

# LEX—Ein juristisches Expertensystem mit natürlichsprachlichem Dialog (Teil 1)

Ein Forschungsprojekt der Universität Tübingen und der IBM Deutschland

Jürgen Sulz/Reimund Baumann

## A. Einleitung

Im September 1984 vereinbarten die Eberhard-Karls-Universität Tübingen und die IBM Deutschland GmbH eine Zusammenarbeit auf dem Gebiet der Expertensysteme. Angestrebt wurde der Entwurf, die Realisierung und Erprobung eines Systems zum Erwerb, zur Speicherung und zur Abfrage von Wissen. Inzwischen ist ein leistungsfähiges Expertensystem, das System Lex, entwickelt.

Lex kann in natürlicher Sprache eingegebene Sachverhalte bearbeiten und die darin enthaltenen juristischen Probleme einer juristischen Lösung zuführen. Auf in natürlicher Sprache eingegebene Fragen zu einem Text, dem eine Originalentscheidung des OLG Nürnberg zu § 142 StGB (Unfallflucht) zugrunde liegt, antwortet Lex, wiederum in natürlicher Sprache, erstaunlich genau und juristisch richtig. Dies alles kann das System nur leisten, weil es den linguistischen Bedeutungsgehalt des eingegebenen Textes „versteht“ und so die juristische Relevanz der im Text enthaltenen Probleme prüfen kann. „Verstehen“ kann das System nur das, was ihm an Allgemeinwissen und linguistischem Wissen von uns eingegeben wurde. Erst das so „Verstandene“ kann das System bearbeiten.

Das System zerfällt in zwei aus juristischer Sicht relevante Komponenten: Die Wissensbasis (unten B), in welcher das eingegebene Wissen abgelegt ist, und eine Wissensverarbeitungskomponente (unten C.), mit welcher sich Lex die Wissensbasis nutzbar macht. Nur diese beiden Komponenten wollen wir hier in vereinfachter Form darstellen.

## B. Wissenspräsentation mittels DRS (Wissensbasis)

Im alltäglichen Leben vollzieht sich die Wissensvermittlung durch den Formalismus der natürlichen Sprache. Für LEX mußte ein Formalismus entwickelt werden, der es ihm erlaubt, natürliche Sprache zu verarbeiten. Einen solchen Formalismus bilden die DRS'en (Discourse Representation Structures)<sup>1</sup>. Mit ihnen ist natürliche Sprache und das darin enthaltene Wissen formal repräsentierbar. Dieses kann dann in der Wissensbasis des Systems abgelegt und verfügbar gemacht werden.

Stark vereinfacht ausgedrückt liegt der Wissensrepräsentation mittels DRS'en folgender Gedanke zugrunde:

Natürliche Sprache gewinnt ihren Bedeutungsgehalt durch Worte, die aneinander gereiht werden zu Sätzen und so aus sich heraus oder in Verbindung mit anderen Sätzen einen Sinnzusammenhang ergeben, wobei sich der Bedeutungsgehalt einzelner Wörter über Sätze und Satzzusammenhänge zu einem Bedeutungsgehalt höherer Ordnung steigert. In ähnlicher Weise erfolgt die Wissensrepräsentation mittels DRS'en. Natürlichsprachlichen Aussagen niedriger Ordnung — meist einzelnen Wörtern — werden sog. „Variable“ zugeordnet, die den Bedeutungsgehalt der jeweiligen Aussage repräsentieren. Durch diese Verbindung von natürlichsprachlichen Aussagen mit formalsprachlichen Argumenten (Variable) entstehen sog. „Prädikate“. Aussagen höherer Ordnung werden wiederum durch Verknüpfung einzelner Prädikate gebildet. Die so verknüpften Prädikate ergeben die „Regel“ der Aussage höherer Ordnung. LEX kennt somit nur drei standardisierte formalsprachliche Ausdrucksmittel: „Variable“, „Prädikate“ und „Regeln“. Es hat sich gezeigt, daß mit diesen formalsprachlichen Ausdrucksmitteln nahezu alle natürlichsprachlichen und juristischen Aussagen formalisiert werden können.

Verdeutlichen lassen sich diese Zusammenhänge an der Umsetzung des Merkmals „Verletzen“ in eine DRS. Die Umsetzung erfolgt in vier Schritten:

### 1. Erster Schritt:

#### *Ausgangsüberlegungen*

„Verletzen“ im juristischen Sinne bedeutet stets die körperliche Beeinträchtigung eines Menschen. Umgangssprachlich dagegen kann auch eine Sache im juristischen Sinne, etwa ein Hund verletzt werden. Umgangssprachlicher und juristischer Bedeutungsgehalt der Aussage „Verletzen“ können somit voneinander abweichen. Daher verfügt LEX über zwei Wissensbereiche in seiner Wissensbasis: Den Bereich des juristi-

<sup>1</sup> Die Theorie zu dieser Sprache (abgekürzt „DRS“) stammt von Kamp „A Theory of Truth and Semantic Representation“, Groenendijk, J. et. al.: Formal Methods in the Study of Language; MC TRACTS 135, Universität Amsterdam, p. 277.

schen Regelwissens und den linguistischen Bereich<sup>2</sup>. Der linguistische Bereich repräsentiert das Wissen über den Bedeutungsgehalt natürlichsprachlicher Aussagen. Der Bereich des Regelwissens repräsentiert dagegen das Wissen über die Struktur und den Bedeutungsgehalt einer Aussage im juristischen Sinne. Im Beispiel soll „Verletzen“ in eine juristische DRS aus dem Bereich des Regelwissens umgesetzt werden. „Verletzen“ aus dem Bereich des Regelwissens wird daher mit „Verletzen\_j“ kenntlich gemacht.

„Verletzen\_j“ ist in § 223 StGB definiert. Danach ist „verletzen“ gegeben, wenn entweder eine „körperliche Mißhandlung“ oder eine „Gesundheitsbeschädigung“ bei einem Menschen vorliegt. Es läßt sich daher allein mit den Merkmalen

„körperliche Mißhandlung“,  
„Gesundheitsbeschädigung“,  
„Mensch“

definieren.

Diesen Merkmalen sind nun „Variable“ zuzuordnen, so daß „Prädikate“ entstehen, die die (juristische) „Regel“ Verletzen\_j ergeben.

## 2. Zweiter Schritt:

### *Variablen-Belegung*

Variable sind Platzhalter für Elemente aus den Grundmengen aller Objekte, Ereignisse, Orte und Zeiten. Die Elemente aus der Grundmenge aller Objekte werden mit den Variablen „ux“ für Objekte wie Haus, Auto aber auch abstrakte Begriffe wie Norm, Pflicht etc. belegt. Entsprechend werden die Elemente der Grundmenge aller Ereignisse mit „ey“, die Elemente aus der Grundmenge aller Orte mit „In“ und die Elemente aller Zeiten mit „tm“ Variablen belegt. Dabei sind „x“, „y“, „m“ und „n“ Platzhalter für natürliche Zahlen. So ist für jedes „u“ und „e“ sichergestellt, daß jedes „ux“ bzw. „ey“ genau das Objekt „x“ bzw. Ereignis „y“ repräsentiert. Auf die eigene Darstellung von Zuständen wurde bei LEX verzichtet. Ein Zustand wird durch zwei Ereignisse definiert, das ihn begründende und das ihn beendende Ereignis.

Die gefundenen Aussagen, die „Verletzen\_j“ definieren, sind nun daraufhin zu untersuchen, inwieweit sie Elemente der genannten Grundmengen enthalten. Zunächst werden die gefundenen Aussagen auf ihren juristischen Bedeutungsgehalt hin analysiert und in ihre einzelnen Komponenten zerlegt. Die Aussage „körperliche Mißhandlung“ impliziert die juristische Bedeutung, daß ein mißhandeltes Objekt, ein weiteres Objekt, das die Mißhandlung verursacht hat, das Resultat der Mißhandlung und das Ereignis der Mißhandlung vorliegen. Diese Implikation vollzieht sich beim Menschen normalerweise automatisch im Kopf. Zur Umsetzung in DRS'en muß sie jedoch rationalisiert und jede juristische Aussage in Objekte, Ereignisse etc. zerlegt werden. Für die Aussage „körperliche Mißhandlung“ ergeben sich somit die Variablen

u1 = mißhandeltes Objekt,  
u2 = mißhandelndes Objekt,  
u3 = Resultat der Mißhandlung und  
e1 = Ereignis der Mißhandlung.

Für die Aussagen „Gesundheitsbeschädigung“ und „Verletzen\_j“ wird ebenso verfahren.

Die Auflösung einzelner Aussagen in Variablen ist notwendig, damit LEX die Beziehungen der in den Aussagen erwähnten Individuen erkennen und bearbeiten kann.

## 3. Dritter Schritt:

### *Prädikatenbildung*

Prädikate sind Zuordnungen (Relationen). Ein Prädikat ordnet einer Aussage n-Argumente zu. Die Zuordnung erfolgt formal durch Klammerung der Argumente im Anschluß an die Aussage. Je nach Anzahl der einer Aussage zugeordneten Argumente spricht man von ein bis n-stelligen Prädikaten.

### a) Einfache Prädikate

Werden einer Aussage als Argumente nur Variable zugeordnet, spricht man von einfachen Prädikaten. Anschaulich wird dies wiederum am vorliegenden Beispiel. Werden der Aussage „körperliche Mißhandlung“ die gefundenen Variablen u1, u2, u3 und e1 zugeordnet, so entsteht das vierstellige Prädikat

„körperliche Mißhandlung (u3, u2, u1, e1)“.

Entsprechend werden die Prädikate

„Gesundheitsbeschädigung (u3, u2, u1, e1)“

und

„Verletzen\_j“ (u1, u2, u3, e1)“

gebildet.

Werden in einer Aussage auch Ort und Zeitbestimmungen getroffen, sind weitere einfache Prädikate notwendig. Deutlich wird dies beispielsweise an dem Satz: „Am 19. November fuhr Müller Maier auf der A-Straße an“. Ordnet man dieser Aussage Argumente in der Form der Variablen u1 für Fahrer (Müller), u2 für Angefahrener (Maier) und e1 für das Ereignis „anfahen“ zu, dann entstehen die Prädikate:

„anfahen (u1, u2, e2)“

„Fahrer (u1)“

„Angefahrener (u2)“

Zur präzisen Beschreibung des Beispielsatzes müssen weitere Prädikate gebildet werden: t1 zur Bestim-

<sup>2</sup> Auf beide Bereiche wird weiter unten noch näher eingegangen.

mung des Zeitpunktes und tim (= time) zur Verknüpfung von Zeitpunkt und Ereignis sowie zur Verknüpfung von Ereignis und Ort, wobei „lp“ local place bedeutet.

„19.11. (t1)“  
 „tim (t1, e1)“  
 „A-Straße (l1)“  
 „lp (l1)“

Nun sind alle im o.g. Satz getroffenen Einzelaussagen mit Argumenten (Variablen) belegt. Die so entstandenen einzelnen Prädikate repräsentieren noch nicht die im Beispielsatz getroffene Gesamtaussage oder die gesamte juristische Implikation des Merkmales „verletzen\_j“. Diese werden erst durch die logische Verknüpfung der Prädikate in der Form von „Regeln“ im vierten Schritt repräsentiert.

#### b) Metaprädikate

Werden einer Aussage als Argumente Variable und/oder weitere Prädikate zugeordnet, dann spricht man von Metaprädikaten. Einfache Prädikate enthalten somit nur unbegrenzt, aber unendlich viele Variable, Metaprädikate zusätzlich noch zumindest ein weiteres Prädikat als Argumente.

Metaprädikate dienen zur Erleichterung des Beweisverfahrens. Beispielsweise wurde das Metaprädikat

„user(Sachschaden(...))“

für den Fall eingeführt, wenn im Text kein Hinweis auf einen wesentlichen Sachschaden enthalten ist, das System fragt dann den Benutzer nach weiteren Informationen.

#### 4. Vierter Schritt: Regelbildung

Regeln sind komplexe Aussagen, die durch die Verknüpfung der in einzelnen Prädikaten getroffenen Aussagen durch Junktoren entstehen.<sup>3</sup> LEX verwendet als Junktoren: „&“, den Konjunktoren (und), „|“, den Disjunktoren (nicht-ausschließendes oder) sowie „¬“ den Negator (nicht) und den Implikator „→“ (folgt).<sup>4</sup>

Deutlich wird dies, wenn die oben gefundenen Prädikate so miteinander verknüpft werden, daß sich die Regel ergibt, mit der „Verletzen\_j“ definiert wird. Werden die gefundenen Prädikate mit dem Disjunktoren (nicht ausschließendes oder) verknüpft, ergibt sich folgende DRS, die von sämtlichen weiteren Formalien befreit ist:

DRS (.....

körperliche\_Mißhandlung (u3, u2, u1, e1)|  
 Gesundheitsbeschädigung (u3, u2, u1, e1)

→ Verletzen\_j (u1, u2, u3, e1).

Nunmehr ist die DRS entstanden, die Verletzen im juristischen Sinne definiert.

### C. Muster — DRS

#### a. DRS — ‚Unfall‘

Der natürlichsprachliche Satz „Unfall ist ein plötzliches Ereignis, welches unmittelbar zu einem Schaden führt“<sup>5</sup> ergibt folgende DRS:

Zeile 1: DRS...(e1. e2. u1. u2...#

|   |                                |
|---|--------------------------------|
| 2 | Schaden_j (u1, u2, e2).        |
| 3 | plötzlich_j (e1).              |
| 4 | unmittelbare_Ursache (e1, e2)) |
| 5 | → Unfall (e1, e2).             |

Dies ist die Struktur einer DRS. Sie ist von allen Zeichen befreit, die nur für eine maschinelle Verarbeitung erforderlich sind. Dennoch ist sie auf den ersten Blick immer noch verwirrend. Nachfolgend seien daher einige Grundprinzipien dieser DRS erklärt.

#### b. Grundprinzipien einer DRS

Diese DRS gibt die Definition des Begriffes „Unfall“ wieder. Mit ihr soll, ausgehend von einem natürlichsprachlich eingegebenen Text, das Vorliegen des Merkmales Unfall mittels der Merkmale „plötzlich“, „Schaden“ und „unmittelbare Ursache“ bewiesen werden.

Die DRS hat einen dreigeteilten Aufbau:

##### aa. Zeile 1

Den ersten Teil bildet die erste Zeile. Sie beginnt mit ‚DRS‘. Dies ist notwendig, damit das Beweissystem die Prädikate und die Regel erkennt.

Nach ‚DRS‘ folgt die Nummer der DRS. Das Beweissystem benötigt diese nicht. Die Nummer ist lediglich Hilfsmittel für den Programmierer. Durch sie kann der Standort der DRS im gesamten juristischen Regelwissen bestimmt werden.

Nach der Nummer der DRS wird eine Klammer eröffnet, die erst am Ende des zweiten Teiles wieder geschlossen wird. Durch die geöffnete Klammer erkennt das Beweissystem, daß nunmehr die Merkmale aufgelistet werden, mit denen „Unfall“ bewiesen werden soll. Innerhalb dieser Klammer erfolgen weitere Klammern. Sofern es sich dabei nicht um die Zuordnung der Variablen zu den Prädikaten handelt, kann hierdurch der Beweis gesteuert werden, etwa um deutlich zu machen, daß ein Disjunktoren Vorrang vor einem Konjunktoren hat.

<sup>3</sup> Schneider, Hans-Jochen (Hrsg.), Lexikon der Informatik und Datenverarbeitung, München Wien 1986, S. 307.

<sup>4</sup> vgl. dazu Koch, Hans-Joachim/Rußmann, Helmut, Juristische Begründungslehre, München 1982, S. 31 ff.; Herberger, Maximilian/Simon, Dieter, Wissenschaftstheorie für Juristen, Frankfurt am Main 1980, S. 34 ff.

<sup>5</sup> BGHStE 8, 263, 264 f.: Unfall ist ein plötzliches Ereignis, das zur Tötung oder Verletzung eines Menschen oder zu einer nicht völlig belanglosen Sachbeschädigung führt.

Unmittelbar hinter der ersten geöffneten Klammer folgt die Liste der in der DRS vorkommenden Variablen<sup>6</sup>. In der Muster-DRS sind dies vier Variable, weil mit diesen „Unfall“ vollständig definiert werden kann.

Die Liste von Variablen endet stets mit ‚#‘, damit das Beweissystem die Trennung der Variablen von den Prädikaten erkennt. Die Liste kann beliebig lang sein und muß grundsätzlich alle Variablen enthalten, die in der DRS vorkommen. Hiervon wird abgewichen, wenn etwa eine Variable in allen Regeln, und nicht nur in einer einzelnen DRS einen festen Wert hat.

bb. Zeile 2 bis einschließlich Zeile 4

Den zweiten Teil der DRS bilden die Zeilen bis zum Folgepfeil ‚->‘. Sie enthalten die Prädikate mit denen „Unfall“ definiert wird. Dies ist das Definiens. In der Muster-DRS sind dies drei Prädikate. Diese sind durch den Konjunktoren ‚&‘ verknüpft. Diese Verknüpfung ergibt die Regel<sup>7</sup>, mit der „Unfall“ definiert ist.

Die Prädikate in den Zeilen 2 bis einschließlich 4 sind demnach wie folgt zu lesen:

aaa. Schaden \_j (u1, u2, e2)

Zeile 2 enthält das Prädikat „Schaden \_j“. „Schaden \_j“ bedeutet Schaden im juristischen Sinne (s.u. # A. III). Schaden ist hier kein Zustand, weil LEX eine derartige Kategorie nicht kennt. Es war daher notwendig, das Prädikat Schaden \_j in Objekte und Ereignisse zu zerlegen. Demnach wurde Schaden \_j definiert durch das Schadensereignis ‚e2‘, und die beiden Objekte ‚u1‘ für den Schädiger und ‚u2‘ für den Geschädigten. Die Bezeichnung des Schadensereignisses mit ‚e2‘ war notwendig, weil Zeile 3 ein weiteres, von diesem zu unterscheidendes Ereignis enthält. Entsprechendes gilt für die Objekte. Das Prädikat führt daher in seiner Klammer die Variablen u1, u2, e2 mit.

bbb. plötzlich \_j (e1)

Zeile 3 enthält das Prädikat plötzlich \_j (im juristischen Sinne) mit der Variablen ‚e1‘. Plötzlich wird da-

mit als Ereignis definiert, nämlich als das schadensverursachende Ereignis.

ccc. unmittelbare \_Ursache (e1, e2)

Zeile 4 enthält das Prädikat unmittelbare \_Ursache mit den Variablen ‚e1‘ für das schadensverursachende Ereignis und ‚e2‘ für das Ereignis des Schadens selbst. Dies bringt zum Ausdruck, daß beide Ereignisse in einem Ursachenzusammenhang stehen müssen.

Die Zeilen 2 bis einschließlich 4 werden durch den Konjunktoren ‚&‘ verknüpft, denn nach der Definition „Unfall“ müssen sämtliche Prädikate kumulativ vorliegen. Die geschlossene Klammer am Ende des zweiten Teils hat zur Folge, daß die in der Klammer enthaltenen Variablen allquantifiziert werden. Allquantifiziert heißt, daß alle geklammerten Variablen frei sind zur Belegung mit bestimmten Objekten, Ereignissen, Zeiten und Orten aus der jeweiligen Grundmenge, d.h. insbesondere mit den in einer konkreten Sachverhaltsdarstellung vorkommenden. So stellt etwa der Autofahrer Maier ein bestimmtes Objekt aus der Grundmenge aller Objekte dar.

cc. Zeile 5

Den dritten Teil bildet die Zeile mit dem Folgepfeil ‚->‘. Sie gibt das Definiendum wieder — hier das Prädikat Unfall. Der Aussage „Unfall“ sind die Argumente in der Form der Variablen ‚e1‘ für das schadensverursachende Ereignis und ‚e2‘ für das Ereignis des Schadens selbst zugeordnet.

Die Muster-DRS enthält somit vier Prädikate, denen zwei Objekte und vier Ereignisse zugeordnet sind. Damit ist formal die Regel bestimmt, die das Merkmal „Unfall“ bestimmt. Das Beweissystem kann damit auf Anforderung das Vorliegen eines Unfalls beweisen<sup>8</sup>.

*(wird fortgesetzt)*

<sup>6</sup> Zur Funktion von Variablen siehe oben B. 2. Variablen-Belegung

<sup>7</sup> siehe oben B. 4. Regelbildung

<sup>8</sup> Zu den Methoden der Beweisführung: Siehe D. 1. Das Beweisverfahren in LEX

In den juristischen Expertensystemen beginnt Anfang 1989 im elektronischen Newsletter „Informatik und Recht“ (NUA: 45612133061) eine Serie, die auch Ergebnisse eines Saarbrücker Seminars dazu (Wahlster/Rüßmann/Herberger) einbezieht.