzung für die Anwendung der ITA darstellt.

Der Vergleich von DAT_1 und DAT_2 zeigt, daß sich nach dem (zufälligen) Entfernen von Lösungspatterns aus dem Datensatz eine Struktur ergibt (DAT_2), die wesentlich mehr Zustände enthält als DAT_1. Dies ist darin begründet, daß der reduzierte Datensatz weniger Information enthält und somit auch weniger Paare von Aufgaben in die abgeleitete Surmise-Relation eingehen können.

Zugleich problematisch und hilfreich bei der Anwendung der ITA ist die Tat-

sache, daß die resultierende Relation bezüglich der (vermuteten) inhaltlichen Struktur eines Wissensgebietes absolut unplausible Elemente enthalten kann. Der Grund dafür kann z. B. sein, daß eine sehr inhomogene Itemmenge verwendet wurde, oder daß Fehler in der Versuchsdurchführung zu einer Verzerrung der Ergebnisse führten. Für die Ableitung empirisch valider Surmise-Relationen kann es jedoch hilfreich sein, die Ergebnisse der ITA mit den Ergebnissen anderer (inhaltlich fundierter) Verfahren (s. u.) zu vergleichen, um potentielle Störfaktoren innerhalb der Validierungsprozedur erkennen zu können.

3.2 Expertenbefragung: Der Algorithmus von Dowling (1993a)

3.2.1 Theoretischer Hintergrund

Wenn eine Surmise-Relation bzw. ein Surmise-System erstellt werden soll, ohne daß auf vorhandene Daten zurückgegriffen werden kann, liegt es nahe, auf externe Ressourcen bezüglich des zu untersuchenden Wissensgebietes zurückzugreifen. Eine solche Ressource ist das Wissen von Experten. Die bereits oben erwähnten Ansätze zur Expertenbefragung von Koppen und Dowling/Müller bieten adaptive Algorithmen, mit deren Hilfe die Antworten von Experten zur ökonomischen Ableitung von Surmise-Relationen und Surmise-Systemen verwendet werden können.

Da die Experten nicht direkt über Wissenszustände in der Aufgabenmenge befragt werden können, werden ihnen Fragen zu Beziehungen zwischen Teilmengen der Aufgabenmenge gestellt. Die Antworten der Experten werden dazu benutzt, Elemente aus der Potenzmenge der Aufgabenmenge auszuschließen, die mit den Expertenantworten nicht kompatibel sind. Das Format der Fragen bezüglich einer Aufgabenmenge Q sieht folgendermaßen aus:

> *„Angenommen eine Person hat alle Aufgaben in einer Teilmenge A aus Q falsch beantwortet. Ist es dann praktisch sicher, daß diese Person auch Aufgabe q aus Q falsch beantworten wird? Leichtsinnsfehler und das Raten von Antworten sollen ausgeschlossen sein."*

Der Experte muß eine solche Frage entweder akzeptieren (d. h. mit 'ja' beantworten) oder zurückweisen (verneinen). Wir notieren eine Frage vom vorgestellten Typ mit $A \rightarrow q$.

Die Menge aller Fragen $A \rightarrow q$, die vom Experten akzeptiert werden, etabliert eine Relation $\rightarrow$ auf $2^Q \setminus \{\emptyset\} \times Q$. Diese Relation wird als „Implikationsrelation" bezeichnet. Falls eine Implikationsrelation $\rightarrow$ die beiden Bedingungen

$$\text{(i)} \quad \forall p \in A \; (A \rightarrow p),$$
$$\text{(ii)} \quad \forall p \in A \; (O \rightarrow p) \wedge (A \rightarrow q) \Rightarrow (O \rightarrow q)$$

erfüllt, bestimmt sie eineindeutig einen Wissensraum $\mathcal{W}$ auf Q (siehe Corollary 2.5 in Dowling, 1993a) durch die folgende Bedingung:

$$W \in \mathcal{W} \Leftrightarrow \forall A \subseteq Q \; \forall q \in Q \; (A \rightarrow q \wedge A \subseteq Q \setminus W \Rightarrow q \in Q \setminus W).$$

Ohne Befragungsalgorithmen mit deren Hilfe Folgerungen aus den Expertenantworten gezogen werden, muß ein Experte im ungünstigsten Fall für eine Aufgabenmenge mit m Aufgaben $(2^m - 1) \times m$ Fragen beantworten. Mit Hilfe der Algorithmen von Dowling (1993a) bzw. Koppen (1993) wird diese Anzahl durch Ausnutzen der Abschlußeigenschaften (i) und (ii) erheblich reduziert. Dennoch werden auch diese Prozeduren ab einer Anzahl von ca. 20 Aufgaben sehr zeitaufwendig und damit kostenintensiv.

Für die hier beschriebene Untersuchung wurde ein Rechnerprogramm verwendet, das auf der Basis des Ansatzes von Dowling (1993a) arbeitet.[6]

3.2.2 Anwendung des Verfahrens

Das verwendete Befragungsprogramm wurde an der Technischen Universität Braunschweig unter Anleitung von PD Dr. C. E. Dowling entwickelt (Programmversion „tr73"). Das Programm übernimmt die Selektion der günstigsten Frage an den Experten, das Ziehen von Schlußfolgerungen aus den Antworten und die Erstellung der Basis[7] des resultierenden Wissensraumes. Für die intendierte Anwendung mit Experten des Wissensgebietes Schach wurde noch zusätzlich die Möglichkeit realisiert, den kompletten Ablauf der Befragung zu protokollieren.

Als Experten dienten drei erfahrene Turnierspieler unterschiedlicher Spielstärke[8] mit unterschiedlicher Erfahrung in der Vermittlung von Schachwissen an Novizen:

- Experte EXP1: Internationaler Meister, wenig Erfahrung in der Wissensvermittlung.
- Experte EXP2: Mittlere Spielstärke, umfangreiche Erfahrung in der Wissensvermittlung.
- Experte EXP3: Mittlere Spielstärke, keine Erfahrung in der Wissensvermittlung.

[6] Das Programm wurde uns freundlicherweise von Frau PD Dr. C. E. Dowling, Technische Universität Braunschweig zur Verfügung gestellt.

[7] Die Basis B eines Wissensraumes $\mathcal{W}$ ist eine bzgl. $\subseteq$ minimale Teilmenge von $\mathcal{W}$ derart, daß sich jedes Element von $\mathcal{W}$ als Vereinigung von Elementen von B darstellen läßt, d. h. $\mathcal{W} = \{B_1 \cup \ldots \cup B_n \mid n \in \mathbb{N} \wedge B_1, \ldots, B_n \in B\}$.

[8] Die Spielstärke von Schachspielern wird in Deutschland durch die sogenannte DWZ (Deutsche Wertungszahl) gemessen. EXP1 besitzt eine DWZ von 2390, EXP2 eine DWZ von 1960, während für EXP3 keine solche Zahl existiert.

Die Befragung der Experten wurde am Psychologischen Institut der Universität Heidelberg durchgeführt. Die Teilnahme an der Untersuchung dauerte ca. 3 Stunden und wurde pauschal mit DM 100.- honoriert.

Nach einer allgemeinen Instruktion (siehe Instruktion 1 in Abbildung 7 im Anhang) wurden alle 16 Aufgaben den Experten in Diagrammform vorgelegt. Die Experten hatten zunächst die Aufgabe, jedes der Schachprobleme zu lösen (siehe Instruktion 2 in Abbildung 7). Ihnen stand dafür ein Schachspiel in Standardgröße zur Verfügung. Keiner der Experten benötigte mehr als 15 Minuten, um alle Aufgaben zu bearbeiten, da die Aufgaben für reine Hobbyspieler konzipiert waren und deshalb von der Schwierigkeit für die Experten kein Problem darstellten.

Anschließend wurden den Experten in einem Trainingsdurchgang drei Fragen vorgelegt, die in ihrem Aufbau denen der Befragungsprozedur entsprechen. Allgemein hatte die Fragen folgende Form (siehe Instruktion 3 in Abbildung 8 im Anhang):

> „Ein Schachspieler hat die Aufgaben X, Y, ... nicht gelöst. Sehen Sie es als praktisch sicher an, daß dieser Spieler auch Aufgabe Z nicht lösen wird?"

In diesem Trainingsdurchgang sollten die Experten mit dem (ungewöhnlichen) Fragenformat vertraut gemacht werden. Es sollte außerdem überprüft werden, ob die Fragen von den Experten richtig verstanden wurden.

Nach einer abschließenden Instruktion (siehe Instruktion 4 in Abbildung 8) startete der Versuchsleiter — sofern seitens des Experten keine Fragen mehr waren — das Befragungsprogramm. Während der Untersuchung laß der Versuchsleiter die Fragen vom Bildschirm des Rechners ab und legte die entsprechenden Diagramme dem Experten vor (getrennt nach Prämissen und Konsequenzen). Die Antwort des Experten wurde vom Vl in den Rechner eingegeben.

Die Befragungsprozedur begann mit den Fragen mit einelementigen Prämissen (also: „Ein Schachspieler hat Aufgabe X nicht gelöst. Sehen Sie es als praktisch sicher an, daß dieser Spieler auch Aufgabe Y nicht lösen wird?"). Nachdem alle Fragen mit einelementigen Prämissen gestellt waren, wurde das Zwischenergebnis gesichert (also die resultierende *Surmise-Relation*) und dann mit der Untersuchung fortgefahren, bis das Programm keine weiteren Fragen mehr zu stellen hatte. Maximal traten Fragen mit drei Aufgaben in der Prämisse auf.

In Anschluß an die Befragung wurden die Experten gebeten, eine hierarchische Struktur der Aufgaben (im Sinne einer Surmise-Relation) aufzuzeichnen (siehe Instruktion 5 in Abbildung 9 im Anhang). Dafür standen Karten mit den Aufgabennummern und vorgefertigte „Verbindungspfeile" zur Verfügung. Der Sinn dieser Prozedur lag, neben der Gegenüberstellung der Verfahren an sich, insbesondere darin, zu überprüfen, welche Grobstruktur die Experten der

Aufgabenmenge zuordneten und ob erhebliche Diskrepanzen zwischen den Resultaten der Befragungsprozedur und den gezeichneten Strukturen bestanden (z. B. wegen Antwortfehlern im Laufe der Befragungsprozedur).

3.2.3 Ergebnisse

Tabelle 2 enthält eine Übersicht der Ergebnisse. Für jeden der Experten sind drei Wissensstrukturen aufgelistet. Die Strukturen mit der Extension „1el" resultierten nach der Beantwortung aller Fragen mit einelementigen Prämissen. Die „end"-Strukturen stellen die Resultate der kompletten Befragungsprozedur dar. Die Ergebnisse des abschließenden Strukturlegens finden sich bei den „su"-Strukturen. Abbildung 5 zeigt als Beispiel für ein Resultat der Expertenbefragung und eine gelegte Struktur die Hasse-Diagramme zu den Strukture EXP2.1el und EXP2.su. Die Gegenüberstellung der Strukturen der Experten EXP1 und EXP3 ist in den Abbildungen 10 und 11 im Anhang zu finden.

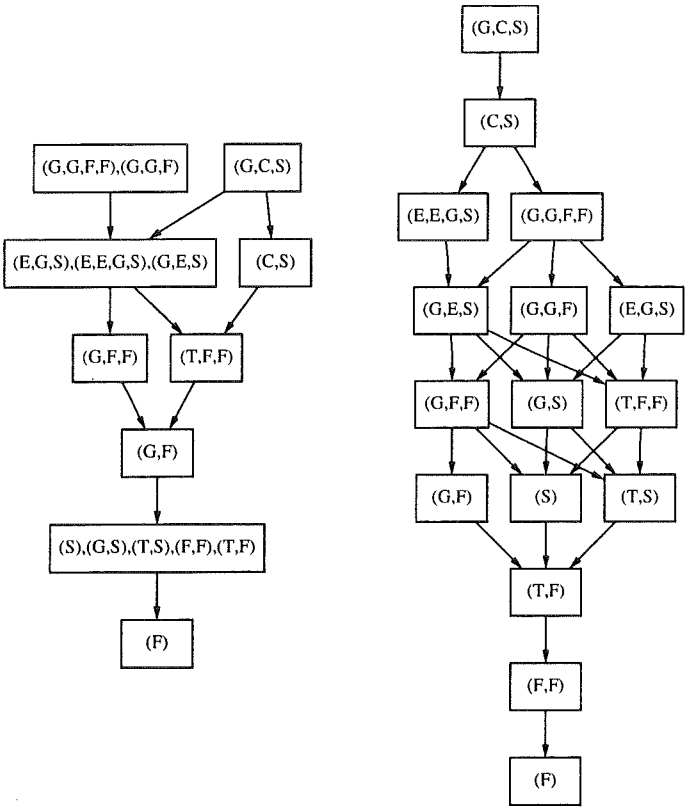


Abbildung 5. Beispiele für eine aus der Befragungsprozedur resultierende Surmise-Relation (EXP2.1el, links) und eine gelegte Struktur (EXP2.su, rechts)

Außerdem wurde für jeden der drei Strukturtypen die Vereinigungsmenge der Einzelergebnisse gebildet („∪"-Strukturen) und die resultierenden Strukturen mit den Daten verglichen. Die „∪∪"-Strukturen stellen den Vereinigungsab-

schluß der ∪-Strukturen dar.[9]

Für alle Strukturen sind die Distanzverteilung, die mittlere symmetrische Mengendifferenz ($\overline{d}$) und der DA-Wert angegeben. Experte EXP1 mußte insge-

Tabelle 2. Ergebnisse Expertenbefragung

Struktur	# Zustände	Distanzen								$\overline{d}$	DA
		0	1	2	3	4	5	6	7		
EXP1.1el	23	17	23	26	17	6	3	–	–	1.79	0.35
EXP1.end	20	17	18	31	16	7	3	–	–	1.86	0.36
EXP1.su	30	20	26	25	12	8	1	–	–	1.62	0.33
EXP2.1el	15	15	24	31	14	5	3	–	–	1.77	0.33
EXP2.end	15	15	24	31	14	5	3	–	–	1.77	0.33
EXP2.su	36	26	40	14	6	6	–	–	–	1.19	0.23
EXP3.1el	11	18	15	21	15	14	7	1	1	2.24	0.41
EXP3.end	10	18	15	21	15	14	7	1	1	2.24	0.41
EXP3.su	11	14	15	19	29	9	5	1	–	2.25	0.41
∪*.1el*	38	20	33	24	10	3	2	–	–	1.45	0.30
∪*.end*	34	20	28	29	9	4	2	–	–	1.51	0.31
∪*.su*	62	30	38	17	4	3	–	–	–	1.04	0.23
∪∪*.1el*	47	20	34	24	10	3	1	–	–	1.40	0.30
∪∪*.end*	42	20	28	30	9	3	2	–	–	1.49	0.29
∪∪*.su*	79	30	39	16	4	3	–	–	–	1.03	0.22

'*EXP2.1el*' und '*EXP2.end*' sind identisch. Die ∪-Strukturen enthalten die Vereinigung der Einzelstrukturen. Die ∪∪-Strukturen enthalten die bzgl. Vereinigung abgeschlossenen ∪-Strukturen.

samt 85 Fragen mit einelementigen Prämissen und 5 Fragen mit mehrelementigen Prämissen beantworten. Bei EXP2 waren es 74 einelementige Prämissen und 4 mehrelementige. Die Befragung von EXP3 war nach 62 einelementigen und einer mehrelementigen Prämisse abgeschlossen.

Die Übereinstimmung der Expertenstrukturen untereinander kann aus der Kardinalität der Schnittmengen der einzelnen Strukturtypen ersehen werden. Bei den 1el-Strukturen gibt es 3 gemeinsame Wissenszustände. Für die end-Strukturen und die su-Strukturen sind es ebenfalls 3 Zustände. Wenn man die Ergebnisse von EXP3 vernachlässigt, so ergeben sich für EXP1 und EXP2 jeweils 6 gemeinsame Zustände.

Die von den Experten nach der Befragungsprozedur gelegten Strukturen weisen bei EXP1 und EXP2 geringere durchschnittliche Distanzen zur Datenmenge und DA-Werte auf. Bei EXP3 steigt $\overline{d}$ geringfügig, wobei der DA-Wert konstant

[9] Die ∪∪-Strukturen wurden mit Hilfe eines Programmes erstellt, das auf einem Algorithmus von Dowling (1993b) basiert. Wir danken Frau PD Dr. C. E. Dowling und Herrn C. Hockemeier (beide TU Braunschweig) dafür, daß sie uns dieses Programm zur Verfügung gestellt haben.

bleibt.

3.2.4 Diskussion des Verfahrens

Das vorgestellte Verfahren ermöglicht eine sehr effiziente und ökonomische Acquisition von Expertenwissen. Die Fragen an den Experten beziehen sich direkt auf das Aufgabenmaterial, das in einer wissensdiagnostischen Prozedur verwendet werden soll. Sofern die vorgegebene Itemmenge nicht zu umfangreich ist, können mit Hilfe der Expertenbefragung Wissensräume erzeugt werden, ohne daß eine Datenerhebung oder eine inhaltliche Analyse des Aufgabenmaterials erforderlich ist. Im Gegensatz zur ITA führt die Expertenbefragung nicht nur zu Surmise-Relationen, sondern auch zu Surmise-Funktionen.

Andererseits ist die vorgestellte Prozedur nur dann effizient, wenn der Experte „fehlerfrei", d. h. während der gesamten Befragung völlig konsistent antwortet. Unterlaufen dem Experten Fehler, so kann es wegen der Schlußfolgerungen, die der Algorithmus zieht, zu starken Abweichungen von den Strukturen kommen, die eigentlich der Expertenmeinung entsprechen. In einer Untersuchung von Schrepp & Held (1994) wird gezeigt, daß bereits eine Fehlerquote von einem Prozent zu einer erheblichen Verzerrung der Ergebnisse führen kann. Auch bei der hier vorgestellten Untersuchung zeigte sich, daß das Auftreten von Urteilsfehlern kaum vermieden werden kann. Als besonders „fehlerträchtige" Phasen sind Anfang und Schlußphase der Befragung anzusehen. Dies liegt zum einen daran, daß die Experten lernen müssen, mit dem ungewöhnlichen Format der Fragen umzugehen und daß zum anderen Ermüdungserscheinungen während einer Befragungsprozedur kaum zu vermeiden sind.

Generell genügt es nicht, Spezialisten für das untersuchte Wissensgebiet als Experten auszuwählen. Gemäß Kriterien für Experten, wie sie z. B. von Einhorn (1986), Gaeth & Shanteau (1986) und Shanteau (1988) angeführt werden, kann eine Person erst als Experte bezeichnet werden, wenn sie sowohl über umfangreiches relevantes Fachwissen verfügt, als auch hochtrainiert bezüglich der bei einer Befragung gegebenen Entscheidungssituation ist. Vor dem Beginn einer Befragung sollte also stets ein intensives Expertentraining erfolgen. Nach Shanteau (1988) ist es für Experten wichtig, die Konsequenzen ihrer Antworten für das Gesamtergebnis der Befragung einschätzen zu können, um eine „Justierung" des Antwortverhaltens zu ermöglichen. Einhorn (1986) stellt die Entscheidungssituation des Experten in der Terminologie der Signalentdeckungstheorie (vgl. Green & Swets, 1966) dar. Gemäß dieser Sichtweise können die Anteile des vorgegebenen Items, die für die Entscheidung des Experten relevant sind als Reiz betrachtet werden, der durch das „Rauschen" der irrelevanten Anteile überlagert wird. Ein optimaler Experte ist in der Lage, das Vorhandensein eines Reizes relativ sicher zu erkennen. Er neigt also im Sinne der Signalentdeckungstheorie weder zu „False-Alarms", noch zu „Misses." Durch solche möglichen Antworttenden-

zen von Experten können die erheblichen Abweichungen der Expertenurteile untereinander teilweise begründet werden. Weitere Ursachen für Abweichungen zwischen Experten können natürlich in unterschiedlichen Auffassungen zu einzelnen Wissensinhalten und im Grad der Expertise liegen.

Im hier vorgestellten Beispiel mit Schachexperten zeigt sich, daß der beste Schachspieler (EXP1) nicht das qualitativ beste Ergebnis liefert. Das Ergebnis, das mit den Daten am ehesten konform geht, stammt von EXP2. Dieser Experte verfügt zwar nicht über die hohe Spielstärke von EXP1, er ist aber im Bereich der Vermittlung elementaren Schachwissens als höher trainiert ausgewiesen. Es ist anzunehmen, daß diese Art von Expertise für den untersuchten Wissensbereich einschlägiger ist als Spielstärke.

Die Ergebnisse von EXP3 weichen sehr stark von den Ergebnissen der anderen beiden Experten ab. Dies mag darin begründet sein, daß EXP3 zum Zeitpunkt der Befragung über eine relativ geringe aktuelle Spielpraxis verfügte und der Experte außerdem keine Erfahrung in der Vermittlung elementaren Schachwissens verfügt. Grundsätzlich sind bei der Expertenbefragung auch motivationale Aspekte zu beachten. Die Länge der Prozedur und die Eintönigkeit der Fragen können zu Fehlantworten der Experten führen. Außerdem ist nicht auszuschließen, daß die Experten aufgrund der Systematik der Prozedur ein Antwortverhalten entwickeln, das darauf abzielt, die Prozedur zu verkürzen. Tatsächlich ist es so, daß eine häufige Bejahung der Fragen mit einelementiger Prämisse dazu führt, daß sich die Prämisse schneller ändert (d. h. Blöcke von Fragen mit identischer Prämisse sind schneller beendet). Für zukünftige Untersuchungen muß somit darauf geachtet werden, daß der Ablauf des Verfahren für die Experten so wenig transparent wie möglich gestaltet wird.

3.3 Aufgabenkonstruktion und Aufgabenanalyse: Sequenzeninklusion (Albert et al., 1994)

3.3.1 Theoretischer Hintergrund

Das hier vorgestellte Verfahren der Sequenzeninklusion gehört zu einer Gruppe von Prinzipien zur Ableitung von Surmise-Relationen aus komponentenbasierten Aufgaben. Für die Darstellung der anderen Verfahren (z. B. Mengeninklusion und Multisetinklusion) sei auf Albert et al. (1994) und Lukas & Albert (1993) verwiesen. Da u. E. die Sequenzeninklusion das geignetste Verfahren zur Strukturierung der gegebenen Aufgabenmenge (siehe Abschnitt 2) ist, beschränken wir uns auf die Darstellung dieses Verfahrens.

Wie bereits in Abschnitt 2 erwähnt, wird aufgrund der Konstruktion bzw. Analyse der verwendeten Schachaufgaben angenommen, daß zur Lösung der Aufgaben besimmte taktische Standardsituationen, die sogenannten Motive, bekannt sein müssen. Eine zentrale Annahme des Prinzips Sequenzeninklusion ist

weiterhin, daß die Sequenz in der die Motive innerhalb einer Aufgabe auftreten maßgeblich für den Lösungsprozeß und damit auch für die Aufgabenschwierigkeit ist.

Im folgenden gehen wir aus von einer endlichen Menge M von Motiven. Der „Aufgabenraum" $P(M)$ ist definiert als die Menge aller Schachaufgaben, die mit Hilfe der Motive in M kodiert werden können. Wir nehmen an, daß eine Aufgabe p schwieriger zu lösen ist als eine Aufgabe q, wenn die Sequenz von Motiven, die eine Person kennen und erkennen muß, um Aufgabe p zu lösen, die Motivsequenz von Aufgabe q enthält.

Für eine Schachaufgabe $p \in P(M)$ sei $H(p)$ das geordnete Tupel von Motiven, die im Lösungsweg von p vorkommen. Wir schreiben $H(p) = (m_1, m_2, \ldots, m_k)$, wenn das erste auftretende Motiv m_1, das zweite m_2 und das letzte Motiv m_k ist. Jedes Motiv kann innerhalb einer Motivsequenz mehrmals auftreten. Wir definieren nun:

$$\begin{aligned} M_k &:= \{(m_1, \ldots, m_k) \mid m_1, \ldots, m_k \in M\}, \\ M_{I\!N} &:= \bigcup_{k \in I\!N} M_k. \end{aligned}$$

M_k ist also die Menge aller k-Tupel von Motiven aus M, während $M_{I\!N}$ die Menge aller Motivtupel beliebiger Länge von Motiven aus M bezeichnet. Unter der Annahme, daß die Schwierigkeit einer Schachaufgabe $p \in P(M)$ nur von $H(p)$ abhängt, identifizieren wir für die Ableitung einer Surmise-Relation eine Aufgabe p mit $H(p)$. Eine Relation $\preceq_{SI}$ auf der Aufgabenmenge $P(M)$ wird mit Hilfe einer Quasi-Ordnung $\sqsubseteq$ auf $M_{I\!N}$ definiert:

$$p \preceq_{SI} q \;:\Leftrightarrow H(p) \sqsubseteq H(q).$$

p und q sind Aufgaben in $P(M)$. Die Quasi-Ordnung auf $M_{I\!N}$ ist für $(m_1, \ldots, m_k)$, $(m'_1, \ldots, m'_l) \in M_{I\!N}$ durch $(m_1, \ldots, m_k) \sqsubseteq (m'_1, \ldots, m'_l)$ genau dann, wenn eine Funktion $f : \{1, \ldots, k\} \to \{1, \ldots, l\}$ existiert, die die folgenden Bedingungen erfüllt:

(i) $\forall i, j \in \{1, \ldots, k\}\ (i < j \to f(i) < f(j))$,
(ii) $\forall j \in \{1, \ldots, k\}\ (m_j \sqsubseteq_M m'_{f(j)})$.

Die Relation $\sqsubseteq_M$ ist entsprechend der Annahme, daß unterschiedliche Motive möglicherweise unterschiedlich leicht erkannt werden können, für die Menge M von Motiven definiert. Wir nehmen an, daß $\sqsubseteq_M$ eine Quasi-Ordnung auf M etabliert und interpretieren $M_1 \sqsubseteq_M M_2$ für zwei Motive $M_1, M_2 \in M$ als „Jede Person, die in der Lage ist, M_2 zu erkennen, ist auch in der Lage, M_1 zu erkennen."[10] Die Relation $\preceq_{SI}$ hängt also von den durch $\sqsubseteq_M$ getroffenen Annahmen bezüglich der einzelnen Motive ab.

[10] Eine genaue Erklärung der Eigenschaften von $\sqsubseteq_M$ findet sich in Albert et al. (1994).

3.3.2 Anwendung des Verfahrens

Die Anwendung der Sequenzeninklusion führt zu den Strukturen, die in Abbildung 6 gezeigt sind. Durchgezogene Kanten stellen im Hassse-Diagramm Struktur SI_2 dar, die sich unter der Annahme ergibt, daß keine Schwierigkeitsunterschiede bezüglich des Erkennens und Verarbeitens der auftretenden Motive bestehen.[11] Für Struktur SI_1 kommen die gepunkteten Kanten noch hinzu. Diese Struktur ist begründet durch die Annahme, daß die Gabel das am leichtesten zu erkennende und verarbeitende Motiv ist.[12] Der mit Struktur SI_1 korrespondierende quasi-ordinale Wissensraum enthält 368 Zustände. Die durch schwächere Annahmen fundierte Struktur SI_2 führt zu einem Raum mit 1025 Zuständen.

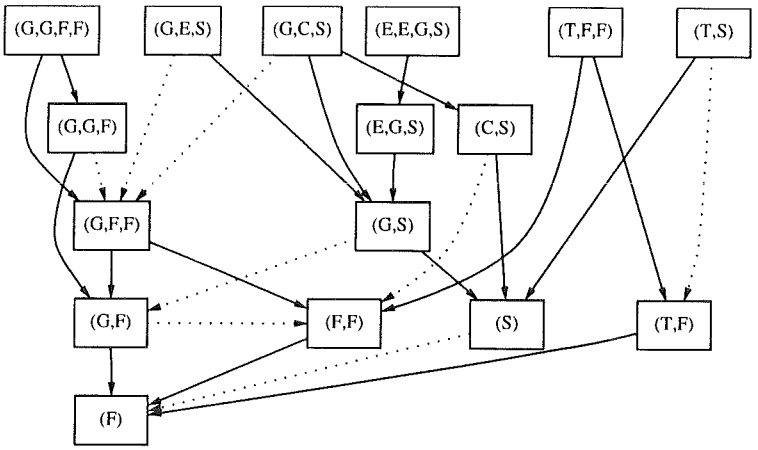


Abbildung 6. Aus der Sequenzeninklusion resultierende Surmise Relationen.

3.3.3 Ergebnisse

In Tabelle 3 sind die Distanzen der Datenpatterns zu den Strukturen SI_1 und SI_2, sowie die durchschnittlichen Distanzen ($\overline{d}$) und die DA-Werte dargestellt. Der Raum SI_2 enthält mehr als zweieinhalb mal so viele Zustände wie SI_1. Dennoch ist der DA-Wert für SI_2 deutlich niedriger als der für SI_1. Der Raum SI_2 kann also gemäß dem DA-Kriterium als der Raum mit der besseren Passung auf die Daten angesehen werden.

3.3.4 Diskussion des Verfahrens

Mit Hilfe des vorgestellten Verfahrens können, ebenso wie mit mit verwandten Verfahren, prinzipiell beliebig große Itemmengen strukturiert werden. Im

[11] D. h. $M_1 \sqsubseteq_M M_2 \Leftrightarrow M_1 = M_2$ für zwei Motive M_1, M_2.
[12] D. h. $M_1 \sqsubseteq_M M_2 \Leftrightarrow M_1 = M_2 \vee M_1 = \mathrm{F} \wedge M_2 \in \{\mathrm{S,C,T,G,E}\}$ für zwei Motive M_1, M_2.

Tabelle 3. Ergebnisse motivbasierte Aufgabenstrukturen (Sequenzeninklusion)

Struktur	# Zustände	Distanzen								$\overline{d}$	DA
		0	1	2	3	4	5	6	7		
SI_1	368	43	31	14	3	1	–	–	–	0.78	0.21
SI_2	1025	53	31	6	2	–	–	–	–	0.53	0.17

Gegensatz zur Expertenbefragung ist eine Erweiterung der Itemmenge unproblematisch, da jede neue durch Komponenten (in diesem Falle sind es die Motive) kodierte Aufgabe mit Hilfe des Ordnungsprinzips in eine bereits bestehende Struktur eingefügt werden kann.

Strukturen, die mit einem aufgabenanalytischen Verfahren erzeugt wurden sind explizit inhaltlich begründet. Die inhaltliche Fundierung der Strukturen basiert im allgemeinen auf domainspezifischem Wissen und/oder auf theoretischen Ergebnissen zur kognitiven Verarbeitung der Aufgaben. Das domainspezifische Wissen muß üblicherweise mit der Unterstützung von Experten gesammelt und strukturiert werden. Zu diesem Zweck ist auch der Einsatz einer Befragungsprozedur wie oben beschrieben denkbar. Die kognitive Verarbeitung der Aufgaben bzw. der Ablauf des Problemlöseprozesses kann in vielen Fällen mit Hilfe psychologischer Modelle (z. B. Prozeßmodelle) vorhergesagt und erklärt werden. Andererseits bieten die aufgabenanalytischen Verfahren die Möglichkeit, psychologische Modelle und Theorien empirisch zu überprüfen.

Es ist jedoch klar, daß aus der Aufgabenanalyse bzw. der Analyse eines Wissensgebietes im allgemeinen nur eine unvollständige Menge von relevanten Aufgabenkomponenten resultiert. Außerdem werden stets allgemeine Regeln formuliert, deren Gültigkeit für alle konstruierbaren Aufgaben eines Wissensbereichs angenommen wird. Im Falle der verwendeten Schachaufgaben kann man sich jedoch leicht vorstellen, daß außer den Motiven noch weitere Faktoren existieren, die zur Aufgabenschwierigkeit beitragen. Diese Faktoren, die insbesondere in der Formulierung einzelner Aufgaben begründet sein können, sind durch die allgemeinen Strukturierungsprinzipien nicht abzudecken.

Weiterhin stellt der Bereich der elementaren Schachprobleme ein inhaltlich relativ klar aufgebautes Wissensgebiet dar. Es gibt jedoch viele Wissensdomänen, in denen sich die Suche nach Aufgabenkomponenten und inhaltlich fundierten Beziehungen zwischen Aufgaben weit schwieriger gestaltet. Für solcher Bereiche können durch die Intuition und Erfahrung eines Experten möglicherweise eher valide Wissensstrukturen gewonnen werden.

Das hier besprochene Verfahren der Sequenzeninklusion führt ausschließlich zu quasi-ordinalen Wissensräumen. In Bezug auf die untersuchte Wissensdomäne liegt die Etablierung von Surmise-Relationen jedoch nahe. Somit muß diese Ei-

genschaft des Verfahrens in diesem Falle nicht als Mangel angesehen werden, da bereits Verfahren existieren, die — sofern es die Domäne erfordert — eine Ableitung von Surmise-Funktionen aus komponentenbasierten Aufgaben ermöglichen (vgl. Held, 1993; Korossy, 1993).

4 Vergleich der Ergebnisse

Zum Vergleich der Ergebnisse werden die DA-Werte, die durchschnittlichen Distanzen der einzelnen Wissensräume zu den Daten, die Kardinalität der Räume, die maximal beobachtete Distanz zu einem Antwortvektor und die Anzahl der Zustände, die ausgewählte Paare von Räumen gemeinsam haben, herangezogen. Tabelle 4 gibt einen Überblick über die Kennwerte Kardinalität, DA und durchschnittliche und maximale Distanz.

Tabelle 4. Vergleich der Ergebnisse

Struktur	# Zustände	$\overline{d}$	DA	d_{max}	Struktur	# Zustände	$\overline{d}$	DA	d_{max}
DAT_1	57	0.84	0.18	4	*EXP2.1el*	15	1.77	0.33	5
DAT_2	496	0.27	0.07	3	*EXP2.end*	15	1.77	0.33	5
DAT_3	232	0.52	0.13	3	*EXP2.su*	36	1.19	0.32	4
DAT_4	121	0.66	0.15	3					
DAT_2^*	496	0.54	0.16	3	*EXP3.1el*	11	2.24	0.41	7
					EXP3.end	10	2.24	0.41	7
SI_1	368	0.78	0.21	4	*EXP3.su*	11	2.25	0.41	6
SI_2	1025	0.53	0.17	3					
					∪.1el	38	1.45	0.30	5
EXP1.1el	23	1.79	0.35	5	∪.end	34	1.51	0.31	5
EXP1.end	20	1.86	0.36	5	∪.su	62	1.04	0.23	4
EXP1.su	30	1.62	0.33	5					
					∪∪.1el	47	1.40	0.30	5
					∪∪.end	42	1.49	0.29	5
					∪∪.su	79	1.03	0.22	4

* Ergebnis für den halben Datensatz (siehe Abschnitt 3.1.2).

Generell führen Datenanalyse und Sequenzeninklusion zu Ergebnissen, die aufgrund der DA-Werte und der durchschnittlichen Distanzen als besser als die Ergebnisse der Experten bezeichnet werden können. Die beste Expertenstruktur EXP2.su besitzt einen DA-Wert von 0.23, während der minimale DA-Wert bei der Sequenzeninklusion 0.17 (SI_2) und bei der Datenanalyse 0.07 (DAT_2) beträgt.

Die Gegenüberstellung der Verfahren wirft zwangsläufig die Frage auf, welcher der Kennwerte eines Raumes als der bedeutsamste angesehen werden soll.

Die Verwendbarkeit eines Wissensraumes innerhalb einer adaptiven diagnostischen Prozedur wie sie von Falmagne & Doignon (1988a) und Falmagne & Doignon (1988b) vorgeschlagen wird, hängt zum einen davon ab, wie gut dieser Raum die Menge der in der Realität vorkommenden Wissenzustände erfaßt (d. h. von $\overline{d}$), und zum anderen davon, wie stark die Anzahl der zu stellenden Fragen durch die Verwendung des Raumes reduziert wird.[13] Die Frage nach dem optimalen Kennwert kann somit zumindest in dieser Untersuchung nicht beantwortet werden.

Ein wichtiger Punkt beim Vergleich der Ergebnisse ist die Betrachtung der Schnittmengen der einzelnen Räume. In Tabelle 5 wird ein vollständiger Überblick über die Mächtigkeit der Schnittmengen aller Paare von Räumen gegeben. Wie bereits in Abschnitt 3.2.3 gezeigt wurde, gibt es kaum eine Übereinstimmung zwischen den Ergebnissen der Experten. Die vereinigten, bzw. vereinigungsabgeschlossenen gelegten Strukturen der Experten (∪.su und ∪∪.su) sind jedoch weitgehend in den theoretischen Strukturen SI_1 und SI_2 enthalten. Bemerkenswert ist in diesem Zusammenhang auch die gelegte Struktur des Experten EXP2 (EXP2.su), auf die im folgenden Abschnitt noch genauer eingegangen wird. Die gelegte Struktur des Experten EXP1 (EXP1.su) ist vollständig in der Struktur SI_2 enthalten. Jedoch sind nur 16 der 30 Zustände in EXP1.su auch Elemente von SI_1. Dies legt die Vermutung nahe, daß Experte EXP1 die zusätzlichen Annahmen, die zu SI_1 führen bei seinen Entscheidungen nicht berücksichtigt hat.

[13] Ein Extrembeispiel für einen Wissensraum ist die Potenzmenge der Aufgabenmenge. Dieser Raum deckt alle vorkommenden Zustände ab — er ist in dieser Hinsicht also optimal. Der Raum ist für eine ökonomische Diagnose aber unbrauchbar, da stets alle Aufgaben vorgegeben werden müssen, um einen Wissenszustand diagnostizieren zu können.

Tabelle 5. Vollständiger Überblick über die Mächtigkeit der Schnittmengen aller in der Untersuchung betrachteten Räume.

Raum	DAT_1 (57)	DAT_2 (496)	DAT_3 (232)	DAT_4 (121)	EXP1.1el (23)	EXP1.end (20)	EXP1.su (30)	EXP2.1el (15)	EXP2.end (15)	EXP2.su (36)	EXP3.1el (11)	EXP3.end (10)	EXP3.su (11)	U.1el (38)	U.end (34)	U.su (62)	UU.1el (47)	UU.end (42)	UU.su (79)	SI_1 (368)	SI_2 (1025)
DAT_1 (57)	–	57	53	45	9	9	10	13	13	19	5	5	5	17	17	24	18	18	29	44	50
DAT_2 (496)	–	–	158	121	13	13	17	13	13	26	5	5	7	21	21	38	26	26	49	97	160
DAT_3 (232)	–	–	–	54	13	13	16	14	14	25	5	5	7	22	22	37	28	28	43	115	146
DAT_4 (121)	–	–	–	–	10	10	10	10	10	21	3	3	5	15	15	26	16	16	33	57	74
EXP1.1el (23)	–	–	–	–	–	20	15	6	6	3	5	5	9	23	20	16	23	20	17	15	22
EXP1.end (20)	–	–	–	–	–	–	15	6	6	3	5	5	9	20	20	16	20	20	17	15	19
EXP1.su (30)	–	–	–	–	–	–	–	5	5	6	5	5	9	15	15	30	15	15	30	16	30
EXP2.1el (15)	–	–	–	–	–	–	–	–	15	11	3	3	4	15	15	13	15	15	15	14	15
EXP2.end (15)	–	–	–	–	–	–	–	–	–	11	3	3	4	15	15	13	15	15	15	14	15
EXP2.su (36)	–	–	–	–	–	–	–	–	–	–	3	3	4	15	15	13	15	15	15	14	15
EXP3.1el (11)	–	–	–	–	–	–	–	–	–	–	–	10	3	11	10	5	11	10	5	4	6
EXP3.end (10)	–	–	–	–	–	–	–	–	–	–	–	–	3	10	10	5	10	10	5	4	6
EXP3.su (11)	–	–	–	–	–	–	–	–	–	–	–	–	–	9	9	11	10	10	11	9	11
U.1el (38)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	34	24	38	34	26	23	32
U.end (34)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	24	34	34	26	23	29
U.su (62)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	62	25	62	43	60
UU.1el (47)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	42	27	27	37
UU.end (42)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	27	27	33
UU.su (79)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	57	76
SI_1 (368)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	368
SI_2 (1025)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

5 Diskussion einer Integration der Verfahren

Jedes der drei vorgestellten Verfahren weist eine Reihe spezifischer Vor- und Nachteile auf. Unser Ziel ist es nun, darzustellen, wie durch eine kombinierte Anwendung der Verfahren die Vorteile genutzt werden können, ohne die Nachteile in Kauf nehmen zu müssen.

Die Datenanalyse, die hier in Form der Item-Tree-Analyse vorgestellt wurde, sollte nur auf große heterogene Datensätze angewandt werden. Außerdem sind die ITA-Strukturen nicht inhaltlich bzw. theoretisch fundiert. Dieser Anspruch an eine valide Wissensstruktur soll hier vorausgesetzt werden.

Jedoch kann die ITA auch auf kleine Datensätze wie den hier verwendeten angewandt werden, um die Struktur eines Datensatzes zu visualisieren und daraus Hinweise auf mögliche Ergänzungen der theoretischen Annahmen zu gewinnen. Wir wollen dies an einem konkreten Beispiel darstellen. Die mit der ITA erzeugte Struktur DAT_1 (siehe Abbildung 3) legt die Annahme nahe, daß nicht nur Motiv F (Gabel) am leichtesten zu erkennen ist (wie für Struktur SI_1 angenommen), sondern daß darüberhinaus Motiv G (Hinlenkung) am schwierigsten zu erkennen ist. Diese Vermutung ergibt sich aus der Tatsache, daß die Aufgaben, die Motiv G enthalten in der Hierarchie der Aufgaben relativ hoch angeordnet sind. Modifiziert man nun die Relation $\sqsubseteq_M$ (siehe Abschnitt 3.3.1) gemäß dieser Annahme, so ergibt sich eine Struktur SI_3 mit den in Tabelle 6 gezeigten Kennwerten.

Tabelle 6. Ergebnisse für Struktur SI_3 (Annahme, daß Motiv G am schwersten zu erkennen ist und Motiv F am leichtesten).

Struktur	# Zustände	0	1	2	3	4	5	6	7	$\overline{d}$	DA
SI_3	99	37	34	16	4	1	–	–	–	0.89	0.20

Wir sehen, daß sich in SI_3 die Anzahl der Zustände im Vergleich zu SI_1 (368 Zustände) auf weit weniger als ein Drittel verringert hat. Dennoch enthält SI_3 nur 6 Zusände weniger, die Distanz 0 zu den Daten aufweisen. Der DA-Wert von SI_3 (0.20) ist geringfügig niedriger als der von SI_1 (0.21). In Hinblick auf die Verwendung innerhalb einer wissensdiagnostischen Prozedur dürfte SI_3 die „ökonomischere“[14] Struktur darstellen. Die Struktur SI_3 ist eine Teilmenge der Struktur SI_1.

Eine weitere Reduktion der Wissensstruktur ist mit Hilfe der Ergebnisse der Expertenbefragung möglich. Wir betrachten dazu die gelegte Struktur des Ex-

[14] Mit dem Begriff „ökonomisch“ ist in diesem Zusammenhang gemeint, daß die durchschnittliche Anzahl von Fragen, die bei einer Verwendung der Struktur in einem adaptiven Diagnoseverfahren (vgl. Falmagne & Doignon, 1988a; Falmagne & Doignon, 1988b) gestellt werden müßten relativ gering wäre.

perten EXP2 (EXP2.su, vgl. Abb. 5). Genau 31 der 36 Zustände in EXP2.su sind auch in der aufgabenanalytischen Struktur SI_1 enthalten (bzw. 34 in SI_2). Der DA-Wert von EXP2.su (0.23) liegt nur knapp über dem DA-Wert von SI_1 (0.21), die aufgetretenen Distanzen sind jedoch höher. Die Struktur EXP2.su stellt im Vergleich zu den anderen Räumen, die nicht per Datenanalyse abgeleitet wurden die „ökonomischste" Lösung dar. Betrachtet man die Übereinstimmung der Struktur mit den Daten, so ist zu beachten, daß der Datensatz 69 unterschiedliche Antwortvektoren enthält, EXP2.su jedoch nur eine Kardinalität von 36 besitzt.

In den beiden Beispielen wurde bereits die Grundtendenz einer möglichen Integration der drei Verfahren zur Ableitung von Wissensräumen erkennbar. Datenanalyse und Expertenbefragung können primär zur *Reduktion* komponentenbasierter Strukturen dienen. Wie wir bereits in Abschnitt 3.3.4 erwähnten, wird es in den wenigsten Fällen möglich sein, theoretisch fundierte Wissensstrukturen abzuleiten, in denen die Gesamtheit der Komponenten, die die Aufgabenschwierigkeit beeinflussen, berücksichigt ist. Datenanalyse und Expertenbefragung können dazu dienen, solche Komponenten zu identifizieren oder festzustellen, daß die theoretisch angenommenen Komponenten revidiert werden müssen.

Expertenbefragung und Aufgabenanalyse sind die Verfahren der Wahl, wenn eine neue Wissensstruktur entwickelt werden soll. Die Konsultation von Experten ist im allgemeinen notwendig, wenn Aufgabenkomponenten zusammengestellt und Ordnungsprinzipien kreiert werden. Strukturelle Eigenschaften einer Komponentenmenge können mit dem hier besprochenen Befragungsalgorithmus identifiziert werden.

Wenn eine erste Version der Wissensstruktur existiert, kann es sinnvoll sein (z. B. aus Kostengründen), noch vor der Datenerhebung mit den konstruierten Aufgaben eine Expertenbefragung durchzuführen. Ergeben sich starke Abweichungen zwischen Expertenstruktur und aufgabenanalytischer Struktur, so sollte das bestehende Komponenten-Konzept nochmals überprüft werden.

Nach der Datenerhebung können datenanalytische Verfahren zum Einsatz kommen. Wie gezeigt, kann ein solches Verfahren als nützliches Werkzeug zur Darstellung der Daten dienen. Falls sich Konsequenzen aus der Datenanalyse ergeben, so müssen diese in theoretischer Hinsicht überprüft werden und im Falle einer theoretischen Begründbarkeit (auch mit Hilfe von Experten) in die Wissensstruktur einbezogen werden. Anschließend kann eine weitere Datenerhebung begonnen werden.

Das vorgeschlagene, aus verschiedenen Methoden kombinierte, Verfahren stellt also einen zirkulären Prozeß dar, dessen Ziel die Konstruktion von sowohl empirisch validen, als auch möglichst ökonomischen Wissensstrukturen ist.

Literatur

Airasian, P. W. & Bart, W. M. (1973). Ordering theory: a new and useful measurement model. *Educational Technology*, **May**, 56–60.

Albert, D. (1989). *Knowledge assessment: choice heuristics as strategies for construting questions and problems*, Paper read at the 20th european mathematical psychology group meeting, Nijmegen, September 1989.

Albert, D. & Held, T. (1994). Establishing knowledge spaces by systematical problem construction. In D. Albert (Ed.), *Knowledge Structures* (pp. 81–115). Berlin, Heidelberg: Springer.

Albert, D., Schrepp, M. & Held, T. (1994). Construction of knowledge spaces for problem solving in chess. In G. H. Fischer & D. Laming (Eds.), *Contributions to Mathematical Psychology, Psychometrics, and Methodology* (pp. 123–135). Berlin, Heidelberg, New York: Springer.

Bart, W. M. & Airasian, P. W. (1974). Determination of the ordering among seven Piagetian tasks by an ordering-theoretic method. *Journal of Educational Psychology*, **66**(2), 277–284.

Bart, W. M. & Krus, D. J. (1973). An ordering-theoretic method to determine hierarchies among items. *Educational and Psychological Measurement*, **33**, 291–300.

Birkhoff, G. (1937). Rings of sets. *Duke Mathematical Journal*, **3**, 443–454.

Bortz, J. (1993). *Statistik für Sozialwissenschaftler* (4th ed.). Berlin, Heidelberg, New York: Springer.

Doignon, J.-P. & Falmagne, J.-C. (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, **23**, 175–196.

Dowling, C. E. (1993a). Applying the basis of a knowledge space for controlling the questioning of an expert. *Journal of Mathematical Psychology*, **37**, 21–48.

Dowling, C. E. (1993b). On the irredundant generation of knowledge spaces. *Journal of Mathematical Psychology*, **37**, 49–62.

Einhorn, H. J. (1986). Expert judgement: Some necessary conditions and an example. In H. R. Arkes & K. R. Hammond (Eds.), *Judgement and decision making* (pp. 480–491). Cambridge: Cambridge University Press.

Falmagne, J.-C. (1989a). A latent trait theory via a stochastic learning theory for a knowledge space. *Psychometrica*, **54**(2), 283–303.

Falmagne, J.-C. (1989b). Probabilistic knowledge spaces: A review. In F. Roberts (Ed.), *Applications of combinatorics and graph theory to the biological and social sciences*, Vol. 17. New York: Springer.

Falmagne, J.-C. & Doignon, J.-P. (1988a). A class of stochastic procedures for the assessment of knowledge. *British Journal of Mathematical and Statistical Psychology*, **41**, 1–23.

Falmagne, J.-C. & Doignon, J.-P. (1988b). A markovian procedure for assessing the state of a system. *Journal of Mathematical Psychology*, **32**(3), 232–258.

Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P. & Johannesen, L. (1990). Introduction to knowledge spaces: how to build, test and search them. *Psychological Review*, **97**, 201–224.

Gaeth, G. J. & Shanteau, J. (1986). Reducing the influence of irrelevant information on experienced decision makers. In H. R. Arkes & K. R. Hammond (Eds.), *Judgement and decision making* (pp. 449–465). Cambridge: Cambridge University Press.

Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Held, T. (1993). *Establishment and empirical validation of problem structures based on domain specific skills and textual properties — A contribution to the 'Theory of Knowledge Spaces'*, Dissertation, Universität Heidelberg.

Held, T. & Korossy, K. (in prep.). Item tree analysis vs. an expert's judgements.

Koppen, M. (1993). Extracting human expertise for constructing knowledge spaces: an algorithm. *Journal of Mathematical Psychology*, **37**, 1–20.

Koppen, M. & Doignon, J.-P. (1990). How to build a knowledge space by querying an expert. *Journal of Mathematical Psychology*, **34**, 311–331.

Korossy, K. (1993). *Modellierung von Wissen als Kompetenz und Performanz: Eine Erweiterung der Wissensstruktur-Theorie von Doignon & Falmagne*, Dissertation, Universität Heidelberg.

Lukas, J. & Albert, D. (1993). Knowledge assessment based on skill assignment and psychological task analysis. In G. Strube & K. F. Wender (Eds.), *The Cognitive Psychology of Knowledge* (pp. 139–159). Amsterdam: Elsevier.

Müller, C. E. (1989). A procedure for facilitating an expert's judgements on a set of rules. In E. E. Roskam (Ed.), *Mathematical Psychology in Progress* (pp. 157–170). Heidelberg, New York: Springer.

Schrepp, M. (1993). *Über die Beziehung zwischen kognitiven Prozessen und Wissensräumen beim Problemlösen*, Dissertation, Universität Heidelberg.

Schrepp, M. & Held, T. (1994). A simulation study concerning the effect of errors on the establishment of knowledge spaces by querying experts.

Schrepp, M., Held, T. & Albert, D. (1994). Component based construction of surmise relations for chess problems.

Shanteau, J. (1988). Psychological characteristics and strategies of expert decision makers. *Acta Psychologica*, **68**, 203–215.

Van Buggenhaut, J. & Degreef, E. (1987). On dichotomisation methods in boolean analysis of questionaires. In E. E. Roskam & R. Suck (Eds.), *Progress in Mathematical Psychology* (pp. 447–453). Amsterdam: Elsevier.

Van Leeuwe, J. F. J. (1974). Item tree analysis. *Nederlands Tijdschrift voor de Psychologie*, **29**, 475–484.

Villano, M. (1991). *Computerized Knowledge Assessment: Building the Knowledge Structure and Calibrating the Assessment Routine*, PhD thesis, New York University.

Anhang

Tabelle 7. Aufstellung der in der Untersuchung verwendeten Schachaufgaben

Nummer	Typ	Stellung	Lösung	Alternative
1	(S)	*Weiß*: Ka1 Df5 *Schwarz*: Kh8 Bg6	1. D beliebig außer Dg6:,Dh5	
2	(G, S)	*Weiß*: Kh1 Lf2 *Schwarz*: Kc7 De5	1. Lg3 Dg3: patt	
3	(E, G, S)	*Weiß*: Kh1 Lf2 Sc3 *Schwarz*: Kc7 Df4 Bc6	1. Sd5+ cd5: 2. Lg3 Dg3: patt	
4	(E, E, G, S)	*Weiß*: Ka1 Lc2 Sf3,d7 *Schwarz*: Kf7 Dc4 Bf6,d6	1. Se5+ de5:/fe5: 2. Se5:+ fe5:/de5: 3. Lb3 Db3: patt	
5	(C, S)	*Weiß*: Ka3 Bb2 *Schwarz*: Kc5 Th2 Ba4,b5,c4	1. b4+ ab3:ep/cb3:ep patt	
6	(G, C, S)	*Weiß*: Kh3 Lg4 Bg2 *Schwarz*: Kg6 Tc2 Bh4,g5,f4	1. Lf5+ Kf5: 2. g4+ hg3:ep/fg3:ep patt	
7	(T, S)	*Weiß*: Kh5 Bf7 *Schwarz*: Kh7	1. f8T Gewinn	
8	(G, E, S)	*Weiß*: Ka1 Lc2 Se2 *Schwarz*: Ke6 Dc4 Be5	1. Lb3 Db3: 2. Sd4+ ed4: patt	
9	(F)	*Weiß*: Kb2 Sf3 *Schwarz*: Kf7 Dc6	1. Se5+ Remis	
10	(G, F)	*Weiß*: Kh1 Le3 Sc3 *Schwarz*: Kc7 De5	1. Lf4 Df4: 2. Sd5+ Remis	
11	(G, F, F)	*Weiß*: Kg1 Lc6 Se4,g4 *Schwarz*: Kg8 De6 Bg7	1. Ld5 Dd5: 2. Sf6+ gf6: 3. Sf6:+ Remis	
12	(G, G, F, F)	*Weiß*: Kb1 Lg3 Sh4,e7 Bf7,c2 *Schwarz*: Kh8 Df6 Lg7 Bh7	1. f8D(T)+ Lf8: 2. Le5 De5: 3. Sg6+ hg6: 4. Sg6:+ Remis	1. Sg6+ hg6: 2. f8D+ Lf8: 3. Le5 De5: 4. Sg6:+ Remis
13	(G, G, F)	*Weiß*: Kb1 Lg3 Sh4 Bf7,e2 *Schwarz*: Kh8 Df6 Lg7	1. f8D(T)+ Lf8: 2. Le5 De5: 3. Sg6+ Remis	
14	(F, F)	*Weiß*: Kb2 Se3,c3 *Schwarz*: Kc7 Df4 Be6	1. Sd5+ ed5: 2. Sd5:+ Remis	
15	(T, F)	*Weiß*: Kc1 Bf7,b2 *Schwarz*: Kh7 Dd7	1. f8S+ Gewinn	
16	(T, F, F)	*Weiß*: Kb1 Se4 Bf7 *Schwarz*: Kh7 Dd7 Lg7	1. f8S+ Lf8: 2. Sf6+ Remis	1. Sf6+ Lf6: 2. f8S+ Remis

Instruktion 1

Diese Untersuchung findet statt im Rahmen des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Forschungsprojekts „Wissensstrukturen". Wir beschäftigen uns in dieser Untersuchung mit der Repräsentation von Wissen am Beispiel von Schachaufgaben. Ihre Aufgabe wird es sein zu beurteilen, wie schwierig Schachaufgaben, die wir ihnen vorlegen werden, für Schachspieler, insbesondere Anfänger, sind. Wir hoffen aus Ihren Urteilen Rückschlüsse auf die zum Lösen solcher Aufgaben notwendigen kognitiven Fähigkeiten ziehen zu können.

Instruktion 2

Sie erhalten nun die von uns verwendeten Aufgaben als Diagramme ausgehändigt. Schauen Sie sich diese bitte gründlich auf Analysieren Sie die Lösung und mögliche Nebenlösungen jeder Aufgabe. Bauen Sie dazu, falls notwendig, die Aufgaben auf dem bereitgestellten Brett auf.

Wenn Sie glauben sich mit den Aufgaben ausreichend vertraut gemacht zu haben, so verständigen Sie bitte den Versuchsleiter. Für die folgenden Phasen der Untersuchung ist es wichtig, daß Sie die Aufgaben möglichst gut kennen. Lassen Sie sich deshalb bitte Zeit beim Analysieren der Aufgaben. Dies wird auch die folgenden Teile der Untersuchung wesentlich beschleunigen.

Abbildung 7. Expertenbefragung: Instruktionstexte 1 und 2

Instruktion 3

Wir werden ihnen im folgenden Fragen zur Schwierigkeit der Ihnen vorgelegten Aufgaben stellen. Diese Fragen werden von der folgenden Form sein :

1) Eine Person kann Aufgabe a nicht lösen. Erwarten Sie dann, daß sie auch Aufgabe b nicht lösen kann ?

oder

2) Eine Person kann die Aufgaben a,b,c nicht lösen. Erwarten Sie dann, daß Sie auch die Aufgabe d nicht lösen kann ?

Sie sollen jede dieser Fragen entweder mit „Ja"oder „Nein"beantworten. Wir demonstrieren Ihnen wie Sie dabei vorgehen sollten am Beispiel der zweiten Frage.

Stellen Sie sich eine Person vor, z.B. einen Schachanfänger oder Hobbyspieler, die die Aufgaben a,b,c nicht lösen kann. D.h. nicht genug Wissen bzw. Übung besitzt diese Aufgaben korrekt lösen zu können. Versuchen Sie dann abzuschätzen, ob diese Person genügend Wissen bzw. Übung besitzen kann, um d lösen zu können. Sind Sie der Meinung, daß dies möglich ist, so antworten Sie bitte mit „Nein". Sind Sie dagegen der Meinung, daß dies unmöglich ist, so antworten Sie bitte mit „Ja".

Um Sie mit dieser Art von Fragen vertraut zu machen, geben wir ihnen zwei konkrete Beispiele. Bearbeiten Sie diese bitte gründlich. Versuchen Sie Ihre Antworten zu begründen.

1) Eine Person kann Aufgabe 17 nicht lösen. Erwarten Sie dann, daß Sie auch Aufgabe 18 nicht lösen kann?

2) Eine Person kann die Aufgaben 19 und 20 nicht lösen. Erwarten Sie dann, daß Sie auch die Aufgabe 17 nicht lösen kann?

Instruktion 4

Wir werden Ihnen nun eine Reihe von Fragen des beschriebenen Typs vorlegen. Diese beziehen sich nur auf die Aufgaben 1 bis 16. Die Aufgaben 17 bis 20 dienten nur zur Übung.

Ihre Antworten werden vom Versuchsleiter in den Rechner eingegeben. Dieser ermittelt dann die Ihnen als nächstes gestellte Frage.

Sie brauchen bei Ihren Antworten nicht auf logische Konsistenz zu achten. Der Rechner stellt Ihnen keine Fragen mehr, deren Antworten aus Ihren vorhergehenden Antworten logisch folgt.

Da der Rechner aus jeder Ihrer Antworten eine große Zahl von logischen Schlußfolgerungen in Bezug auf die nächsten zu stellenden Fragen zieht, sollten Sie Ihre Antworten genau überdenken. Eine „Ja" Antwort kann unter Umständen einen völlig anderen Fortgang der Befragungsprozedur bewirken als eine „Nein" Antwort.

Viel Spaß !

Abbildung 8. Expertenbefragung: Instruktionstexte 3 und 4

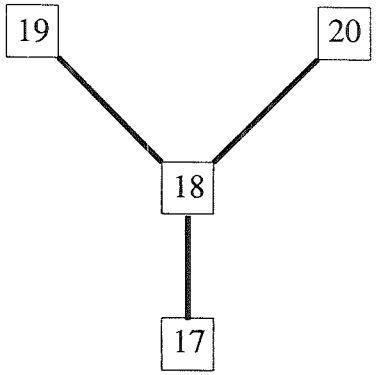
Instruktion 5

Wir kommen nun zum letzten Abschnitt unserer Untersuchung. Ihre Aufgabe besteht nun darin aus den Schachaufgaben eine Aufgabenstruktur zu legen. Wir bitten Sie daher die Aufgabenkarten auf dem Tisch so anzuordnen, daß diese Anordnung Ihren Überzeugungen bezüglich der Folgerungsbeziehungen zwischen den Aufgaben entspricht. Gehen Sie dabei bitte nach den folgenden Regeln vor:

1. Wenn sie erwarten, daß eine Person, die die Aufgabe a löst auch Aufgabe b löst, so ordnen Sie Aufgabe a *über* Aufgabe b an und verbinden beide Aufgaben mit einem Papierstreifen.
   (Beachten Sie bitte, daß hier eine andere Fragerichtung vorliegt als bei der vorherigen Befragung.)
2. Wenn Sie erwarten, daß eine Person, die Aufgabe a nicht löst auch Aufgabe b nicht löst, so ordnen Sie bitte Aufgabe a *unter* Aufgabe b an und verbinden beide Aufgaben mit einem Papierstreifen.
3. Sind drei Aufgaben a, b und c durch eine derartige Folgerungsbeziehung miteinander verknüpft, so legen sie bitte nur die jeweils direkte Beziehung. Wenn z. B. aus der Lösung von Aufgabe a die Lösung von Aufgabe b und von Aufgabe c folgt und aus der Lösung von Aufgabe b die Lösung von Aufgabe c folgt, dann ordnen sie bitte Aufgabe a über Aufgabe b und diese wiederum über Aufgabe c an. Sie verbinden Aufgabe a mit Aufgabe b und Aufgabe b mit Aufgabe c. Die Verbindung zwischen Aufgabe a und Aufgabe c wird nicht gesondert gelegt (diese Eigenschaft wird noch an einem Beispiel erläutert werden).
4. Falls Sie der Auffassung sind, daß eine Folgerungsbeziehung für zwei Aufgaben in beiden Richtungen gilt, wenn also z. B. aus dem Lösen von Aufgabe a das Lösen von Aufgabe b folgt und aus dem Lösen von Aufgabe b das Lösen von Aufgabe a folgt, so ordnen Sie die beiden Aufgaben bitte *nebeneinander* an, und zwar so, daß die Aufgabenkarten einander berühren.

Hinweis: Es ist möglich, daß Aufgabenpaare nicht miteinander verknüpft sind. Dies gilt, wenn aus der Lösung von Aufgabe a nichts für die Lösung von Aufgabe b folgt, und aus der Lösung von Aufgabe b nichts für die Lösung von Aufgabe a folgt.

Lassen sie uns nun noch das Beschriebene kurz am Beispiel der vier Übungsaufgaben näher erläutern. Die folgende Grafik stellt eine mögliche Struktur für die vier Beispielaufgaben dar:



Aus der Lösung von Aufgabe 18 folgt die Lösung von Aufgabe 17 und daher ist Aufgabe 18 über Aufgabe 17 angeordnet und mit ihr verbunden (vergleiche Regel 1). Aus dem Nicht-Lösen von Aufgabe 18 folgt das Nicht-Lösen von Aufgabe 20 und daher ist Aufgabe 18 unter Aufgabe 20 angeordnet und mit ihr verbunden (vergleiche Regel 2). Aus dem Lösen von Aufgabe 19 folgt das Lösen von Aufgabe 18 und Aufgabe 17. Die Verbindung zwischen Aufgabe 19 und Aufgabe 17 wird jedoch nicht gesondert dargestellt (vergleiche Regel 3). Der Experte hat weiterhin die Behauptung aufgestellt, daß die Aufgaben 19 und 20 in keiner Folgerungsbeziehung zueinander stehen. Daher sind die beiden Aufgaben nicht miteinander verknüpft.

Sollten Sie noch Fragen bezüglich des Legens der Aufgabenstruktur haben so wenden Sie sich bitte an den Versuchsleiter.

Abbildung 9. Expertenbefragung: Instruktionstext 5

Tabelle 8. Antwortvektoren aus der experimentellen Untersuchung (Reihenfolge der Aufgaben: (S), (G,S), (E,G,S), (E,E,G,S), (C,S), (G,C,S), (T,S), (G,E,S), (F), (G,F), (G,F,F), (G,G,F,F), (G,G,F), (F,F), (T,F), (T,F,F))

Häufigkeit	Antwortvektor	Häufigkeit	Antwortvektor
8	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1	1 0 0 1 0 0 1 1 1 1 0 0 0 1 1 1
6	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0	1	1 0 0 1 0 0 0 0 1 1 1 0 0 1 0 0
3	1 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0	1	1 0 0 0 1 1 1 1 1 0 0 0 0 0 1 1
3	1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0	1	1 0 0 0 1 1 0 0 1 0 0 0 0 0 1 1
3	0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0	1	1 0 0 0 1 0 1 1 1 0 0 0 0 1 1 0
2	1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1	1	1 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0
2	1 0 0 0 1 0 1 0 1 0 0 0 0 1 1 0	1	1 0 0 0 1 0 1 0 1 0 0 0 0 1 1 1
2	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1	1 0 0 0 1 0 0 0 1 0 0 0 0 1 1 0
2	0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0	1	1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0
2	0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0	1	1 0 0 0 0 0 1 1 1 0 0 0 0 0 1 0
1	1 1 1 1 1 1 1 0 1 1 0 1 0 1 1 0	1	1 0 0 0 0 0 1 0 1 1 1 0 1 1 1 1
1	1 1 1 1 1 0 0 1 1 1 0 0 0 1 1 0	1	1 0 0 0 0 0 1 0 1 1 1 0 0 1 1 1
1	1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1	1	1 0 0 0 0 0 1 0 1 1 0 0 0 1 0 0
1	1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1	1	1 0 0 0 0 0 1 0 1 0 1 0 0 1 1 1
1	1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1	1	1 0 0 0 0 0 1 0 1 0 0 0 0 1 1 1
1	1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1	1	1 0 0 0 0 0 1 0 1 0 0 0 0 1 1 0
1	1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1	1	1 0 0 0 0 0 1 0 1 0 0 0 0 1 0 0
1	1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1	1	1 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0
1	1 1 1 1 0 0 1 0 1 1 1 1 1 1 1 1	1	1 0 0 0 0 0 0 0 1 1 1 1 1 1 0 1 1 1
1	1 1 1 1 0 0 1 0 1 1 0 1 1 1 1 1	1	1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0
1	1 1 1 0 1 1 0 1 1 1 1 0 1 1 1 1	1	1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0
1	1 1 1 0 1 1 0 1 1 1 1 0 0 1 1 1	1	1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0
1	1 1 1 0 0 0 1 1 1 1 1 1 1 1 1 1	1	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
1	1 1 1 0 0 0 1 1 1 1 1 1 1 0 1 1	1	0 0 0 0 1 0 1 0 1 0 0 0 0 1 1 0
1	1 1 1 0 0 0 1 1 1 1 1 1 0 0 1 1	1	0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0
1	1 1 1 0 0 0 1 1 1 1 0 1 0 1 1 0	1	0 0 0 0 1 0 0 0 1 0 0 0 0 1 1 0
1	1 1 0 1 0 0 1 1 1 1 1 1 0 0 1 1	1	0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0
1	1 1 0 0 0 1 1 1 1 1 1 1 1 0 1 0	1	0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0
1	1 1 0 0 0 0 1 1 1 1 0 0 0 1 1 1	1	0 0 0 0 0 0 0 1 0 1 0 0 0 0 1 1 1
1	1 1 0 0 0 0 1 1 1 0 0 0 0 0 1 1 0	1	0 0 0 0 0 0 0 0 1 1 0 0 0 0 1 1 0
1	1 1 0 0 0 0 1 0 1 1 1 0 0 1 1 0	1	0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 1 1
1	1 1 0 0 0 0 0 0 0 1 1 1 0 0 1 0 0	1	0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1
1	1 1 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0	1	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0
1	1 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1	1	0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
1	1 1 0 0 0 0 0 0 0 0 0 1 0 1 0 1		

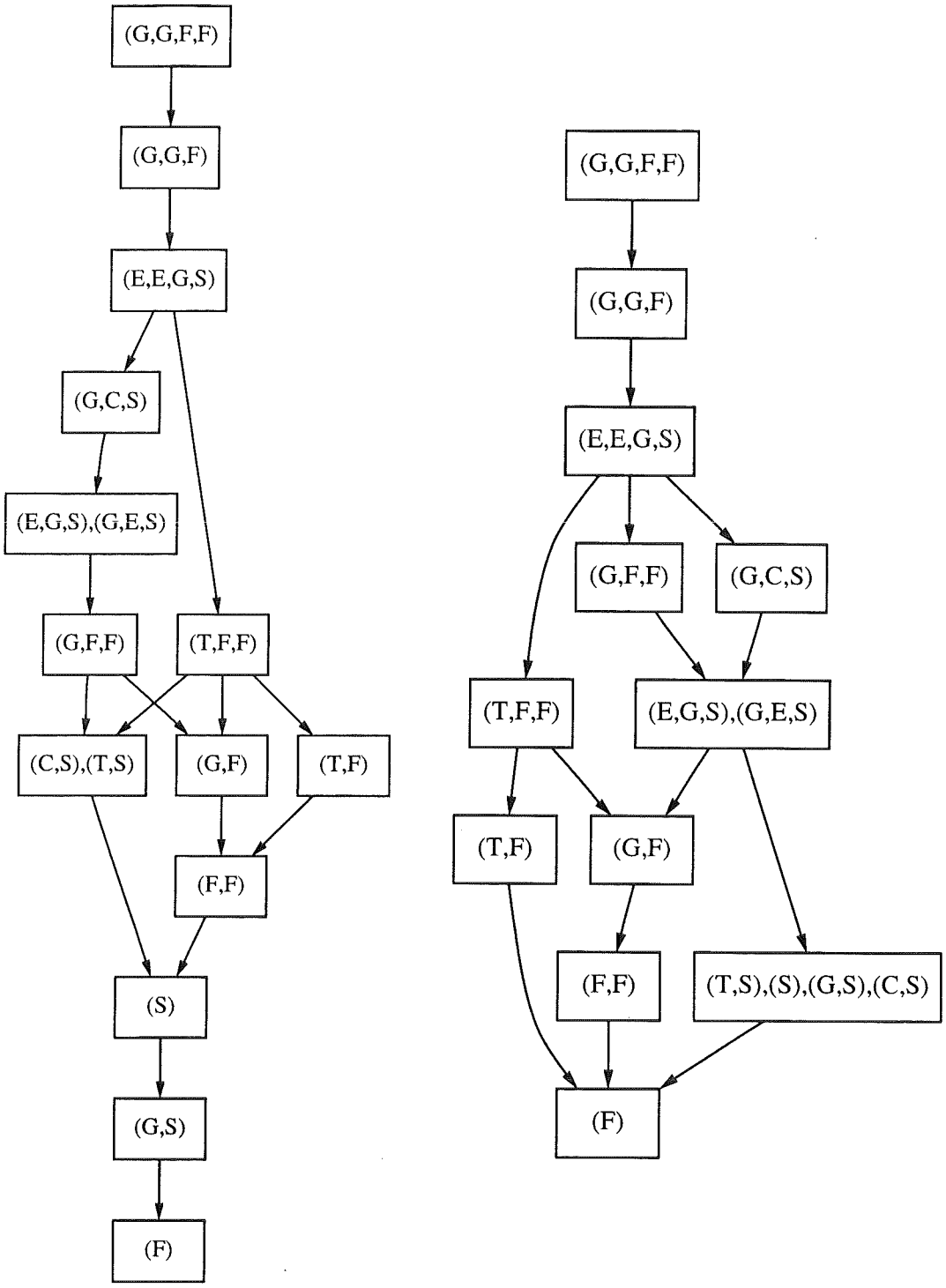


Abbildung 10. Experte EXP1: Resultat der Befragung mit einelementigen Prämissen (links) und gelegte Struktur (rechts)

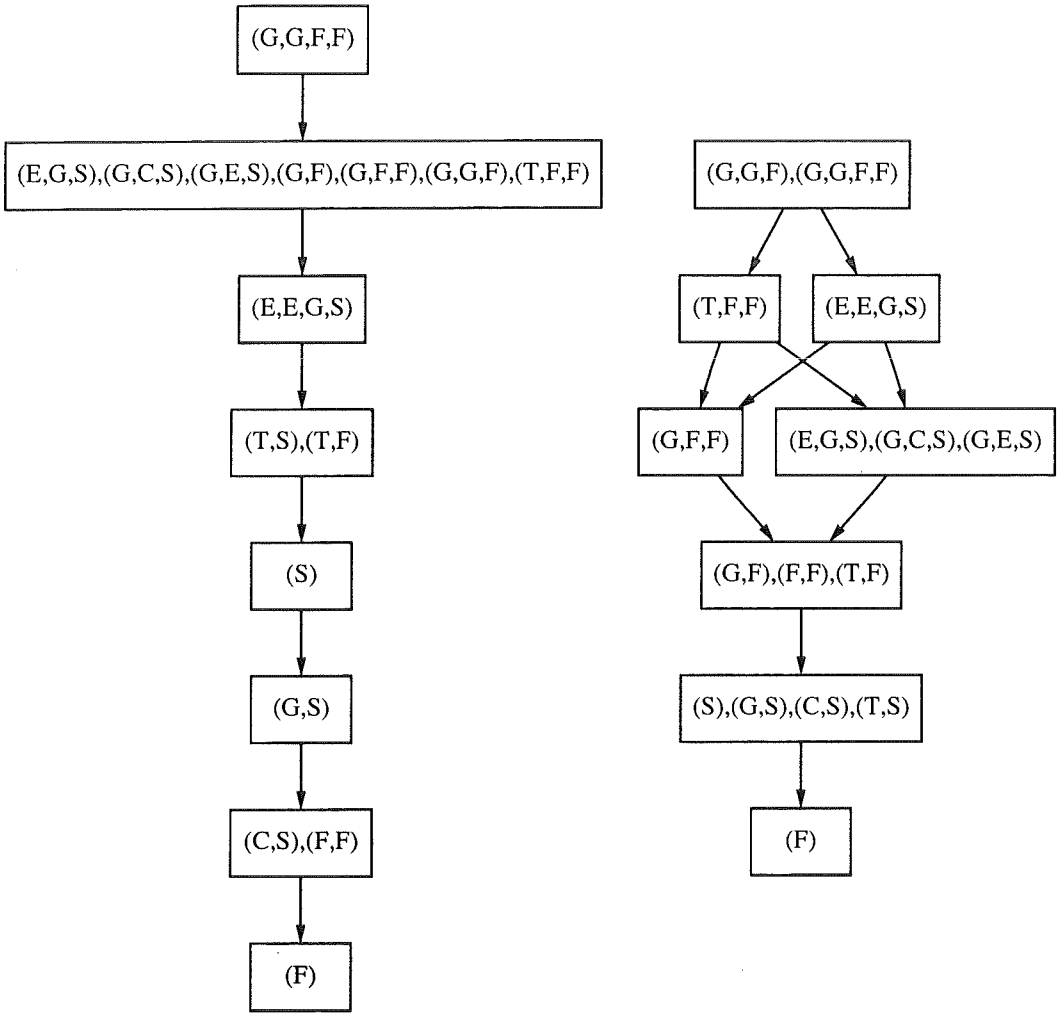


Abbildung 11. Experte EXP3: Resultat der Befragung mit einelementigen Prämissen (links) und gelegte Struktur (rechts)

Bisher erschienene Institutsberichte

Diskussionspapier Nr. 1: **GROEBEN, N.:** Vom behavioralen zum epistemologischen Subjektmodell: Paradigmawechsel in der Psychologie? September 1975

Diskussionspapier Nr. 2: **MÖBUS, C.& SIMONS, H.:** Zur Fairness psychologischer Intelligenztests gegenüber ethnischen und sozialen Gruppen: Kritik klassischer Konzepte. Oktober 1975

Diskussionspapier Nr. 3: **WOTTAWA, H.:** Skalenprobleme bei probabilistischen Meßmodellen. März 1976

Diskussionspapier Nr. 4: **Treiber, B. & Petermann, F.:** Zur Interaktion von Lernermerkmalen und Lehrmethoden: Rekonstruktion und Normierung des ATI-Forschungsprogramms. April 1976

Diskussionspapier Nr. 5: **MÖBUS, C. & WALLASCH, R.:** Zur Erfassung von Hirnschädigungen bei Kindern: Nichtlineare Entscheidungsregeln auf der Basis von Veränderungsmessungen. August 1976

Diskussionspapier Nr. 6: **SCHEELE, B. & GROEBEN, N.:** Voraussetzungs- und zielspezifische Anwendung von Konditionierungs- vs. kognitiven Lerntheorien in der klinischen Praxis. Dezember 1976

Diskussionspapier Nr. 7: **MÖBUS, C.:** Zur Analyse nichtsymmetrischer Ähnlichkeitsurteile: Ein dimensionales Driftmodell, eine Vergleichshypothese, TVERSKY's Kontrastmodell und seine Fokushypothese. Juni 1977

Diskussionspapier Nr. 8: **Simons, H. & Möbus, C.:** Veränderung von Berufschancen durch Intelligenztraining. Juli 1977

Diskussionspapier Nr. 9: **Braunmühl, C. v. & Grimm, H.:** Zur Kommunikationspsychologie: Über Versuche der methodischen Konstitution eines genuin humanwissenschaftlichen Forschungsansatzes zur Entwicklung der Verständigungsfähigkeit. November 1977

Diskussionspapier Nr. 10: **Hofer, M.:** Entwurf einer Heuristik für eine theoretisch geleitete Lehrer- und Erzieherbildung. November 1977

Diskussionspapier Nr. 11: **Scheibler, D. & Schneider, W.:** Probleme und Ergebnisse bei der Evaluation von Clusteranalyse-Verfahren. Juni 1978

Diskussionspapier Nr. 12: **Scheele, B.:** Kognitions- und sprachpsychologische Aspekte der Arzt-Patient-Kommunikation. September 1978

Diskussionspapier Nr. 13: **Treiber, B. & Schneider, W.:** Mehrebenenanalyse sozialstruktureller Bedingungen schulischen Lernens. Oktober 1978

Diskussionspapier Nr. 14: **Ahrens, H.-J. & Kordy, H.:** Möglichkeiten und Grenzen der theoretischen Aussagekraft von multidimensionalen Skalierungen bei der Untersuchung menschlicher Informationsverarbeitung. Teil I: Formale und wissenschaftstheoretische Grundlagen. März 1979

Diskussionspapier Nr. 15: **Groeben, N.:** Entwurf eines Utopieprinzips zur Generierung psychologischer Konstrukte. Juni 1979

Diskussionspapier Nr. 16: **Weinert, F.E. & Treiber, B.:** School Socialization and Cognitive Development. Juni 1979

Diskussionspapier Nr. 17: **Gundlach, H.:** Inventarium der älteren Experimentalapparate im Psychologischen Institut Heidelberg sowie einige historische Bemerkungen. 1978

Diskussionspapier Nr. 18: **Scheele, B. & Groeben, N.:** Zur Rekonstruktion von subjektiven Theorien mittlerer Reichweite. Eine Methodik-Kombination von halbstandardisiertem Interview (einschließlich Konfrontationstechnik) und Dialog-Konsens über die Theorie-Rekonstruktion mittels der Struktur-Lege-Technik (SLT). Dezember 1979

Diskussionspapier Nr. 19: **Gloger-Tippelt, G.:** Subjektive Theorien von Frauen über ihre erste Schwangerschaft: Theoretische Konzepte und methodische Möglichkeiten. Januar 1980

Diskussionspapier Nr. 20: **Kämmerer, A.:** Das Konzept 'psychotherapeutische Strategie' am Beispiel des Problemlösens. Juli 1980

Diskussionspapier Nr. 21: **Scheele, B.:** (unter Mitarbeit von B. Tuschen und C. Maier): Subjektive Theorien über Ironie - als Heuristik für einen wissenschaftlichen Hypothesenkorpus. August 1980

Diskussionspapier Nr. 22: **Treiber, B.:** Erklärung von Förderungseffekten in Schulklassen durch Merkmale subjektiver Unterrichtstheorien ihrer Lehrer. Oktober 1980

Diskussionspapier Nr. 23: **Röhrle, B. & Kommer, D.:** Handlungstheoretische Betrachtungen zur primären Prävention psychischer Störungen. Februar 1981

Diskussionspapier Nr. 24: **Voigt, F.:** Die Entwicklung des Zahlbegriffs. Teil I: Entwicklungslinien des Zahlbegriffs im Vorschulalter: Übersicht über theoretische Probleme und empirische Untersuchungen, mit einer Bibliographie zur Zahlbegriffsentwicklung. Teil II: Entwicklungslinien des Zahlbegriffs im Vorschulalter: Deskriptive Untersuchung des kindlichen Zahlverständnisses und verwandter Konzepte. April 1981. Teil III: Trainingsstudien zum Erwerb konkreter Operationen (unter besonderer Berücksichtigung von Modellen der Invarianzaufgabe). Teil IV: Die Trainierbarkeit ordinaler und kardinaler Konzepte und ihre Beziehung zum Zahlbegriff. Juli 1982

Diskussionspapier Nr. 25: **Schneider, G. & Weimer, E.:** Aspekte der Kategorisierung städtischer Umwelt - Eine empirische Untersuchung. Juni 1981

Diskussionspapier Nr. 26: **Schneider, W. & Scheibler, D.:** Zur Evaluation numerischer Klassifikation: Probleme beim Vergleich von Clusteranalysen. August 1981

Diskussionspapier Nr. 27: **Drinkmann, A. & Groeben, N.**: Techniken der Textorganisation zur Verbesserung des Lernens aus Texten: Ein metaanalytischer Überblick. November 1981

Diskussionspapier Nr. 28: **Graumann, C.F.**: Theorie und Geschichte. November 1982, Historische Reihe Nr. 1

Diskussionspapier Nr. 29: **Woodward, W.R.**: From the Science of Language to Völkerpsychologie: Lotze, Steinthal, Lazarus and Wundt. November 1982, Historische Reihe Nr. 2

Diskussionspapier Nr. 30: **Sommer, J.**: Dialogische Forschungsmethoden. Dezember 1982

Diskussionspapier Nr. 31: **Wintermantel, M. & Christmann, U.**: Textverarbeitung: Empirische Untersuchung zum Verstehen einer Personbeschreibung. Januar 1983

Diskussionspapier Nr. 32: **Schmalhofer, F.**: Text Processing with and without Prior Knowledge: Knowledge- versus Heuristic-Dependent Representations. Februar 1983

Diskussionspapier Nr. 33: **Métraux, A.**: Victor de l'Aveyron oder Zum Streit zwischen Kulturalisten und Biologisten am Anfang des 19. Jahrhunderts. Mai 1983, Historische Reihe Nr. 3

Diskussionspapier Nr. 34: **Graumann, C.F.**: Wundt - Bühler - Mead - Zur Sozialität und Sprachlichkeit menschlichen Handelns. Mai 1983, Historische Reihe Nr. 4

Diskussionspapier Nr. 35: **Gundlach, H.**: Folk Psychology and Social Psychology oder Das Los des Ausdrucks 'Völkerpsychologie' in den englischen Übersetzungen der Werke Wundts. Mai 1983, Historische Reihe Nr. 5

Diskussionspapier Nr. 36: **Woodward, W.R.**: Hermann Lotze's Concept of Function: Its Kantian Origin and its Impact on Evolutionism in the United States. Mai 1983, Historische Reihe Nr. 6

Diskussionspapier Nr. 37: **Schneider, G.**: Reflexivität als Grenzproblem einer kognitiven Psychologie. August 1983

Diskussionspapier Nr. 38: **Geuter, U.**: 'Gleichschaltung' von oben? Universitätspolitische Strategien und Verhaltensweisen in der Psychologie während des Nationalsozialismus. Oktober 1983, Historische Reihe Nr. 11

Diskussionspapier Nr. 39: **Kruse, L.**: Drehbücher für Verhaltensschauplätze oder: Scripts for Settings. Dezember 1983

Diskussionspapier Nr. 40: **Graumann, C.F.**: The individualisation of the social and the desocialisation of the individual - Floyd H. Allport's Contribution to Social Psychology -. Mai 1984, Historische Reihe Nr. 10

Diskussionspapier Nr. 41: **Kruse, L. & Graumann, C.F.**: Environmental Psychology in Germany. November 1984

Diskussionspapier Nr. 42: **Kany, W. & Schneider, G.**: Ein linguistisch fundiertes inhaltsanalytisches System zur Erfassung des referentiellen und prädikativen Gehalts verbaler Daten. Mai 1985

Diskussionspapier Nr. 43: **Hormuth, S.E.**: Methoden für psychologische Forschung im Feld: Erfahrungsstichprobe, Autophotographie und Telefoninterview. Februar 1985

Diskussionspapier Nr. 44: **Haeberle, E.J.**: Die Anfänge der Sexualwissenschaft in Berlin. April 1985, Historische Reihe Nr. 12

Diskussionspapier Nr. 45: **Schmalhofer, F. & Schäfer, I.**: Lautes Denken bei der Wahl zwischen benannt und beschrieben dargebotenen Alternativen. Juni 1985

Diskussionspapier Nr. 46: **Zielinski, W. & Rott, C.H.**: Analyse der Entwicklung des Wortleseprozesses bei erfolgreichen und schwachen Lesern der Grundschule. Februar 1986

Diskussionspapier Nr. 47: **Waller, M.**: Metasprachliche Entwicklung: Forschungsgegenstand, Schwerpunkte, Desiderate und Perspektiven der empirischen Forschung. Juli 1986

Diskussionspapier Nr. 48: **Gundlach, H.**: Inventarium der älteren Experimentalapparate im Psychologischen Institut Heidelberg sowie einige historische Bemerkungen (zweite, vermehrte Auflage). September 1986, Historische Reihe Nr. 9

Diskussionspapier Nr. 49: **Klüpfel, J. & Graumann, C.F.**: Ein Institut entsteht - Zur Geschichte der Institutionalisierung der Psychologie an der Universität Heidelberg -. Oktober 1986, Historische Reihe Nr. 13

Diskussionspapier Nr. 50: **Drinkmann, A.**: Private und öffentliche Self-Consciousness: Eine Zwischenbilanz ihrer empirischen Bewährung. Oktober 1986

Diskussionspapier Nr. 51: **Blickle, G. & Groeben, N.**: Gegen einen objektivistisch halbierten Kognitivismus: Kognitiv-konstruktives Sprachverstehen und nicht-paradoxe Wirkungen von Lob und Tadel. November 1986

Diskussionspapier Nr. 52: **Scheele, B. & Groeben, N.**: Eine Dialog-Konsens-Variante der Ziel-Mittel-Argumentation. Dezember 1986

Diskussionspapier Nr. 53: **Batz, W.-D., Bickes, C., Bickes, H., Busse, D. & Lörch, B.**: Konzeptuelle Strukturen in der Sprache des Vorurteils. Dezember 1986

Diskussionspapier Nr. 54: **Röhrle, B.**: Soziale Netzwerke und Unterstützung. Januar 1987

Diskussionspapier Nr. 55: **Sommer, J.**: Der Signifikanztest in der psychologischen Forschung. Ein Falsifikationsinstrument im Sinne des Kritischen Rationalismus? März 1987

Diskussionspapier Nr. 56: **Batz, W.-D.**: Kodierung und Repräsentation - über hypothetische Mechanismen in Gedächtnistheorien. Dezember 1987

Diskussionspapier Nr. 57: **Bastine, R.**: Psychotherapeutische Prozeßanalyse. September 1987

Diskussionspapier Nr. 58: **Amelang, M. & Krüger, C.**: Kindesmißhandlung. November 1989

Diskussionspapier Nr. 59: **Amelang, M.**: An Investigation of the Factorial Structure and External Validity of Social Intelligence. Dezember 1987

Diskussionspapier Nr. 60: **Bastine, R.**: Klinische Psychodiagnostik. März 1988

Diskussionspapier Nr. 61: **Waller, M.**: Die Entwicklung der Beurteilung fehlerhafter Äußerungen - Eine Pilotstudie. Juni 1988

Diskussionspapier Nr. 62: **Schahn, J. & Holzer, E.**: Untersuchungen zum individuellen Umweltbewußtsein. August 1989

Diskussionspapier Nr. 63: **Stössel, A. & Scheele, B.:** Nomothetikorientierte Zusammenfassung Subjektiver Theorien zu übergreifenden Modalstrukturen. Januar 1990

Diskussionspapier Nr. 64: **Aschenbrenner, K.M., Laier, R. & Albert, D.:** Wichtigkeit als Wissen über die Variation der Merkmalsattraktivität bei der Verhaltenswahl. Dezember 1989

Diskussionspapier Nr. 65: **Albert, D., Gertzen, H., Bürgy, R., Bannert, M. & Schneyer, Th.:** Abruf semantisch strukturierter Informationen beim binären Wählen zwischen beschriebenen Alternativen. Dezember 1989

Diskussionspapier Nr. 66: **Albert, D., Lages, M., Gertzen, H. & Aschenbrenner, K.M.:** Beeinflussen Struktureigenschaften von Wissen das Wahlverhalten? Dezember 1989

Diskussionspapier Nr. 67: **Gertzen, H., Bettinger, C., Körner, Chr. & Albert, D.:** Bewertende Vergleiche und Informationsabruf in Abhängigkeit von beurteilter Dimensionswichtigkeit bei unvollständig beschriebenen Alternativen. Dezember 1989

Diskussionspapier Nr. 68: **Kane, G., Rotter, B. & Waller, M.:** Konstruktion und Erprobung einer Entwicklungsskala zur Erfassung vorsprachlich-gestischer Äußerungen bei geistig behinderten Kindern. Ergebnisse einer Pilotstudie. Januar 1991

Diskussionspapier Nr. 69: **Krüger, C. & Amelang, M.:** Arbeitslosigkeit und Kriminalität. Mai 1991

Diskussionspapier Nr. 70: **Groeben, N. & Erb, E.:** Reduktiv-implikative versus elabrativ-prospektive Menschenbildannahmen in psychologischen Forschungsprogrammen. Dezember 1991

Diskussionspapier Nr. 71: **Albert, D., Schrepp, M., Held, Th.:** Construction of Knowledge Spaces for Problem Solving in Chess - Two Experimental Investigations. März 1992

Diskussionspapier Nr. 72: **Kany, W., Waller, M.:** Desiderate einer entwicklungspsychologischen Theorie des Spracherwerbs: Eine Positionsbestimmung gegenüber der nativistischen Auffassung Chomskys. Februar 1992

Diskussionspapier Nr. 73: **Kadijk, M.**: Plotting Activations in Neural Networks. Oktober 1992

Diskussionspapier Nr. 74: **Unnewehr, J.:** Benutzerhandbuch Prozeduren zur Wissensdiagnose. Dezember 1992

Diskussionspapier Nr. 75: **Erb, Egon:** Die Kontraststruktur menschlichen Denkens zwischen Dogmatismus als kurzschlüssiger Polarisierung und polarer Integration als Entwicklungsziel. Dezember 1992

Diskussionspapier Nr. 76: **Christmann, U., Dienes, A., Groeben, N., Kolz, A., Schreier, M., Shaery, K.**: Ist unser Land wirklich ein `volles Boot'? November 1993

Diskussionspapier Nr. 77: **Erb, E.**: Selbstentwicklung/-sozialisation statt Selbstaktualisierung/-verwirklichung? Zur konzeptuellen Kritik und Weiterführung des zentralen Motiv-Konstrukts der Humanistischen Psychologie. Dezember 1993

Diskussionspapier Nr. 78: **Schahn, J., Erasmy, P., Trimpin, A., Ditschun, K.**: Psychologische Maßnahmen zur Förderung von Hausmüllvermeidung und Hausmülltrennung. Februar 1994

DiskussionspapierNr.79: **Nüse, R.**: New New Representationalism. Eine philosophisch-psychologische Wahrnehmungstheorie. Juni 1994

Diskussionspapier Nr.80: **Schahn, J., Dinger, J., Bohner, G.**: Eine Rolle von Rationalisierungen und Neutralisationen für die Rechtfertigung umweltschädigenden Verhaltens. Mai 1994

Diskussionspapier Nr.81: **Held, T., Schrepp, M., Fries, S.**: Vergleich von Methoden zur Ableitung von Wissensstrukturen. Dezember 1994