Behebung von Inkonsistenzen in der probabilistischen Expertensystem-Shell SPIRIT

W. Rödder, L. Xu

FernUniversität in Hagen

Kurzfassung: Die Expertensystem-Shell SPIRIT ist ein professionelles Instrument zur Akquisition und Deduktion von Wissen. Diese Schritte vollziehen sich durch Aufbau und Transformation von Wahrscheinlichkeitsverteilungen, unter Wahrung des Prinzips maximaler Entropie bzw. minimaler relativer Entropie. Die Kommunikationssprache zwischen Benutzer und Wissensbasis besteht in der Zuweisung von Wahrscheinlichkeiten zu Konditionalen über endlichwertigen propositionalen Variablen sowie zu Konjunktionen, Disjunktionen und Negationen solcher Konditionale. Sogar bedingte Konditionale sind bildbar, womit dem Benutzer bei Akquisition und Wissensverarbeitung ein reiches linguistischen Kommunikationsmittel zur Verfügung steht. Der Reichtum der Sprache birgt jedoch das Risiko inkompatibler oder inkonsistenter Sprachelemente und in Folge das Risiko fehlerhafter Inferenz. Drei Vorgehensweisen zur Behebung von Inkonsistenzen werden vorgestellt. Die erste besteht in der Zuweisung von Wahrscheinlichkeitsintervallen statt Wahrscheinlichkeiten zu Konditionalen, die zweite im Abgleich verschiedener, in sich jeweils konsistenter Wissenswelten und die dritte in der Minimierung eines Inkonsistens-maßes. Ein Demonstrationsbeispiel führt durch die drei Alternativen in SPIRIT.

Schlüsselwörter: Wissensverarbeitung, Entropie, relative Entropie, Inferenz, Inkonsistenz, SPIRIT.

1. Einleitung

Seit der bahnbrechenden Arbeit von Lauritzen und Spiegelhalter [LSP] wird das Wissen über Abhängigkeiten zwischen Variablen und werden Deduktionen daraus –bei Evidentwerden gewisser Tatsachen– gern in Form von Wahrscheinlichkeitsverteilungen und deren Transformationen modelliert. Zur Handhabung großer Wahrscheinlichkeitsräume bedient man sich dabei gewisser Unabhängigkeitsannahmen, die eine Faktorisierung der Verteilungen ermöglichen und in azyklischen gerichteten Graphen –sogenannten Bayesnetzen– visualisiert werden. Unbefriedigend bei dieser Vorgehensweise bleibt die aufwendige Konstruktion eines solchen Netzes ebenso wie die oft durch nichts zu begründenden Unabhängigkeitsannahmen; vergleiche hierzu auch die Anmerkungen über Schwächen von Bayesnetzen in [ROX]. Seit einigen Jahren hat sich eine andere Form der Wissensakquisition und -verarbeitung etabliert, die auf die Vorgabe graphischer Strukturen verzichten kann. Sie benötigt jedoch andere „Skelette", um aus partieller Information über die Wissensdomäne eine eindeutige Verteilung aufzubauen oder eine Deduktion durchzuführen. Diese Skelette sind das Prinzip maximaler Entropie bzw. minimaler relativer Entropie, siehe z. B. [ROM], [ROK]. Obwohl in den zitierten Arbeiten vorgestellt, werden die Prinzipien in Abschnitt 2.1 kurz entwickelt. Sie bieten große Vorteile bei der Verarbeitung von Wissen:

- Anders als in Bayesnetzen können Verteilungen aufgrund partiellen Wissens aufgebaut werden.
- Als Kommunikationsmittel zwischen Benutzer und Wissensbasis steht eine Sprache mit reicher Syntax zur Verfügung.
- Das intendierte Wissen wird unverzerrt angelegt, die Wissensakquisition ist informationstreu.

Die beiden ersten Aspekte werden in Abschnitt 2.2 herausgearbeitet, zu dem letzten studiere man [ROX].

Die Vorteile einer reichen Sprache zwischen Benutzer und Wissensbasis bergen jedoch die Gefahr der Formulierung von in sich inkonsistenten Wissensteilen. *Wenn 99 % aller Vögel fliegen, Pinguine zwar Vögel, jedoch mit an Sicherheit grenzender Wahrscheinlichkeit (99,9 %) flugunfähig sind, so steht dies im Widerspruch zur Annahme einer 5 %-igen Pinguin-Häufigkeit in der Gesamtpopulation; nicht aber einer 0.5 %-igen*. Bei komplexeren als dem vorgestellten Kontexten sind solche Inkonsistenzen schwer identifizierbar. Wie man Widersprüche korrigiert bzw. auflöst, ist Gegenstand des Kapitels 3. Kapitel 4 fasst zusammen und verweist auf zukünftige Arbeiten.

2. Mathematische Grundlagen

2.1 Konditionallogik

Grundelement der Modellierung einer Wissensbasis ist eine endliche Menge endlichwertiger Variabler $V = \{V_1,\ldots,V_L\}$. Ausdrücke der Form $V_e = v_e$ – Variable gleich Wert – heißen Literale. Literale sind logische Formeln, die für ein Individuum oder Objekt *wahr* oder *falsch* sein können. Die Menge aller logischen Formeln, aus Literalen mittels der Konnektive Negation, Konjunktion und Disjunktion gebildet, ist die Sprache L. Elemente aus L werden mit großen, evtl. indizierten, Buchstaben bezeichnet: A, B, C, ..., A_i, E_j, G_k. Ist auf den Vollkonjunktionen $v = v_1 \ldots v_L$ eine Wahrscheinlichkeitsfunktion P angegeben, lässt sich für jedes $A \in L$ seine Wahrscheinlichkeit $P(A) = \sum_{v \subset A} P(v)$ berechnen, wobei die Summation über alle solche Vollkonjunktionen läuft, die Implikant von A sind. Damit ist ein Wahrscheinlichkeitsraum auf der Menge aller logischen Formeln geschaffen. Die Wahrscheinlichkeit von Ω, der Menge *aller* Vollkonjunktion, ist 1; die Wahrscheinlichkeit einer widersprüchlichen Formel ist 0. Ist $v \subset A$ schreiben wir $A\ (v) = t$ (*wahr*), im gegenteiligen Fall schreiben wir $A\ (v) = f$ (*falsch*); die logische Formel A ist unter v *wahr* bzw. *falsch*.

Calabrese [CAL] definiert Konditionale

$$B|A(v) = \begin{cases} t & \text{falls} \quad v \subset AB \quad (= A \wedge B) \\ f & \text{falls} \quad v \subset A\overline{B} \quad (= A \wedge \overline{B}) \\ u & \text{falls} \quad v \subset \overline{A} \end{cases} \quad . \tag{1}$$

Hier bezeichnen $\overline{}$ die Negation, $\wedge$ die Konjunktion und bedeutet u „nicht definiert". Mit $B|\Omega = B$ ist L in die Menge aller Konditionale, L^*, eingebettet. Haben $\overline{}$, $\wedge$, $\vee$ (oder) die üblichen Wirkungen und gilt ferner $t \vee u = t \wedge u = t$, $f \vee u = f \wedge u = f$, $u \vee u = u \wedge u = u$, $\overline{u} = u$, so können *komponierte Konditionale* punktweise definiert werden:

$$[(B|A) \wedge (D|C)](v) = [(B|A)(v)] \wedge [(D|C)(v)],$$
$$[(B|A) \vee (D|C)](v) = [(B|A)(v)] \vee [(D|C)(v)], \tag{2}$$

$$\left(\overline{B|A}\right)(v) = \overline{(B|A)(v)}\ .$$

Hat der Konditionaloperator | die Wirkungen

$t|t = t|u = t,\ \ f|t = f|u = f\ , \cdot|f = u|\cdot = u$,

so können *bedingte Konditionale* punktweise definiert werden:

$$[(B|A)|(D|C)](v) = (B|A)(v)\,|\,(D|C)(v). \tag{3}$$

Zwar sind alle komponierten und bedingten Konditionale auf *einfache* rückführbar, siehe hierzu [CAL] oder [ROE], jedoch gestatten sie in der Form (2) oder (3) Aussagen großer Komplexität über die Wissensdomäne.

Beispiel 1

Mit den mnemotechnischen Kürzeln M, K_1, K_2, D, $Ö$, B, L, V sind folgende Aussagen selbsterklärend.

$[M|(K_1 \wedge K_2)] \wedge [(K_1 \wedge K_2)|M](v)$: Unter v ist die Maschine M funktionsfähig genau dann, wenn die Komponenten K_1 und K_2 funktionsfähig sind.

$[(D \vee Ö) \wedge (B|D)]|S(v)$: Unter v ist Herr Schmidt Deutscher oder Österreicher; ist er Deutscher, kommt er aus Berlin.

$[C \vee (V|\overline{C})]|L(v)$: Herr Lee ist unter v Chinese; falls nicht, ist er Vietnamese.

Für ein Konditional und für $P(A) > 0$ definieren wir $P(B|A) = P(AB)|P(A)$. Wegen der Rückführbarkeit bedingter oder komponierter Konditionaler auf einfache ist auch für sie die Wahrscheinlichkeit definiert.

2.2 Konditionallogik und Inferenz

Mit Konditionalen $B_i|A_i$ und deren in einer Domäne erwünschten Wahrscheinlichkeiten x_i kann man nun eine Verteilung maximaler Entropie oder minimaler relativer Entropie zur Gleichverteilung P^0 durch Lösung der folgenden Aufgabe erzeugen.

$$P^* = arg\ min\, R\left(Q, P^0\right) = \sum_v Q(v)\log_2 Q(v)\big/P^0(v) \text{ u. d. N. } Q(B_i|A_i) = x_i \qquad i = 1...I. \tag{4}$$

Für eine axiomatische Begründung der Berechnung von P^* beim Aufbau einer Wissensbasis siehe u. a. [KIS].

Angesichts evidenten oder situativen Wissens in Form weiterer Konditionale $F_j|E_j$ und deren Wahrscheinlichkeiten y_j löst man:

$$P^{**} = arg\ min\, R\left(Q, P^{*}\right) = \sum_{v} Q(v) \log_2 \frac{Q(v)}{P^{*}(v)} \text{ u. d. N. } Q\left(F_j \middle| E_j\right) = y_j\,; \quad j = 1...J. \tag{5}$$

Die axiomatische Begründung zu (4) umfasst auch den Fall (5).

Wertet man ein Konditional $H|G$ unter P^{**} aus, d. h. bezeichnet man $P^{**}(H|G) = z$, so ist z die aus dem Basiswissen P^{*} unter evidenten Annahmen deduzierte Antwort zur Frage $H|G$ = ? Diese Antwort ist informationstreu!

Der in diesem Abschnitt entwickelte Inferenzmechanismus steht in der Expertensystem-Shell zur Deduktion von Wissen aus Tausenden von Konditionalen zur Verfügung, vgl. [ROM].

3. Behebung von Inkonsistenzen

Ist das Benutzerwissen $Q(B_i|A_i) = x_i \quad i = 1...I$ in (4) inkonsistent, bieten sich drei Möglichkeiten zur Behebung an, die nun entwickelt werden.

<u>Abschwächung zugunsten von Intervallen</u>

Mittels einer binären Hilfsvariablen W_i wird die Restriktion $Q(B_i|A_i) = x_i$ durch $Q(B_i|A_i\,W_i) = x_i'$ *und*

$$Q(B_i|A_i\overline{W_i}) = x_i''\,. \tag{6}$$

mit $x_i' \leqq x_i \leqq x_i''$ ersetzt. Wie man leicht nachrechnet, erzwingt (6) die Gleichung $Q(B_i|A_i) = x_i'\,Q(W|A_i) + x_i''\,Q(\overline{W}|A_i)$. $Q(B_i|A_i) = x_i$ ist also eine Konvexkombination von x_i' und x_i'' und kann im Intervall [x_i', x_i''] variieren. Durch (6) wird die ursprüngliche Forderung $Q(B_i|A_i) = x_i$ abgeschwächt, bei Lösung von (4) mit der Restriktion (6) stellen sich im Optimum $P^{*}(B_i|A_i) = \overline{x_i}$ ein.

<u>Abgleich in sich konsistenter Restriktionsmengen</u>

Bei dieser Behebungsform wird unterstellt, dass eine Partition $U = \{U_1,...,U_k\}$ mit $U_k \subset \{1,...,I\}$ in sich konsistenter Restriktionsmengen $Q(B_i|A_i) = x_i \quad i \in U_k$ existiert, das Restriktionensystem in (4) insgesamt jedoch inkonsistent ist. In [ROX] entwickeln die Autoren die Vorstellung von K Experten, von denen jeder über in sich konsistentes Wissen verfügt, das mit dem der übrigen jedoch nicht kompatibel ist. Durch Einführung von K binären Expertenvariablen W_K werden diese Teilwissensdomänen entzerrt:

$$Q(B_i|A_i\,W_k) = x_i \qquad i \in U_k \qquad k = 1...K\,. \tag{7}$$

Jede Teilwissensbasis ist in der *Welt* W_k erfüllt, und durch eine parametrische Analyse nähert man nun die Welten an:

$$P^*(\bar{x}) = arg\ min\, R(Q, P^0) \text{ u. d. N. } Q(B_i | A_i\, W_k) = x_i \qquad i \in U_k \qquad k = 1 \ldots K$$

$$Q\left(\wedge_k W_k \,\middle|\, \vee_k W_k\right) = \bar{x}\,. \tag{8}$$

Je größer $\bar{x}$, umso wahrscheinlicher sind die übrigen Welten, falls auch nur eine Welt wahr ist. Ein weiterer Rechenschritt gestattet nun den Abgleich der x_i zu $\overline{x_i}$ so, dass Gesamtkonsistenz erreicht wird: Mit der Lösung der Aufgabe

$$Q^* = arg\ min\, R(Q, P^*(\bar{x})) \text{ u. d. N. } Q\left(\wedge_k W_k\right) = 1 \tag{9}$$

rechnet man $\overline{x_i} = Q^*(B_i | A_i)$. Die $\overline{x_i}$ heißen U-Kompromiss zum Niveau $\bar{x}$.

<u>Minimierung eines Inkonsistenzmaßes</u>

Sind die Restriktionen in (4) nicht erfüllbar, wird man versuchen, die geforderten Wahrscheinlichkeiten x_i durch solch x_i zu ersetzen, die eine zulässige Verteilung Q ermöglichen und möglichst nahe an den x_i liegen. Durch die folgende Aufgabe wird dieses Ziel erreicht, wie wir in den anschließenden Begründungen nachweisen.

$$Min\ Z = \sum_i y_i \log_2 \frac{y_i}{z_i} + z_i \log_2 \frac{z_i}{y_i} \text{ u. d. N. } y_i = (1 - x_i)\, Q(B_i A_i),\ z_i = x_i Q(\overline{B}_i A_i) \quad i = 1 \ldots I \tag{10}$$

Begründung: a) Löst $\mathbf{y}^*, \mathbf{z}^*, Q^*$ das Problem (10), so ist mit $\overline{x_i} = Q(B_i | A_i)$ die intendierte Aufgabe gelöst, falls nur $Q(A_i) > 0$ für alle i.

b) $Z = D(\mathbf{y}, \mathbf{z}) + D(\mathbf{z}, \mathbf{y}) = \left(\sum y_i \log_2 \frac{y_i}{z_i} - y_i + z_i\right) + \left(\sum z_i \log_2 \frac{z_i}{y_i} - z_i + y_i\right)$. D heißt verallgemeinerte Divergenz der positiven Vektoren x und y, ist nicht negativ und stellt ein gegenüber der relativen Entropie verallgemeinertes informationstheoretischen Abstandsmaß dar.

c) Wegen b) ist $z \geqq 0$. Sind alle $Q(A_i) > 0$, so gilt $z = 0$ genau dann, wenn $Q(B_i | A_i) = x_i$ für alle i. $Z = 0$ entspricht also dem Erfülltsein der ursprünglichen Restriktionen in (4) – Konsistenz.

d) $\frac{y_i}{z_i} = \frac{1 - x_i}{x_i} \cdot \frac{x_i}{1 - x_i} = 1$, die Minimierung von z bedeutet also die gleichzeitige Minimierung der Abstände von $(x_1, \ldots, x_I)$ und $(x_1, \ldots, x_I)$. Man beachte hierzu, dass $\frac{1 - x_i}{x_i} \cdot \frac{x_i}{1 - x_i}$ genau dann, wenn $x_i = x_i$. Zur Lösung von (10) wurde in SPIRIT ein Algorithmus implementiert, dessen zentraler Schritt die Aufgabe $\overline{Q} = arg\ min\, R(Q, P)$ u. d. N. $Q(A) = x$ ist, kurz $\overline{Q} = PS(P, A[x])$.

<u>Algorithmus</u>

1. $Q^1 = P^0$, die Gleichverteilung.

Für $l = 1 \ldots$ und $k = l\ mod\ I$ bis Abbruch führe aus:

2. Löse in r_k (mittels Newton-Verfahren) die Aufgabe

$$\bar{r}_k = arg\ min\ F(r_k) = \sum_i y_i(r_k)\log_2 \frac{y_i(r_k)}{z_i(r_k)} + z_i(r_k)\log_2 \frac{z_i(r_k)}{y_i(r_k)}$$

u. d. N. $y_i(r_k) = (1-x_i)\,\bar{a}(B_i A_i)\ \ z_i = x_i\, \overline{Q}(\overline{B}_i A_i)$

$$Q = PS\left(Q^k, A_k B_k[r_k]\right).$$

Weise zu $\overline{Q}^k = PS\left(Q_l^k, A_k B_k[\bar{r}_k]\right)$.

3. Löse in r_k (mittels Newton-Verfahren) die Aufgabe

$$\bar{r}_k = arg\ min\ F(s_k) = \sum_i y_i(s_k)\log_2 \frac{y_i(s_k)}{z_i(s_k)} + z_i(s_k)\log_2 \frac{z_i(s_k)}{y_i(s_k)}$$

u. d. N. $y_i(s_k) = (1-x_i)\overline{\overline{Q}}(B_i|A_i)\ \ z_i = x_i\, \overline{\overline{Q}}(\overline{B}_i A_i)$

$$\overline{\overline{Q}} = PS\left(\overline{Q}^k, A_k \overline{B}_k[s_k]\right)$$

Weise zu $\overline{\overline{Q}}^k = PS\left(\overline{Q}, A_k \overline{B}_k[\bar{s}_k]\right)$.

4. Ist Q^* nach Abbruch die errechnete Grenzverteilung, berechne $\bar{x}_i = Q^*(B_i|A_i)$ für alle i.

Bemerkung a) In den Schritten 2 und 3 des Algorithmus konvergiert das Newton-Verfahren wegen der Konvexität der Aufgabe (10) in Q. Der Beweis ist aufwendig und kann hier nicht geführt werden.

b) Die Optimalität des Algorithmus ist nicht beweisbar, die Optimalität des gefundenen $\overline{\overline{Q}}$ kann jedoch mittels der Kuhn-Tucker-Bedingungen zu (10) überprüft werden.

Zum Abschluss sei der in der Einleitung vorgestellt Widerspruch in der Vogel-Population mit den drei Behebungsformen *Intervall*, *U-Kompromiss* und *Inkonsistenzmaß* aufgelöst.

Beispiel 2 Mit den mnemotechnischen Kürzeln V = Vogel, P = Pinguin, F = fliegen hatten wir $Q(F|V) = x_1 = .99$, $Q(V|P) = x_2 = 1.0$, $Q(F|P) = x_3 = .001$, $Q(P) = x_4 = .0$. Bei Intervallvorgaben $[x_1', x_1''] = [.96,.99]$, $[x_3', x_3''] = [0.00, 0.02]$, $[x_4', x_4''] = [0.00, 0.07]$ und Beibehaltung des sicheren Wissens, dass Pinguine Vögel sind, hat man

	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$
Intervall	0.970	0.010	0.020
U-Kompromiss	0.990	0.005	0.004
Inkonsistenzmaß	0.977	0.002	0.023

Alle drei Behebungsformen sind wichtige Instrumente zum Aufbau von Wissensbasen bei komplexen Zusammenhängen. Mit jeder Form verfolgt der Benutzer eine *andere* Philosophie der Auflösung von Widersprüchen, mithin mit verschiedenen Ergebnissen.

Der Leser ist aufgefordert, mit der Expertensystem-Shell SPIRIT zu experimentieren und die Auflösungsmöglichkeiten auszutesten. Sie finden SPIRIT-Versionen unter

http://www.fernuni-hagen.de/BWLOR/forsch.htm

Literatur

[CAL] P. G. Calabrese: *Deduction and Inference Using Conditional Logic and Probability*, in: Conditional Logic in Expert Systems, I. R. Goodman, M. M. Gupta, H. T. Nguyen and G. S. Rogers (editors). Elsevier Science Publishers B. V., 71 – 100, (1991).

[KIS] G. Kern-Isberner: *Characterising the principle of minimum cross-entropy within a conditional-logical framework*, Artificial Intelligence, Vol. 98, 169 – 208, (1998).

[LSP] S. L. Lauritzen and D. J. Spiegelhalter: *Local computations with probabilities in graphical structures and their applications to expert systems*, Journal of the Royal Statistical Society 13 (2), 415 – 448, (1988).

[ROE] W. Rödder: *Conditional Logic and the Principle of Entropy*, Artificial Intelligence 117, 83 – 106, (2000).

[ROK] W. Rödder and G. Kern-Isberner: *Léa Sombé und entropie-optimale Informationsverarbeitung mit der Expertensystem-Shell SPIRIT*, OR Spektrum 19, 41 – 46, (1997).

[ROM] W. Rödder and C.-H. Meyer: *Coherent knowledge processing at maximum entropy by SPIRIT*, 12 th Conference on Uncertainty in Artificial Intelligence, E. Horvitz and F. Jensen (editors), Morgen Kaufmann, San Francisco, California: 470 – 476, (1996).

[ROX] W. Rödder and L. Xu: *Entropy-driven Inference and Inconsistency*, Artificial Intelligence and Statistics, Fort Lauderdale, Florida, 272 – 277, (1999).