

Zeitschrift: Bulletin des Schweizerischen Elektrotechnischen Vereins, des Verbandes Schweizerischer Elektrizitätsunternehmen = Bulletin de l'Association Suisse des Electriciens, de l'Association des Entreprises électriques suisses

Band: 83 (1992)

Heft: 13

Artikel: Zuverlässiges Schliessen mit unzuverlässigen Argumenten : eine Einführung in die Evidenztheorie

Autor: Kohlas, Jürg

DOI: <https://doi.org/10.5169/seals-902842>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 13.10.2024

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Zuverlässiges Schliessen mit unzuverlässigen Argumenten

Eine Einführung in die Evidenztheorie

Jürg Kohlas

In zwei Beiträgen wurde der Umgang mit unscharfen Begriffen (Fuzzy Logic) und unsicherer Information (Kausales Schliessen) vorgestellt. Eine nochmals andere Sicht bietet die Theorie des evidentiellen Schliessens, welche den Umgang mit unzuverlässiger Information zum Inhalt hat.

Les principaux aspects liés à la manipulation de connaissances floues (fuzzy logic) ainsi que la manière de traiter des informations incertaines (raisonnement causal) ont été présentés dans deux autres contributions. La théorie du raisonnement évidentiel représente encore un autre point de vue. Cette théorie s'occupe du raisonnement avec des informations qui ne sont pas fiables.

Der Umgang mit unzuverlässiger Information

Das kausale Modellieren, wie es im Beitrag «Wahrscheinlichkeit und kausales Schliessen» [2] dargestellt wurde, ist keineswegs die einzig mögliche Art, unsichere Zusammenhänge darzustellen. Eine andere Sicht der Dinge betrachtet Informationen als Evidenz¹ für bestimmte Aussagen; Evidenz allerdings, die nicht unfehlbar ist. In einer ersten approximativen Darstellung dieses Ansatzes kann man davon ausgehen, dass Evidenz zu Aussagen führt, die nicht ganz sicher sind. Solche Aussagen können dazu verwendet werden, gewisse Folgerungen zu ziehen. In dem Masse, als die Aussagen oder Argumente nicht völlig zuverlässig sind, sind dann auch die Schlüsse nur mit einer beschränkten Wahrscheinlichkeit gültig.

Die Beweise von Hypothesen mit Hilfe von unsicheren Argumenten sind vergleichbar mit technischen Systemen, deren Komponenten versagen und damit den Ausfall des Systems verursachen können. Im gleichen Sinn, wie es eine Zuverlässigkeitsrechnung technischer Systeme gibt, entsteht aus dieser Sicht eine *Zuverlässigkeitstheorie des Schliessens* mit unsicheren Argumenten. Dieser Ansatz des evidentiellen Modellierens ist weit weniger bekannt und verbreitet, als derjenige des kausalen Schliessens. Letzterer ist verwurzelt im Hauptstrom der Wahrscheinlichkeitstheorie, während der erstere neue Ansätze der Wahrscheinlichkeitsrechnung erfordert. Die Theorie des evidentiellen Schliessens wurde 1976 vom Mathematiker G. Shafer [3] eingeführt, wobei dieser von Arbeiten

des Statistikers A. Dempster [4] ausging. Die Erkenntnis des Zusammenhangs mit der Zuverlässigkeitstheorie komplexer Systeme ist noch neueren Datums.

Zur Einführung in diese alternative Denkweise kann das Beispiel der Alarmanlage aus dem zweiten Beitrag [2] betrachtet werden. Ein Alarm ist sicherlich Evidenz dafür, dass ein Einbruch stattfindet oder stattgefunden hat. Diese Evidenz wäre unfehlbar, wenn die Alarmanlage mit Sicherheit nur auf Einbrüche reagieren würde. Wenn umgekehrt kein Alarm ertönt, dann ist dies Evidenz dafür, dass kein Einbruch stattfindet. Auch diese Evidenz ist jedoch nicht ganz unfehlbar. Es ist nicht ausgeschlossen, dass die Alarmanlage bei einem Einbruch nicht anspricht.

Für eine solche Situation wollen wir ein recht allgemeines Modell, ein sogenanntes *Sensormodell* einführen. Dabei wird angenommen, dass bei einer Alarmanlage zwei Arten von Fehlerzuständen oder von Fehlfunktionen vorkommen können, jede mit einer bestimmten Wahrscheinlichkeit:

Fehler 1. Art: Die Alarmanlage spricht aus einem anderen Grund als einem Einbruch an (Wahrscheinlichkeit 0,2).

Fehler 2. Art: Die Alarmanlage spricht nicht an, obwohl ein Einbruch stattfindet (Wahrscheinlichkeit 0,05).

Ein Alarm ist somit eine Evidenz, die mit einer Wahrscheinlichkeit von 0,8 auf einen Einbruch schliessen lässt. Die Gegenwahrscheinlichkeit von 0,2 lässt aber logisch einen Einbruch nicht ausschliessen: obwohl die Anlage aus einem anderen Grund als einem Einbruch anspricht (z.B. infolge von Erschütterungen), könnte ja immer noch gleichzeitig ein Einbruch stattfinden. Die Evidenz unterstützt die Hypothese eines Einbruchs mit ei-

Adresse des Autors

Prof. Dr. Jürg Kohlas, Institut für Automation und Operations Research, Universität Freiburg, 1700 Freiburg.

¹ Evidenz [lat.] Einleuchtende Gewissheit.

nem Grad von 0,8. Dagegen unterstützt sie die Gegenhypothese keines Einbruchs gar nicht. Und da soweit kein Grund gegen einen Einbruch spricht, bleibt dieser völlig plausibel (plausibel zum Grad 1). Zu diesen Graden der Unterstützung oder Glaubwürdigkeit und der Plausibilität wird später noch einiges zu sagen sein.

Dieses Sensormodell der Alarmanlage kann auch durch zwei Regeln:

- R1 Wenn Alarm, dann Einbruch (Zuverlässigkeit 0,8);
- R2 Wenn kein Alarm, dann kein Einbruch (Zuverlässigkeit 0,95)

beschrieben werden. Jede dieser Regeln hat eine beschränkte Zuverlässigkeit, d.h. eine positive Wahrscheinlichkeit, dass sie nicht gültig ist. Die Schlüsse (Beweise) aus diesen Regeln haben dementsprechend auch nur eine begrenzte Zuverlässigkeit. Man beachte, dass die Angabe des Zuverlässigkeitsgrades nicht Teil der Regel ist, sondern eine Aussage über die Gültigkeit der Regel. Angenommen nun, die ängstliche Frau Meier ruft Herrn Huber in seinem Büro an und sagt ihm, sie glaube einen Alarm in seinem Haus gehört zu haben. Dieses Telefon ist Evidenz für einen Alarm. Die Überängstlichkeit von Frau Meier macht sie jedoch zu einem relativ unzuverlässigen Zeugen. Falls man mit 70% Sicherheit davon ausgehen kann, dass Frau Meier irgend ein anderes Geräusch für den Alarm hält, dann gilt die folgende Regel nur mit einer Zuverlässigkeit von 0,3:

- R3 Wenn Frau Meier anruft, dann ist Alarm (Zuverlässigkeit 0,3).

Der Anruf von Frau Meier erlaubt, das Ereignis eines Einbruchs abzuleiten, wenn sowohl Regel R1 als auch Regel R3 als gültig angenommen werden darf. Die Wahrscheinlichkeit dafür ist $0,3 \cdot 0,8 = 0,24$ – man darf hier multiplizieren, da die Korrektheit des Alarmsystems und diejenige von Frau Meier unabhängig voneinander sind. Der Schluss des Einbruchs aus dem Anruf von Frau Meier funktioniert nur, wenn beide Regeln R1 und R3 als gültig betrachtet werden können. Das ist genau wie bei einem Seriesystem in der technischen System-Zuverlässigkeit, bei der die Komponenten-Zuverlässigkeiten multipliziert werden.

Wenn nun Herr Huber, um sicherer zu sein, seine andere Nachbarin, Frau Müller, anläutet und diese den Alarm bestätigt, dann ist dies eine weitere

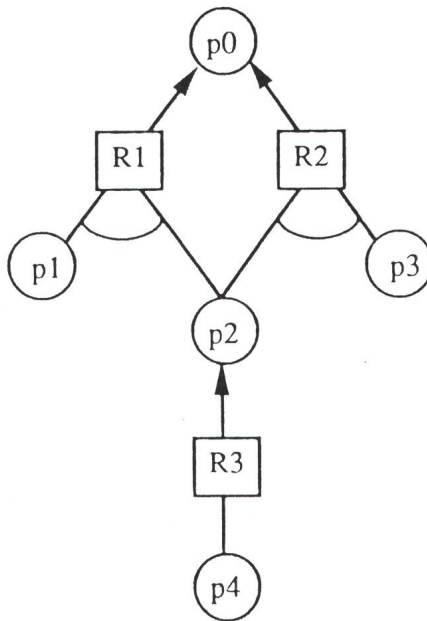


Bild 1 Graphische Darstellung eines Regelsystems

Aussagen werden durch Kreise, Regeln durch Quadrate dargestellt

Evidenz für einen Alarm – und damit indirekt auch für einen Einbruch. Da die Frau Müller zuverlässig ist, darf hier das Regelsystem:

- R4 Wenn Frau Müller Alarm bestätigt, dann Alarm (Zuverlässigkeit 0,9);
- R5 Wenn Frau Müller Alarm nicht bestätigt, dann kein Alarm (Zuverlässigkeit 0,4)

vorausgesetzt werden. Die Zeugin Frau Müller wird wie Frau Meier mit Hilfe eines Sensormodells beschrieben.

Mit welcher Sicherheit muss Herr Huber nun den Alarm als Tatsache betrachten? Dazu genügt, dass mindestens eine der beiden Regeln R3 und R4 gültig ist. Das entspricht exakt einem Parallelsystem der technischen Systemzuverlässigkeit. Die Wahrscheinlichkeit dafür ist nach der bekannten Formel

$$1 - (1-0,3)(1-0,9) = 0,93 \quad (1)$$

zu berechnen². Die solide Evidenz der Aussage von Frau Müller macht einen Alarm fast sicher. Damit wird die Gewissheit eines Einbruchs gleich

² Wahrscheinlichkeit, dass die Regeln R3 und R4 nicht gleichzeitig falsch sind, d.h., dass sich die beiden Damen nicht gleichzeitig irren.

$0,93 \cdot 0,8 = 0,744$, wenn man noch die Zuverlässigkeit des Alarmsystems mitberücksichtigt.

Regelbasierte Wissensdarstellung ist sehr beliebt in einfacheren Expertensystemen. Die obigen simplen Formeln der Seriekombination für die Verkettung von Regeln und der Parallelkombination für Regeln mit gleichem Schluss sind in einigen dieser Systeme als Ungewissheitskalkül eingebaut worden. Der Haken dabei ist jedoch, dass die Sache doch nicht ganz so einfach ist. Eine erste Schwierigkeit ist, dass oft mehrere, nicht unabhängige Regelverkettungen zu einem Schluss möglich sind. In Bild 1 ist ein einfaches Regelschema graphisch dargestellt, bei dem dies der Fall ist. Die Problematik ist ähnlich derjenigen bei Bayesschen Netzwerken mit Zyklen. Es stehen auch analoge Ansätze zu ihrer Bewältigung zur Verfügung, nämlich das sogenannte hypothetische Schliessen und die Gruppierung der Variablen zur Bildung eines sogenannten Markoffschen Baumes (ähnlich einem Bayesschen Baum).

Die andere, neue Schwierigkeit ist, dass Evidenzen widersprüchlich sein können. Dazu ein Beispiel: Wir nehmen an, dass Frau Meier angerufen und einen Alarm gemeldet hat, der beim Kontrollanruf bei Frau Müller nicht bestätigt wird. Damit steht Regel R3 im Gegensatz zu Regel R5; es können nicht beide Regeln gleichzeitig gültig sein, denn das würde zu einem Widerspruch führen. Die Wahrscheinlichkeit für diesen Widerspruch ist $0,3 \cdot 0,4 = 0,12$. Auf einen Alarm kann man nur schliessen, wenn man annimmt, dass Frau Meier zuverlässig (Regel R3 gültig) und gleichzeitig Frau Müller unzuverlässig (Regel R5 nicht gültig) ist. Die Wahrscheinlichkeit dafür ist $0,3 \cdot 0,6 = 0,18$. Einen Alarm kann man ausschliessen, wenn man Frau Meier als unzuverlässig und Frau Müller als zuverlässig annimmt. Dafür ist die Wahrscheinlichkeit $0,7 \cdot 0,4 = 0,28$. Wenn man beide Personen als unzuverlässig annimmt (Wahrscheinlichkeit $0,7 \cdot 0,6 = 0,42$), dann kann man weder einen Alarm annehmen, noch diesen ausschliessen; unter dieser Konstellation ist die Evidenz ohne Inhalt, sie lässt keinen Schluss zu. Da nun aber logisch kein Widerspruch vorkommen kann, ist die erste Kombination, in der beide Zeugen als zuverlässig angenommen werden, auszuschliessen und die Wahrscheinlichkeiten der restlichen drei Fälle sind auf die Tatsache, dass nur

einer dieser drei Fälle möglich ist, zu bedingen. Damit können die beiden Hypothesen eines Alarms bzw. keines Alarms mit den folgenden Wahrscheinlichkeiten bewiesen werden:

$$\begin{array}{ll} \text{Alarm} & 0,18/(1-0,12) = 0,205. \\ \text{Kein Alarm} & 0,28/(1-0,12) = 0,318. \end{array}$$

Die Wahrscheinlichkeit für die Richtigkeit der Hypothese, dass man den Alarm weder annehmen noch ausschliessen kann, beträgt $0,42/(1-0,12) = 0,477$.

Jetzt gibt es mehr oder weniger zuverlässige Argumente für und gegen einen Alarm und damit auch für und gegen einen Einbruch. Die widersprüchlichen Zeugenaussagen führen dazu, dass beide Hypothesen keine sehr grosse Glaubwürdigkeit erlangen und somit eine grosse Ungewissheit zurückbleibt. Diese ist in einem solchen Fall wohl auch angebracht. Die hier dargestellte Art der Kombination widersprüchlicher Evidenzen wird Regel von Dempster genannt, nach dem Statistiker A. Dempster, der sie 1967 – allerdings in einem anderen Zusammenhang – eingeführt hat [4]. Dazu wird im folgenden Abschnitt noch mehr zu sagen sein.

Das Modell kann wie im vorangehenden Teil dahingehend verfeinert werden, dass ein Alarm ebenso als Evidenz für einen Einbruch als auch für ungewöhnliche Erschütterungen, etwa infolge eines Erdbebens, gelten kann. Dann kann ein Alarm soweit nur als Argument für Einbruch *oder* Erdbeben gelten. Er lässt für sich allein genommen keinerlei Unterscheidung zwischen Einbruch und Erdbeben zu. Dazu müssen weitere Hinweise gesucht werden. Zum Beispiel kann Herr Huber die Erdbebenzentrale anrufen. Angenommen, man teilt ihm dort mit, dass in der Tat ein schwaches Beben in der Gegend des Wohnsitzes von Herrn Huber registriert wurde, so kann man dies als eine unfehlbare Aussage nehmen und damit mit Sicherheit auf ein Erdbeben schliessen. Diese Information erlaubt aber nicht, einen Einbruch auszuschliessen; es ist immer noch möglich, dass gleichzeitig ein Einbruch stattfindet. Der Alarm enthält in diesem Fall schlicht keine Information mehr bezüglich eines Einbruchs.

Hier unterscheidet sich das evidentielle Schliessen stark vom kausalen Schliessen des vorangehenden Teils. Dort erklärte das Erdbeben den Alarm und reduzierte damit die Glaubwürdigkeit eines Einbruchs.

Dieser Mechanismus hing jedoch stark damit zusammen, dass eine kleine A-priori-Wahrscheinlichkeit für einen Einbruch vorausgesetzt wurde. Da bei einem Erdbeben der Alarm keine Information für oder gegen die Hypothese eines Einbruchs enthält, kann man diese kleine A-priori-Wahrscheinlichkeit als Argument verwenden, um einen Einbruch praktisch auszuschliessen.

Anders sieht es aus, wenn die Erdbebenzentrale kein Beben registriert hat. Das kann als nicht ganz zuverlässiges Argument gegen die Existenz eines Bebens genommen werden. Angenommen, diese Auskunft der Erdbebenzentrale erlaube mit einer Zuverlässigkeit von 0,9 ein Beben auszuschliessen, dann kann unter Annahme des weiter vorne benutzten Sensormodells die Hypothese eines Einbruchs mit einer Wahrscheinlichkeit von $0,8 \cdot 0,9 = 0,72$ bestätigt werden. Mit einer Wahrscheinlichkeit von $0,8 \cdot 0,1 = 0,08$ kann auf einen Einbruch oder ein Erdbeben geschlossen werden und mit der Wahrscheinlichkeit von 0,2 kann nichts geschlossen werden (die Auskunft und der Alarm sind ohne jegliche Information).

Dieses Beispiel illustriert, wie Evidenzen als unzuverlässige Argumente für logische Schlüsse betrachtet werden können. Das erlaubt die Zuverlässigkeit der Ableitung oder Folgerung gewisser Hypothesen oder ihrer Gegenhypothesen zu berechnen. Daran kann der Grad der Glaubwürdigkeit und der Grad der Plausibilität der Hypothesen im Lichte der vorliegenden Information gemessen werden. Es versteht sich, dass dieser Ansatz formalisiert und generalisiert werden muss, wenn er eine grössere Tragweite erhalten soll. Im nächsten Abschnitt wird skizziert, wie das gemacht werden kann.

Das Konzept von Hinweisen, Grade der Glaubwürdigkeit und der Plausibilität

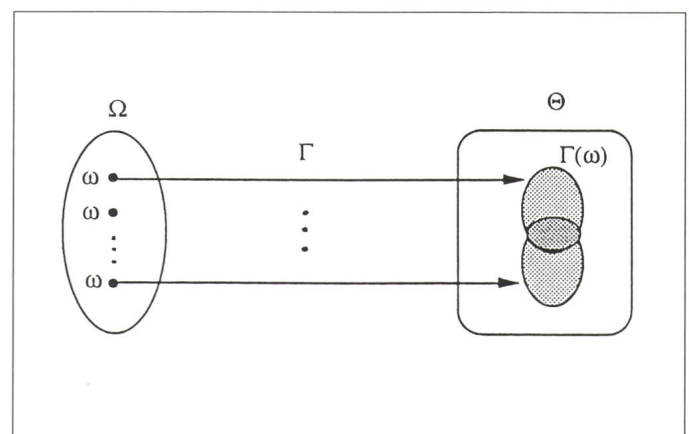
Information ist immer in bezug auf eine bestimmte Frage, die man gerne beantworten möchte, zu analysieren. Eine klar abgegrenzte Information soll hier im folgenden zur sprachlichen Vereinfachung *Hinweis* genannt werden. Wenn man also einen bestimmten Hinweis formal beschreiben will, muss man zuerst die zugehörige Frage fixieren. Sodann sind die möglichen Antworten dazu zu beschreiben, ähnlich wie bei der Darstellung vager Informationen. Diese bilden eine Menge Θ , deren Elemente θ Aussagen sind, die mögliche Antworten zur gestellten Frage darstellen (Bild 2). Dabei wird vorausgesetzt, dass Θ erschöpfend ist, d.h., dass die wahre, wenn auch unbekannte Antwort, ein Element von Θ ist. Ferner müssen die Elemente von Θ sich gegenseitig ausschliessen, d.h. nur eines kann die wahre Antwort sein.

Ein Hinweis beinhaltet im allgemeinen eine gewisse Ungewissheit; er wird nur in seltenen Fällen die gesuchte Antwort eindeutig und zweifelsfrei festlegen. Mehrere verschiedene Interpretationen der Information werden im allgemeinen möglich sein, und es wird nicht sicher sein, welches die richtige Interpretation ist. Daher wird zur Formalisierung eine Menge Ω eingeführt, deren Elemente ω die möglichen Interpretationen des Hinweises darstellen. In den meisten Fällen sind nicht alle Interpretationen gleich wahrscheinlich. Dies wird durch Wahrscheinlichkeiten $p(\omega)$ beschrieben. Diese Wahrscheinlichkeiten werden als bekannt vorausgesetzt.

Wenn nun eine bestimmte Interpretation ω als die richtige angenommen wird, dann erlaubt diese Interpreta-

Bild 2
Darstellung der Elemente eines Hinweises

Ω : Menge der möglichen Interpretationen ω ;
 Θ : Menge der möglichen Antworten;
 $\Gamma(\omega)$: Teilmenge von Θ der noch möglichen Antworten, wenn ω die korrekte Interpretation ist



tion im allgemeinen den Bereich der richtigen Antwort in Θ einzuschränken. Es kann somit unter dieser Voraussetzung die richtige Antwort mit Bestimmtheit in einer Teilmenge $\Gamma(\omega) \subset \Theta$ lokalisiert werden (Bild 2). Im Extremfall erlaubt eine Interpretation sogar, die Antwort eindeutig festzulegen. Dann besteht die Menge $\Gamma(\omega)$ nur aus einem einzigen Element θ von Θ . Es kann aber umgekehrt auch sein, dass eine Interpretation gar keine Eingrenzung der Antwort erlaubt, dann ist $\Gamma(\omega) = \Theta$. Das bedeutet, dass der Hinweis bei dieser Interpretation überhaupt keine Information zur gestellten Frage enthält. Es ist sogar möglich, dass der Hinweis unter allen seinen möglichen Interpretationen keine Information zur betrachteten Frage enthält. Ein solcher Hinweis wird als leer in bezug auf die gestellte Frage bezeichnet. Man beachte jedoch, dass der Hinweis durchaus Information in bezug auf andere Fragen enthalten kann!

Leere Hinweise gestatten es, die vollständige Ignoranz formal korrekt darzustellen. Das ist wichtig. Im Rahmen des Bayesschen Ansatzes ist es nicht möglich, die Ignoranz korrekt zu beschreiben. Üblicherweise wird vorgeschlagen, die Laplacesche Hypothese der Gleichwahrscheinlichkeit aller möglichen Antworten zu adoptieren. Das ist jedoch keine Darstellung der Ignoranz, sondern eine Default-Annahme für eine unbekannte Wahrscheinlichkeitsverteilung.

Mit Hilfe der skizzierten Konstruktion wird dem Begriff des Hinweises eine mathematische Form gegeben. Das Sensormodell des vorangehenden

Tabelle II
Formale Darstellung der Information des km-Zählers als mathematischer Hinweis zur Frage, ob der Tank leer ist oder nicht

Mögliche Antworten:

$$\Theta = \{ \text{"Tank leer"}, \text{"Tank nicht leer"} \}$$

Mögliche Interpretationen des Kilometerzählers:

$$\Omega = \{ \text{"korrekt gestellt"}, \text{"irrtümlich zurückgestellt"} \}$$

Wahrscheinlichkeiten der Interpretationen:

$$p(\text{korrekt gestellt}) = 0,5$$

$$p(\text{irrtümlich zurückgestellt}) = 0,5$$

Feststellung der Antworten unter den einzelnen Interpretationen:

$$\Gamma(\text{korrekt gestellt}) = \{ \text{"Tank nicht leer"} \}$$

$$\Gamma(\text{irrtümlich zurückgestellt}) = \Theta$$

Abschnitts kann in diese Form gekleidet werden. Ein weiteres Beispiel ist ein Benzinstandsanzeiger in einem Auto. Wenn dieser auf Null steht, dann ist dies ein Hinweis dafür, dass der Benzintank leer ist. Allerdings ist es auch möglich, dass der Zeiger defekt ist. Daher ist der Hinweis nicht völlig sicher. In der Tabelle I ist die formale Darstellung dieses Hinweises beschrieben. Der Autofahrer kann auch den Tages-Kilometerzähler zur Kontrolle des Benzinstandes beziehen. Er wird dann beim Tanken seinen Kilometerzähler jeweils auf Null zurückstellen, falls er dies nicht vergisst. Ein kleiner Kilometerstand ist für ihn ein Hinweis darauf, dass der Tank noch nicht leer sein kann, es sei denn, jemand hätte den Zähler irrtümlich zu einem späteren Zeitpunkt zurückgestellt. Die Tabelle II zeigt die

zu Tabelle I analoge Darstellung für diesen Hinweis. Die Kombination der beiden Hinweise vom Benzinstandsanzeiger und Kilometerzähler wird weiter hinten besprochen.

Es wird nun möglich, Hypothesen im Lichte eines Hinweises zu beurteilen. Eine Hypothese ist eine Aussage über die Antwort zur betrachteten Frage und kann als Teilmenge H von Θ definiert werden. Eine solche Hypothese kann richtig oder falsch sein. Das erstere ist der Fall, wenn die wahre Antwort ein Element von H ist, das zweite ist der Fall, wenn die wahre Antwort kein Element von H ist. Was kann nun zugunsten und gegen eine Hypothese gesagt werden?

Man kann sich fragen, unter welchen Interpretationen ω die Hypothese H notwendig richtig ist, d.h. sich schlüssig beweisen lässt. Das ist genau der Fall für diejenigen Interpretationen ω , für die $\Gamma(\omega) \subset H$ gilt. Unter einer solchen Interpretation liegt die wahre Antwort in $\Gamma(\omega)$ und damit notwendigerweise auch in H . Falls also eine solche Interpretation korrekt ist, muss H richtig sein. Nun weiss man zwar nicht, ob eine solche Interpretation korrekt ist, aber man kann mindestens die Wahrscheinlichkeit dafür ausrechnen. Dazu muss man nur die Wahrscheinlichkeiten aller entsprechenden Interpretationen mit $\Gamma(\omega) \subset H$ aufsummieren. Diese Wahrscheinlichkeit sei mit $sp(H)$ (Support of H) bezeichnet. $sp(H)$ gibt an, mit welcher Wahrscheinlichkeit die Hypothese H mit Hilfe der Information im Hinweis bewiesen werden kann. Aus diesem Grund wird $sp(H)$ Grad der Unterstützung oder der Glaubwürdigkeit von H genannt.

Tabelle I
Formale Darstellung der Information des Benzinstandsanzeigers als mathematischer Hinweis zur Frage, ob der Tank leer ist oder nicht

Mögliche Antworten:

$$\Theta = \{ \text{"Tank leer"}, \text{"Tank nicht leer"} \}$$

Mögliche Interpretationen des Benzinstandsanzeigers:

$$\Omega = \{ \text{"funktionstüchtig"}, \text{"defekt"} \}$$

Wahrscheinlichkeiten der Interpretationen:

$$p(\text{funktionstüchtig}) = 0,9$$

$$p(\text{defekt}) = 0,1$$

$$p(\text{funktionstüchtig}) + p(\text{defekt}) = 1$$

Festlegung der Antworten unter den einzelnen Interpretationen:

$$\Gamma(\text{funktionstüchtig}) = \{ \text{"Tank leer"} \} \subset \Theta$$

$$\Gamma(\text{defekt}) = \{ \text{"Tank leer"}, \text{"Tank nicht leer"} \} = \Theta$$

Man kann aber auch fragen, unter welchen der möglichen Interpretationen des Hinweises die Hypothese H überhaupt möglich, d.h. nicht ausgeschlossen ist. Das ist für alle Interpretationen der Fall, für die $\Gamma(\omega) \cap H \neq \emptyset$ gilt. Da unter einer solchen Interpretation die wahre Antwort in $\Gamma(\omega)$ liegen muss, ist es möglich, dass sie auch in H liegt; H ist damit nicht ausgeschlossen. Man kann erneut die Wahrscheinlichkeit solcher Interpretationen berechnen. Sie sei mit $pl(H)$ bezeichnet. $pl(H)$ wird auch Grad der Plausibilität der Hypothese H bezeichnet. Auf diese Weise wird jede mögliche Hypothese durch einen Hinweis mit zwei verschiedenen Graden versehen: dem Grad (oder der Zuverlässigkeit), mit dem die Hypothese bewiesen werden kann und dem Grad der Plausibilität. Letzterer ist auch das Komplement der Zuverlässigkeit, mit dem das Gegenteil der Hypothese bewiesen werden kann. Eine Hypothese darf im Lichte eines Hinweises um so eher akzeptiert oder als richtig betrachtet werden, als der Grad der Unterstützung sich dem Wert 1 nähert. Die Hypothese ist um so eher abzulehnen oder als falsch zu beurteilen, je kleiner der Grad der Plausibilität ist.

Oft hat man zwei oder mehrere Hinweise zu einer Frage (man erinnere sich an den vorangehenden Abschnitt). Es stellt sich dann die Frage, wie solche Hinweise zu kombinieren sind. Da man davon ausgehen kann, dass in jedem der Hinweise genau eine Interpretation die korrekte ist, muss es genau eine korrekte kombinierte Interpretation geben. Bei zwei Hinweisen besteht diese aus einem Paar (ω_1, ω_2) worin ω_1 die korrekte Interpretation des ersten Hinweises und ω_2 diejenige des zweiten Hinweises ist. Da man aber die jeweiligen korrekten Interpretationen nicht kennt, muss man zunächst alle solchen Paare von Interpretationen als mögliche Interpretationen des kombinierten Hinweises betrachten. Wenn man darüber hinaus die beiden Hinweise als voneinander unabhängig betrachten darf, dann kann man die Wahrscheinlichkeiten der Interpretationen ω_1 und ω_2 multiplizieren und mit $p_1(\omega_1)$ $p_2(\omega_2)$ die Wahrscheinlichkeit der kombinierten Interpretation (ω_1, ω_2) erhalten.

Da unter der Interpretation ω_1 die wahre Antwort in $\Gamma_1(\omega_1)$ liegen muss und unter der Interpretation ω_2 in $\Gamma_2(\omega_2)$, muss sie unter der kombinierten Interpretation in beiden Mengen,

also im Durchschnitt $\Gamma_1(\omega_1) \cap \Gamma_2(\omega_2)$ liegen. Es ist nun allerdings möglich, dass dieser Durchschnitt leer ist. Das bedeutet aber, dass die beiden Interpretationen widersprüchlich sind. Sie können nicht beide zugleich korrekt sein. Man muss alle solchen Paare ausschließen und dann die bedingten Wahrscheinlichkeiten unter der Bedingung keines Widerspruchs nehmen. Das ist genau was bei der Kombination der widersprüchlichen Aussagen von Frau Meier und Frau Müller im letzten Abschnitt getan wurde. Auf diese Weise entsteht aus zwei Hinweisen ein neuer, kombinierter Hinweis. Diese Art Hinweise zu kombinieren, wird die Regel von Dempster genannt. Mit ihr entsteht ein eigentlicher Kalkül mit Hinweisen, der es gestattet, Evidenz aus unterschiedlichen Quellen zu verarbeiten und zu integrieren.

In der Tabelle III ist das Beispiel der Tabellen I und II zur Illustration weitergeführt. Es werden nun beide Hinweise – vom Benzinstandsanzeiger und vom Kilometerzähler – verwendet, um die Grade der Unterstützung und der Plausibilität der Hypothese, dass der Tank leer ist, zu überprüfen.

Es versteht sich, dass dies nur eine knappe Darstellung der Grundzüge einer mathematischen Theorie der Hinweise ist. Die Theorie muss darüber hinaus Elemente enthalten, die einen dynamischen Wechsel des Brennpunkts der Diskussion und der Argumentation zwischen verschiedenen, aber verwandten Fragen erlaubt. Vor allem müssen auch effiziente Rechenschemata zur Automatisierung des evidentiellen Schliessens verfügbar gemacht werden. Alles dies und noch mehr ist mittlerweile erarbeitet

worden. Für weitergehende Information hierzu sei auf die einführende Darstellung in [5] verwiesen. Im anschließenden Abschnitt soll nun noch eine typische Anwendung des evidentiellen Schliessens im Ingenieurbereich aufgezeigt werden.

Evidentielles Schliessen in der modellgestützten Diagnostik

Wenn man ein Rechenwerk wie in Bild 3 betrachtet und dabei die beobachteten Eingabe- und Ausgabewerte analysiert, dann wird sofort klar, dass gewisse Module des Netzwerkes defekt sein müssen. Die Frage ist nur, welche es sind. Das ist mittlerweile eine bereits klassische Fragestellung der künstlichen Intelligenz, vgl. [6], dem auch das Beispiel entnommen ist. Eine typische Diagnose könnte lauten, dass das Addierwerk A1 oder das Multiplizierwerk M1 defekt ist, es ist aber auch möglich, dass die beiden Multiplizierwerke M2 und M3 oder M2 und das Addierwerk A2 defekt sind. Um zu solchen Antworten zu kommen (insbesondere zu den letzten beiden), muss allerdings das Modell weiter spezifiziert werden; vor allem die möglichen Fehlermodi müssen genauer definiert werden. Bei obigen Antworten ist zum Beispiel vorausgesetzt, dass auch ein defektes Werk richtige Ergebnisse produzieren kann.

Man hat aber, unabhängig von den undefinierten Details, doch den Eindruck, dass nicht alle Diagnosen gleich wahrscheinlich oder gleich plausibel sind. Und eine genauere Analyse dieser Wahrscheinlichkeiten könnte durchaus nützlich, sogar unumgänglich sein. Man kann wohl da-

Kombinierte Interpretationen		Antwort zur kombinierten Interpretation	Wahrscheinlichkeit
Benzinstandsanzeiger	Kilometerzeiger		
funktionsstüchtig	korrekt gestellt	Widerspruch!	$0,9 \times 0,5 = 0,45$
funktionsstüchtig	irrtümlich zurückgestellt	Tank leer	$\frac{0,9 \times 0,5}{(1 - 0,45)} = 0,82$
defekt	korrekt gestellt	Tank nicht leer	$\frac{0,1 \times 0,5}{(1 - 0,45)} = 0,09$
defekt	irrtümlich zurückgestellt	Ø, d.h. Tank leer oder nicht leer	$\frac{0,1 \times 0,5}{(1 - 0,45)} = 0,09$

Tabelle III Anwendung der Regel von Dempster bei der Kombination von zwei verschiedenen Hinweisen, nämlich einem Benzinstandsanzeiger, der auf leer steht und einem Kilometeranzeiger, der eine kleine Kilometerzahl angibt

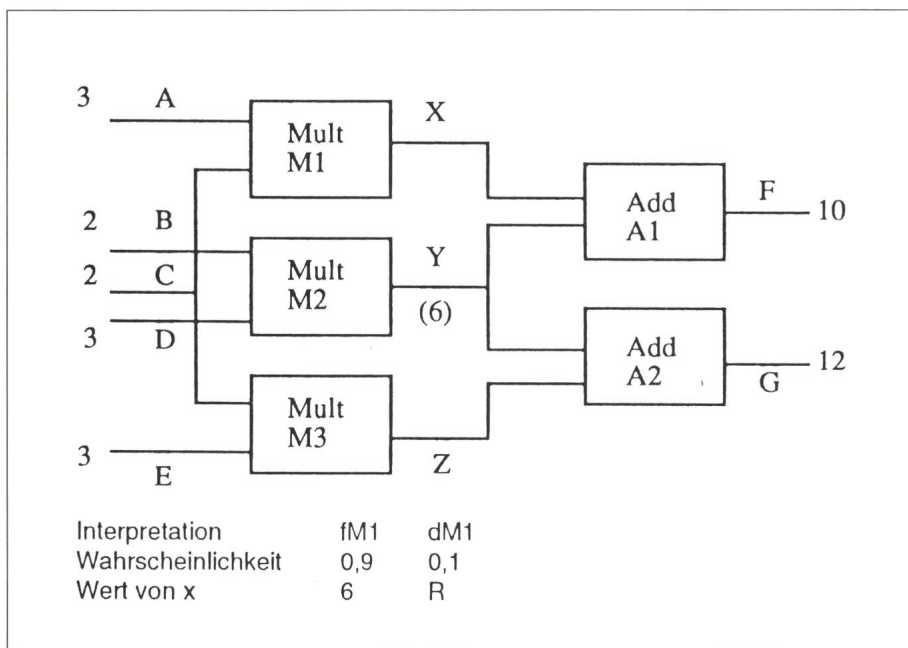


Bild 3 Ein defektes Rechenwerk: Welche Komponenten sind defekt?

Man darf annehmen, dass die gleichzeitigen Ausfälle von M2 und M3 oder M2 und A2 unwahrscheinlich sind. Das bedeutet, dass der Ausgang Y von M2 das korrekte Resultat 6 enthält

von ausgehen, dass Erfahrungswerte für die Fehlerwahrscheinlichkeiten der Addier- und der Multiplizierwerke vorliegen. So sei die Fehlerwahrscheinlichkeit eines Addierwerkes 0,01 und diejenige eines Multiplizierwerkes 0,1. Jedes Werk kann also in einem der beiden Zustände f = «funktionstüchtig» oder d = «defekt» sein. Je nach Zustand eines Werkes ist die Input-Outputbeziehung entweder (bei f) durch die entsprechende arithmetische Relation $in1 + in2 = out$ bzw. $in1 \cdot in2 = out$ definiert oder (beim Zustand d) völlig un spezifiziert, d.h. alle Tripel $(in1, in2, out)$ sind möglich. Damit erhält man einen Hinweis im Sinne des vorangehenden Abschnitts über die Funktionsart eines Werkes. Die möglichen Zustände f und d des Werkes entsprechen den möglichen Interpretationen des Hinweises. Andere Funktionsmodelle sind auch denkbar, zum Beispiel könnte vorausgesetzt werden, dass bei einem Defekt das korrekte Resultat mit Sicherheit nicht vorkommt, oder dass der Output eines Werkes auf einem bestimmten Wert fixiert ist usw.

Die gleichzeitigen Ausfälle von M2 und M3 oder M2 und A2 sind relativ unwahrscheinlich und werden zur Vereinfachung der Rechnung im folgenden ausgeschlossen. Es soll aber betont werden, dass die Analyse durchaus auch ohne diese vereinfachende Annahme durchgeführt werden könnte. Unter diesen Annahmen ist M2 sicher funktionstüchtig, so dass für die Variablen y der korrekte Wert 6 angenommen werden darf (Bild 3).

Die Frage ist nun, wie weit die vorliegende Information für den Ausfall des Addierwerkes A1 spricht. Diese Frage wird mit Hilfe der Regel von Dempster analysiert: Falls M1 korrekt funktioniert ($fM1$), muss der Wert der Variablen x am Ausgang von M1 gleich 6 sein; falls M1 defekt ist ($dM1$), kann dieser Wert beliebig sein (\mathcal{R}). Der Hinweis über die Funktion des Werkes transformiert sich damit in einen Hinweis über den möglichen Wert der Variablen x (Bild 3). Diese Information ist nun mit dem Hinweis über das Funktionieren von A1 zu kombinieren: Zuerst ist festzustellen, dass die Annahme der Funktionstüch-

tigkeit von M1 und A1 ($fM1$ und $fA1$) zu einem Widerspruch führt und damit auszuschliessen ist. Die nachfolgenden Wahrscheinlichkeiten sind daher mit dem Komplement $1 - 0,9 \cdot 0,99 = 0,109$ zu dividieren, um die entsprechenden bedingten Wahrscheinlichkeiten zu erhalten. Die Kombination der Zustände $(fM1, dA1)$, $(dM1, fA1)$ und $(dM1, dA1)$ sind alle möglich und erhalten die Wahrscheinlichkeiten 0,083, 0,908 und 0,009. Die erste und dritte Kombination implizieren einen Defekt im Addierwerk A1, womit die Wahrscheinlichkeit des Ausfalls dieses Werks gleich 0,092 wird.

In einer analogen Analyse kann man die Wahrscheinlichkeit des Ausfalls des Multiplizierwerkes M1 berechnen. Diese wird gleich 0,917. Man beachte, dass die Ausfallwahrscheinlichkeiten der beiden Werke zusammen auf einen Wert über Eins summieren, da ja auch die Möglichkeit des gleichzeitigen Ausfalls beider Werke besteht. In diesem einfachen Netz hätte man diese Wahrscheinlichkeiten mit genügender Näherung einfach aus dem Verhältnis 1:10 der Ausfallwahrscheinlichkeiten der beiden Werke schliessen können. Diese einfache Rechnungsart versagt jedoch in komplexeren Netzen. Dann wird eine Kombination der vorliegenden Information in der dargelegten Art unumgänglich. Natürlich kann das zugrundeliegende Rechenschema verallgemeinert und automatisiert werden. Das ist ein Beispiel, bei dem evidentielles Schliessen die Grundlage für einen Inferenzmechanismus eines Expertensystems werden kann.

Literatur

- [1] J. Kohlas: Fuzzy Logic – Die Darstellung und Verarbeitung vager Information. Bull. SEV/VSE 83(1992)1, S. 19–25.
- [2] J. Kohlas: Wahrscheinlichkeit und kausales Schliessen – Darstellung von ungewisser Information. Bull. SEV/VSE 83(1992)9, S. 49–52.
- [3] G. Shafer: A Mathematical Theory of Evidence. Princeton University Press 1976.
- [4] A.P. Dempster: Upper and Lower Probabilities Induced by a Multivalued Mapping. Annals of Mathematical Statistics, pp. 325–339.
- [5] J. Kohlas, P.A. Monney: Modeling and Reasoning with Hints. University of Fribourg, Institute for Automation and Operations Research, Working Paper No. 174, 1990.
- [6] J. De Kleer, B.C. Williams: Diagnosing Multiple Faults. Artificial Intelligence 32(1987), pp. 97–130.