

**PROPOSITIONALE UND ANALOGE REPRÄSENTATION  
VON EREIGNISSEN**

Bernd Neumann

FBI-HH-M-149/87

Juni 1987

Universität Hamburg  
Fachbereich Informatik  
Bodenstedtstr. 16  
2000 Hamburg 50

# Propositional and Analogical Representation of Events

Bernd Neumann

Fachbereich Informatik, Universität Hamburg  
Bodenstedtstr. 16, 2000 Hamburg 50

## Abstract

This report deals with computational representations of events. We propose representations which are particularly suited for events in the visual world where object motion plays a significant part. The first part of the report deals with propositional event models. They support an effective event recognition strategy by means of a hierarchical structure and an algebraic treatment of temporal relations. In the second part we present analogical event models complementary to the propositional ones. An analogical event model supports event visualization using a TAF (trajectory accumulation frame) which contains traces of experienced events in a compressed and generalized form.

# Propositionale und analoge Repräsentation von Ereignissen

Bernd Neumann

Fachbereich Informatik, Universität Hamburg

Bodenstedtstr. 16, 2000 Hamburg 50

## Zusammenfassung

Dieser Bericht behandelt die computerinterne Repräsentation von Ereignissen, also von zeitlichen Vorgängen, die als sinnvolle Einheit verstanden werden. Die dafür vorgestellten Lösungen orientieren sich an der Struktur von Ereignissen in der sichtbaren Welt, insbesondere an Objektbewegungen. Im ersten Teil werden propositionale Ereignismodelle beschrieben, die eine effektive Ereigniserkennung aufgrund ihrer hierarchischen Struktur und einer algebraischen Behandlung von Zeitbeziehungen erlauben. Im zweiten Teil wird eine analoge Repräsentationsform vorgestellt, die komplementär zur propositionalen ist. Sie unterstützt den zur Ereigniserkennung inversen Prozeß, die Visualisierung von Ereignissen. Als Ereignismodell wird ein 'Trajektorien Akkumulierungsfeld' (TAF) vorgeschlagen, in dem die möglichen räumlich-zeitlichen Ausprägungen eines Ereigniskonzeptes in Form von komprimierten und generalisierten Erfahrungen vorliegen.

## 1. Einleitung

Zeit ist ein im menschlichen Denken fest verankertes Konzept. Wir nehmen unsere Umwelt in Gestalt von zeitveränderlichen Phänomenen wahr, wir sprechen mühelos über Ereignisse und Vorgänge, wir folgern aus zeitveränderlichen Fakten, wir passen unser Wissen sich verändernden Verhältnissen an, kurz: wir scheinen mühelos alle zeitbezogenen Probleme zu meistern, mit denen heutige KI-Systeme Schwierigkeiten haben. In der Tat wurden temporale Aspekte von KI-Forschern bis vor wenigen Jahren nur sehr nebenbei behandelt. Dies gilt sowohl für anwendungsorientierte Arbeiten als auch für Grundlagenforschung, etwa im Bereich Wissensrepräsentation. Waren in einem Anwendungssystem zeitveränderliche Vorgänge zu repräsentieren, so verwendete man häufig Folgen von Zustandsbeschreibungen, deren innerer Zusammenhang durch die Repräsentationsform in unzureichender Weise wiedergegeben wurde. Beispielsweise erhielten Objekte im Expertensystem MOLGEN mehrere Namen, für jeden Zustand einen anderen. Dadurch wurde die Identität von Objekten in verschiedenen Zuständen verschleiert, fehleranfällige Programmierung und überflüssige Verwaltungsprozesse waren die Folge. Die Programmiersprache PLANNER und die in ihr programmierte Planung von Aktionsfolgen in der Blockswelt bietet ein weiteres typisches Beispiel. In PLANNER übernimmt der Assoziativspeicher die unglückliche Doppelrolle einer Faktenbasis und einer Hypothesenverwaltung. Das führt dazu, daß Veränderungen auf Grund von Realzeitvorgängen (ein Klotz ist bewegt worden) konzeptuell nicht von hypothetischen Veränderungen während der Planung (angenommen, der Klotz würde bewegt) unterschieden werden können.

Die Forderung nach einer konzeptuellen Basis ist keineswegs lediglich eine akademische Marotte, sondern eine wichtige Voraussetzung für Anwendungen. Sie wird umso dringlicher, je komplexer die zu lösenden Probleme sind. MCDERMOTT 82 gibt einen Vorgeschmack von den verwickelten Abläufen, die bei unsauberer Behandlung zeitlicher Aspekte entstehen können. Er skizziert eine dramatische Situation,

in der ein KI-System eine an die Schienen gefesselte schöne Frau vor einem herannahenden Eisenbahnzug retten möchte. Das KI-System erkennt folgerichtig, daß es die Frau von den Schienen entfernen muß und konstruiert einen Aktionsplan. Währenddessen überwacht eine Konsistenzkomponente seine Wissensbasis. Sobald klar wird, daß ein Rettungsplan ausgeführt werden soll und erfolgreich sein wird, schließt diese Komponente, daß die schöne Frau nicht mehr in Gefahr ist und beseitigt damit den Anlaß für den Rettungsplan. Der Plan wird aufgegeben und die schöne Frau ist folgerichtig wieder in Gefahr. Von dieser Stelle an wiederholt sich der Vorgang so oft, bis die Notwendigkeit einer Rettung nicht mehr besteht. Offenbar fehlt dem System das rechte Verständnis von Realzeit versus Problemlösungszeit.

McDermott entwickelt in demselben Bericht eine umfassende Zeitlogik mit dem Ziel, gesichert mit Fakten, Ereignissen, Plänen und chronologischen Abläufen in KI-Programmen umgehen zu können. Bemerkenswert ist die beträchtliche Komplexität des vorgeschlagenen Formalismus im Vergleich zur scheinbar formal unbefrachteten Vorgehensweise des Menschen. McDermott braucht immerhin 44 Axiome, so daß bereits vor der Verwendung seiner Theorie Zweifel an ihrer Konsistenz angebracht sind. Vermutlich hat die unplausible Komplexität dieses Ansatzes und auch anderer Ansätze (BRUCE 72, ALLEN 82) dazu geführt, daß Zeitlogiken in anwendungsbezogenen KI-Systemen kaum verwendet worden sind. Auf eine neuere Arbeit von Allen und Hayes (ALLEN und HAYES 85), die unseren Intuitionen eher entspricht, wird am Ende des zweiten Abschnittes näher eingegangen. Für Interessenten an weiteren formal-logischen Ansätzen sei auf eine ausführliche Übersicht in TURNER 84 verwiesen.

In diesem Bericht werden Verfahren zum Beschreiben zeitveränderlicher Phänomene aufgezeigt, deren Ausgangspunkt nicht in einer axiomatischen Basis liegt sondern in der Welt des Gegenständlichen, wie sie vom Menschen wahrgenommen wird. Die Verfahren sind also nicht formal-logisch basiert. Sie erhalten ihre Konsistenz aus den

inhärenten Gesetzmäßigkeiten der zu repräsentierenden Realwelt, sie sind also extensional orientiert. Wir gehen davon aus, daß die zu repräsentierenden raum-zeitlichen Objekte Ereignisse sind, wie sie z.B. in MILLER und JOHNSON-LAIRD 76 definiert werden. Ereignisse sind im wesentlichen Ausschnitte aus der als vierdimensional aufgefaßten Realwelt. Im ersten Teil des Berichtes entwickeln wir propositionale Ereignismodelle, deren hierarchische Struktur sich besonders zum Erkennen von Ereignissen in einer Realweltszene eignet. Primitive der Ereignisrepräsentation basieren auf zeitveränderlichen perzeptuellen Merkmalen, also beispielsweise Position und Orientierung eines Gegenstands in einem Zeitintervall. Damit werden zunächst nur sichtbare Vorgänge erfaßt. Wir werden aber argumentieren, daß dieses Konzept auch auf verallgemeinerte Vorgänge mit möglicherweise abstrakten Merkmalen ausgedehnt werden kann, z.B. eine Geldanlage.

Der zweite Teil dieses Berichtes befaßt sich mit der Repräsentation von Erfahrungswissen zur Visualisierung von Ereignissen, wiederum im Kontext von Vorgängen im Gegenständlichen. Die Repräsentationsanforderungen gehen in verschiedener Hinsicht über die im ersten Teil diskutierten Möglichkeiten hinaus. Beispielsweise ist es für Erfahrungswissen typisch, daß zwischen wahrscheinlichen und unwahrscheinlichen Vorgängen unterschieden wird. Erfahrungswissen erlaubt es, Vorerwartungen über zeitliche Vorgänge zu erzeugen und spielt offenbar bei einer Vielzahl von kognitiven Prozessen eine wichtige Rolle, z.B. beim Visualisieren sprachlich beschriebener Vorgänge oder bei Pfadplanungsproblemen in der Robotik. Die hier vorgeschlagene analoge Repräsentationsform ist in gewisser Hinsicht komplementär zu der propositionalen Ereignisrepräsentation des ersten Teils: Sie unterstützt andere kognitive Prozesse und sollte daher eher als eine Ergänzung denn eine Alternative aufgefaßt werden. Dabei wird deutlich, daß die Wahl einer Repräsentationsform für zeitliche Vorgänge im Hinblick auf die jeweilige Aufgabe erfolgen muß.

## 2. Propositionale Ereignismodelle für zeitveränderliche Szenen

In diesem Abschnitt wird eine propositionale Form der Ereignisrepräsentation beschrieben, deren Primitive perzeptueller Natur sind und im Bildverstehen wurzeln, deren Ausdruckskraft aber sprachlichen Konzepten entspricht, z.B. Bewegungsverbren wie "anhalten" oder "abbiegen". Die Ergebnisse gehen im wesentlichen auf Arbeiten zusammen mit H.-J. Novak im Rahmen von Projekt NAOS zurück (NEUMANN und NOVAK 83, NEUMANN und NOVAK 86). Ziel des Projektes war es, eine in Gestalt einer Bildfolge vorgegebene Verkehrsszene automatisch sprachlich zu beschreiben.

Das Erkennen eines Ereignisses kann als die Transformation einer quantitativen in eine qualitative Szenenbeschreibung aufgefaßt werden. Dabei spielen verschiedene konzeptuelle Zwischenebenen eine Rolle (s. Abbildung 1).

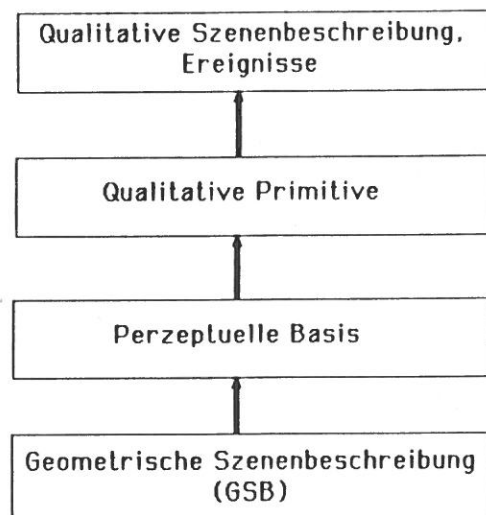


Abbildung 1: Repräsentationsebenen zur Ereigniserkennung

Die quantitative Szenenbeschreibung besteht aus Daten über die räumlich-zeitliche Lage der in der Szene enthaltenen Objekte, sowie

ihr Aussehen und ihre Klassenzugehörigkeit. Diese Informationen stellen eine kanonische Szenenbeschreibung dar, wie sie beispielsweise durch ein Bildanalyse-System gewonnen werden könnte. In NEUMANN UND NOVAK 83 wurde dafür die Bezeichnung "Geometrische Szenenbeschreibung" (GSB) eingeführt. Bei einer Szene mit bewegten Objekten enthält die GSB insbesondere Informationen über Objekttrajektorien in räumlichen und zeitlichen Koordinaten. Diese könnte z.B. in folgender Form repräsentiert werden:

(LAGE <Objektbezeichner> <xyz-Koordin.> <Orient.-Vektor> <Zeitpunkt>)

z.B.                   (LAGE VW1 (93 158 0) (.9912 -.1322 .0) 18)  
                         (LAGE VW1 (77 159 0) (.9981 -.0624 .0) 19)  
                         (LAGE VW1 (60 159 0) (1.0000 .0000 .0) 20)  
                         etc.

Die relationalen Tupel beschreiben Schwerpunktslage und Orientierung eines Objektes für fortlaufende Zeitpunkte. Objekteigenschaften wie Form und Farbe sind in objektzentrierten Schemata enthalten, auf die hier nicht weiter eingegangen wird.

Vollständige Angaben über die exakte räumlich-zeitliche Lage aller Objekte einer Szene stehen wohl weder in technischen noch biologischen Systemen zur Verfügung. Dennoch ist die GSB eine realistische Basis für abgeleitete Beschreibungen. Die in ihr enthaltene Information - eine möglichst genaue quantitative Beschreibung der sichtbaren Umgebung - ist Ziel und Voraussetzung für zahlreiche kognitive Prozesse, z.B. Pfadplanung, Visualisierung, räumliches Schließen. Dies gilt sowohl für den Menschen (soweit man seine kognitive Architektur versteht) als auch für KI-Systeme.

Wir betrachten nun Szenenausschnitte (repräsentiert durch eine GSB), die aufgrund charakteristischer Eigenschaften ein "Ereignis" darstellen. Es gibt keine rein formalen Kriterien dafür, was ein Ereignis ausmacht. Wir orientieren uns hier an der natürlichen



Sprache, die mit ihren Begriffen widerspiegelt, welche Szenenausschnitte als sinnvolle konzeptuelle Einheiten aufzufassen sind. Bei zeitlichen Vorgängen spielen die Verben eine entscheidende Rolle. Wir definieren daher ein Ereignis als einen vierdimensionalen Unterraum des Ort-Zeit-Kontinuums, der mit einem Bewegungsverb beschreibbar ist. Das Problem, Ereignismodelle zu definieren, ist damit gleichbedeutend mit der Definition einer extensionalen Semantik für Bewegungsverben. Wenn wir wissen, welche Szeneneigenschaften beispielsweise ein "überholen" ausmachen, haben wir sowohl die Semantik von "überholen" definiert als auch den Weg für die Repräsentation eines entsprechenden Ereignismodells gewiesen.

Bei der Untersuchung von Bewegungsverben für Straßenverkehrsszenen zeigte sich, daß die durch diese Verben definierten Ereignisse eine hierarchisches System bilden. Komplexe Ereignisse wie z.B. "überholen" implizieren einfachere Ereignisse, z.B. "sich nähern" und "sich entfernen". Eine ausführliche Diskussion dieser Verbhierarchie findet sich in NOVAK 86. Die Hierarchie basiert auf Primitiven, die interessante Gesetzmäßigkeiten offenbaren.

- Ereignisprimitive sind durative Prädikate. Sie beschreiben Zeitintervalle, in denen bestimmte Szeneneigenschaften ununterbrochen gültig sind.
- Ereignisprimitive sind qualitative Prädikate über beobachtbaren Größen, die durch die GSB quantitativ vorgegeben sind.

Als beobachtbare Größen kommen nur bestimmte Szeneneigenschaften in Frage. Dies sind relative Position und Orientierung eines Objektes bezüglich geeigneter Referenzobjekte, sowie die zugehörigen zeitlichen Ableitungen, jeweils als Funktion der Zeit. Diese Größen, die bei einer Szene mit vielen Objekten in beträchtlicher Zahl anfallen, stellen die perzeptuelle Basis einer qualitativen Szenenbeschreibung dar (s. Abbildung 1).

Ereignisprimitive erfassen verschiedene Arten von Konstantheiten, die geeignet sind, den zeitlichen Verlauf der beobachteten Größen zu charakterisieren:

- konstanter Wert  
z.B. Stillstand, Geradeausbewegung, konstante Geschwindigkeit
- eingeschränkter Wertebereich  
z.B. parallel, nahe, neben, auf
- größer/kleiner  
z.B. Komparative, ungewöhnliche Werte (rasen)
- monotoner Verlauf  
z.B. beschleunigen, drehen, sich nähern

Für den Aufbau der ca. 50 Verben umfassenden Verbhierarchie sind 19 derartige Ereignisprimitive verwendet worden. Eine vollständige Liste findet sich in NEUMANN und NOVAK 86. Wir zeigen nun, wie ein komplexes Ereignis aus Komponenten (und letztendlich aus Primitiven) aufgebaut wird. Wir betrachten dazu Ereignismodelle, also generische Ereignisbeschreibungen, die jeweils eine Klasse von Ereignissen definieren. Ein primitives Ereignismodell, z.B. für eine Bewegung, läßt sich wie folgt notieren:

(BEWEGEN OBJ T1 T2)

T1 und T2 beschreiben Anfangs- und Endpunkt eines Zeitintervalls, währenddessen sich ein Objekt OBJ bewegt. Die Durativität impliziert die Gültigkeit des Ereignisses in allen Teilintervallen:

(BEWEGEN OBJ T1 T2) --> (BEWEGEN OBJ T1' T2')  
T1 =< T1' < T2' =< T2

Zwei Ereignisse werden zusammengesetzt, indem ihre relative zeitliche Lage spezifiziert wird. Im einfachsten Fall sind die Ereignisse gleichzeitig:

(E1 ... T1 T2)

(E2 ... T1 T2)

Oder sie folgen einander:

(E1 ... T1 T2)

(E2 ... T2 T3)

Durch die Verwendung identischer Zeitmarken kann die Koinzidenz von Intervallgrenzen ausgedrückt werden. Fast alle für die Beschreibung von Verkehrsszenen erforderlichen Ereignisse sind so darstellbar. Das folgende ist ein Ereignismodell für "überholen":

(UEBERHOLEN OBJ1 OBJ2 T1 T2) <-->  
    (BEWEGEN OBJ1 T1 T2)  
    (BEWEGEN OBJ2 T1 T2)  
    (NAEHERN OBJ1 OBJ2 T2 T3)  
    (HINTER OBJ1 OBJ2 T2 T3)  
    (NEBEN OBJ1 OBJ2 T3 T4)  
    (ENTFERNEN OBJ1 OBJ2 T4 T2)  
    (VOR OBJ1 OBJ2 T4 T2)

Um auch ein "vor" oder "während" von Zeitintervallen ausdrücken zu können, wird das degenerierte primitive Ereignis D ("dauern") eingeführt. Es beschreibt verstrichene Zeit, ohne eine Aussage über die Szene zu machen. Ein Ereignis, das (nicht unmittelbar) vor einem zweiten stattfindet, wird dann wie folgt spezifiziert:

(E1 ... T1 T2)

(D T2 T3)

(E2 ... T3 T4)

Wir betrachten im folgenden die Intervallgrenzen (T1 T2) eines Ereignisses (E1 ... T1 T2) als Referenzzeitmarken und untersuchen, in welchen zeitlichen Beziehungen der Anfangszeitpunkt T3 eines

zweiten Ereignisses (E2 ... T3 T4) zu den Referenzzeitmarken stehen kann. Durch Koinzidenz und durch Anbindung über D-Ereignisse ergeben sich acht verschiedene Möglichkeiten, die in Abbildung 2 gezeigt werden. Links ist die jeweilige Festlegung von T3 durch Kreuz oder durchgezogene Linie auf einer Zeitachse mit den Referenzzeitmarken T1 und T2 dargestellt. Rechts steht die dazugehörige Notation für das Ereignis E2.

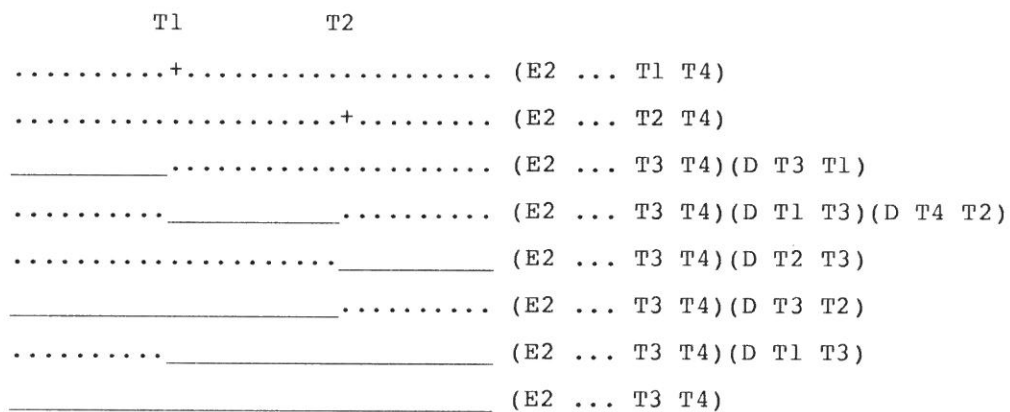


Abbildung 2: Acht zeitliche Beziehungen zwischen dem Startzeitpunkt von E2 und einem Ereignis (E1 ... T1 T2)

Kombiniert man die Möglichkeiten für den Startzeitpunkt mit denen für den Endzeitpunkt eines Intervalls, so ergeben sich insgesamt 27 mögliche Zeitbeziehungen zwischen zwei Ereignissen. Dies wird weiter unten noch im Hinblick auf die 13 Zeitbeziehungen in ALLEN 81 kommentiert.

Ein zusammengesetztes Ereignis ist in der Regel nicht mehr durativ, d.h. es trifft nicht gleichermaßen für alle Teilintervalle seines Gültigkeitsbereichs zu. Man kann sehen, daß die Intervallgrenzen (T1 T2) eines zusammengesetzten Ereignisses einem System linearer Ungleichungen genügen, die von den durativen Primitiven herrühren,

aus denen das Ereignis aufgebaut ist. Eine Lösung hat stets die Form:

$$\begin{aligned} T_{1\min} &= < T_1 = < T_{1\max} \\ T_{2\min} &= < T_2 = < T_{2\max} \\ T_1 &< T_2 \end{aligned}$$

Anfangs- und Endzeitpunkt einer Ereignisinstanz liegen also im allgemeinen in individuell beschränkten Zeitintervallen. Ein konkretes Überholereignis, beispielsweise, kann durch eine Proposition wie diese beschrieben werden:

(UEBERHOLEN BMW1 VW1 (13 19) (54 58))

Das Intervall (13 19) beschreibt die möglichen Werte für  $T_1$ , das Intervall die möglichen Werte (54 58) für  $T_2$ . Inchoative Ereignisse (z.B. "losfahren") oder resultative Ereignisse (z.B. "anhalten") sind Sonderfälle, bei denen das Start- bzw. Endintervall punktförmig ist. Ein duratives Ereignis ist aus dieser Sicht auch ein Sonderfall: Start- und Endzeitpunkt unterliegen denselben Beschränkungen.

Die so definierten Ereignismodelle ermöglichen eine effektive Ereigniserkennung. Das Verfahren basiert darauf, die sich durch Teilinstantiierungen akkumulierenden Beschränkungen für Zeitmarken in einem Beschränkungsnetz zu propagieren, so daß der Gültigkeitsbereich einer Zeitmarke zu jeder Zeit ersichtlich ist und Sackgassen bei der Suche frühzeitig erkannt werden können (Details s. NOVAK 85). Darüberhinaus unterstützt die Ereignisstruktur eine hierarchische Suche, deren Effektivität bei einem Strukturvergleich seit langem bekannt ist (BARROW et al. 72).

Wir kehren nun zu den in der Einführung angesprochenen Zeitlogiken zurück und versuchen, einige Unterschiede zwischen jenen und dem hier vorgestellten Ansatz zu beleuchten. Zunächst ist festzustellen, daß unsere Ereignisse strikt intervallbasiert sind. Dies ist keine

zufällige formale Entscheidung, sondern trägt zum einen der Vorstellung Rechnung, daß Ereignisse einen endlichen Unterraum des vierdimensionalen Ort-Zeit-Kontinuums einnehmen. Zum anderen sind die hier gewählten Ereignisprimitive aus verschiedenen Gründen naheliegend, nicht zuletzt mit dem Blick auf menschliche Perzeption. Sie sind intervallbasiert und ergeben durch die beschriebenen Verknüpfungsmöglichkeiten intervallbasierte komplexe Ereignisse. Wir befinden uns dabei in Übereinstimmung mit den Ansätzen von McDermott (MCDERMOTT 82) und Allen (ALLEN 81, ALLEN und HAYES 85), die in der KI bisher den nachhaltigsten Einfluß gehabt haben.

Interessant ist der Vergleich mit ALLEN und HAYES 85. Dort wird eine Zeitlogik mithilfe einer einzigen axiomatischen Beziehung zwischen Zeitintervallen entwickelt, der unmittelbaren Folge (MEETS). Dies entspricht der hier verwendeten Kompositionstechnik durch koinzidierende Intervallgrenzen. Die Autoren heben hervor, daß alle 13 Zeitrelationen zwischen zwei durativen Ereignissen, die in ALLEN 82 noch einzeln eingeführt werden mußten, mithilfe der MEETS-Beziehung definiert werden können. Unsere Analyse ergab, daß auch nicht-durative Ereignisse als Ergebnis einer konjunktiven Verknüpfung von durativen Komponenten entstehen können. Bei dieser allgemeineren Ereignisform können Anfangs- und Endzeitpunkt verschiedene Zeitbereiche einnehmen. Daher ergeben sich 27 statt 13 Zeitrelationen, die alle durch die Spezifikation koinzidierender Intervallgrenzen dargestellt werden können. Wir meinen, daß diese allgemeinere Ereignisdefinition sinnvoll ist, denn sie folgt zwangsläufig aus der Verknüpfung von durativen Primitiven und deckt sich auch mit sprachlichen Ereigniskonzepten (z.B. "überholen").

Ein weiterer Vergleichspunkt ist die Ereigniserkennung. Der logisch-axiomatische Ansatz erfordert deduktive Methoden, um ein Ereignis in einer Faktenbasis zu verifizieren. Hier wird dagegen von einer reellwertigen Zeitachse ausgegangen, auf der Intervallgrenzen algebraisch berechnet werden. Obwohl ein direkter Vergleich nicht durchgeführt wurde, halten wir den algebraischen Ansatz für

überlegen, weil durch ihn die inhärenten Eigenschaften der Zeitachse besser ausgenutzt werden.

Die letzten vergleichenden Anmerkungen gelten punktuellen Ereignissen. Ohne Zweifel werden Ereignisse im natürlichen Sprachgebrauch und in unserem Denken häufig wie Zeitpunkte behandelt. Man betrachte als Beispiel "als es knallte, lief er los". Hier sind Knall und Loslaufen offenbar punktuelle Ereignisse. Es muß jedoch nicht wundern, wenn ein an solchem Sprachgebrauch orientiertes formales System kompliziert oder inkonsistent wird. Offenbar kann ein Knall in einem anderen Kontext als zeitlich ausgedehnter Vorgang angesehen werden, was zu einem Konflikt mit der anfänglichen Typisierung führt. Andererseits ist der Zeitpunkt des Loslaufens als Grenze zwischen Stehen und Laufen auch bei feinerer zeitlicher Auflösung konzeptuell eine Grenze, also ein Zeitpunkt. Allenfalls erhält er dadurch Intervallcharakter, daß er sich schwer definieren oder feststellen läßt. In ALLEN und HAYES 85 werden deshalb auch zwei Arten von Zeitpunkten unterschieden, solche die Punkt ereignissen und solche die Ereignisgrenzen entsprechen. Interessanterweise weisen die Autoren nach, daß Punkt ereignisse nicht mit einer reellwertigen Zeitachse konsistent sind. Wir meinen, daß Punkt ereignisse überhaupt nicht in die axiomatische Basis einer Zeitlogik gehören, sondern im Zusammenhang mit approximativen Verarbeitungsformen und informellem Sprachgebrauch eingeführt werden sollten. Zeitpunkte als Ereignisgrenzen sind jedoch ein natürlicher Bestandteil von intervallbasierten Formalismen.

Zum Abschluß dieses Abschnittes soll diskutiert werden, inwieweit sich das hier vorgestellte, für eine perzeptuelle Basis entwickelte Ereigniskonzept auf beliebige Vorgänge mit Ereignischarakter übertragen läßt. Formal gesehen spielt es offenbar keine Rolle, ob die Ereignisprimitive perzeptueller Natur sind. Entscheidend ist, daß sich die charakteristischen Eigenschaften von Vorgängen in der betrachteten Domäne auf durative Primitive zurückführen lassen. Dies scheint nicht ausgeschlossen, wie der folgende Versuch eines

Ereignismodells für "kaufen" zeigt:

(KAUFEN AGENT WARE GELD T1 T2) <-->  
(BESITZT AGENT GELD T1 T3)  
(GIBT AGENT GELD T1 T3)  
(NIMMT AGENT WARE T3 T2)  
(BESITZT AGENT WARE T3 T2)

Das Problem dabei ist offenbar, geeignete semantische Primitive zu finden. Im Gegensatz zur Domäne des Gegenständlichen konnten für allgemeinere Domänen bisher keine zufriedenstellenden Primitive gefunden werden. Die konzeptuelle Dependenztheorie (SCHANK 73), beispielsweise, ist für viele Zwecke nicht genügend differenziert. Vorausgesetzt, eine geeignete Basis von durativen Primitiven wäre vorhanden, dann könnten damit Ereignismodelle in der beschriebenen Weise aufgebaut und Ereignisse effektiv erkannt werden.



### 3. Ereignismodelle zur Visualisierung von Ereignissen

Die im vorhergehenden Abschnitt diskutierten propositionalen Ereignismodelle eignen sich besonders gut zur Ereigniserkennung, beispielsweise zum Zweck der Transformation einer geometrischen (quantitativen) in eine sprachliche (qualitative) Szenenbeschreibung. Für den umgekehrten Vorgang, der Transformation einer sprachlichen Beschreibung in eine räumlich-zeitliche Darstellung, werden Ereignismodelle mit anderen Eigenschaften gebraucht. In diesem Abschnitt wird eine Repräsentationsform vorgestellt, die dafür geeignet ist. Sie wurde zusammen mit M. Mohnhaupt entwickelt (MOHNHAUPT und NEUMANN 87), um ein bildverstehendes System sprachlich steuern zu können.

Im Gegensatz zur Ereigniserkennung haben wir es jetzt mit einer prinzipiell mehrdeutigen Transformation zu tun: Zu einer sprachlichen Beschreibung, z.B. "Ein gelber VW biegt in die Bieberstraße ein", passen im allgemeinen unendlich viele verschiedene räumlich-zeitliche Vorgänge. Wir konzentrieren uns im folgenden auf die durch ein Bewegungsverb charakterisierte Trajektorie und beschreiben ein Ereignismodell, das die möglichen Ausprägungen dieser Trajektorie repräsentiert. Die Leitidee besteht darin, ein Ereignismodell aus Einzelbeobachtungen erlernbar zu machen. Wir gehen also davon aus, daß das Wissen über die zu einem Verb passenden Trajektorien auf Grund von Erfahrungen entsteht. Darüberhinaus sollen die Ereignismodelle charakteristische Eigenschaften von Visualisierungen unterstützen. Diese sind:

#### a) Explizite Topologie

Visualisierungen sollen die räumlichen und zeitlichen Beziehungen eines Ereignisses explizit machen. Dazu sind analoge Repräsentationsformen grundsätzlich besser geeignet als propositionale.

b) Quantitativer Charakter

Im Gegensatz zu einer propositionalen, qualitativen Beschreibung sollen Visualisierungen quantitative, geometrische Informationen liefern. Sie sollen also den Charakter einer GSB haben. Beim Menschen stehen Visualisierungen hinsichtlich Form und Repräsentationsmedium in der Tat in engem Zusammenhang mit tatsächlich wahrgenommene Szenen (FINKE 85).

c) Unschärfe

Die mit einem sprachlichen Ereignisbegriff verbundenen räumlich-zeitlichen Vorstellungen sind in der Regel unscharf, d.h. sie spezifizieren Objekttrajektorien (und andere Situationsmerkmale) nicht exakt sondern lassen Alternativen zu. Andererseits kann es durchaus eindeutige Merkmale geben, z.B. die Bewegungsrichtung.

d) Adaptierbarkeit

Visualisierungen müssen an konkrete Randbedingungen angepaßt werden. Ein Ereignismodell wird in der Regel von Situationsdetails abstrahieren, z.B. von gelegentlichen Hindernissen, denen ausgewichen werden muß. Visualisierte Trajektorien müssen deshalb erst an die im konkreten Fall vorliegenden Hindernisse angepaßt werden.

Wir haben für diese Zwecke ein 'Trajektorien-Akkumulierungsfeld' (TAF) entwickelt. Ein TAF ist ein mehrdimensionales Zählerfeld für Zustandsvektoren zur Beschreibung von Objektbewegungen. Wir betrachten im folgenden ein vierdimensionales Zählerfeld  $C(x,z,r,b)$ , das ein bestimmtes Gebiet der  $xy$ -Ebene abdeckt, z.B. eine bestimmte Kreuzung. Jede Zelle des vierdimensionalen Feldes repräsentiert eine bestimmte Kombination von:

$xy$	Ort
$r$	Geschwindigkeitsrichtung
$b$	Geschwindigkeitsbetrag

Die Zähler enthalten die Spuren von beobachteten Trajektorien einer

(fiktiven) Lernphase und repräsentieren damit die Wahrscheinlichkeit der ihnen zugeordneten Zustandsvektoren.

Die Visualisierung einer Trajektorie erfolgt im Prinzip dadurch, daß zu einer Startzelle die wahrscheinlichste Fortsetzung aus dem TAF entnommen wird: Dabei muß die Fortsetzung in einer qualifizierten Nachbarschaft, einem Kontinuitätskriterium entsprechend, gesucht werden.

Abbildung 3a zeigt ein TAF für Abbiegevorgänge an einer bestimmten Kreuzung, projiziert auf die xy-Ebene. Das TAF enthält die Spuren von 10 "gelernten" Trajektorien (hier von Hand eingegeben). Vorhersagen folgen - wie zu erwarten - den gelernten Trajektorien (Abbildung 3b), wenn entsprechende Startwerte (schwarzer Kreise) vorgegeben sind.

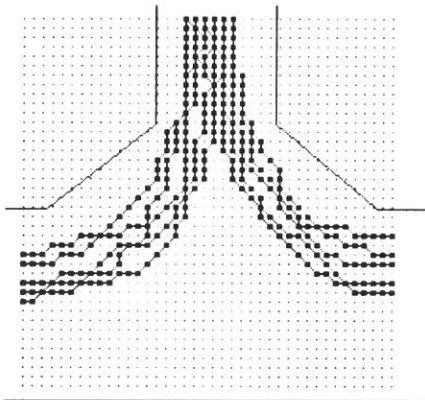


Abbildung 3a:  
TAF mit 10 Einzeltrajektorien

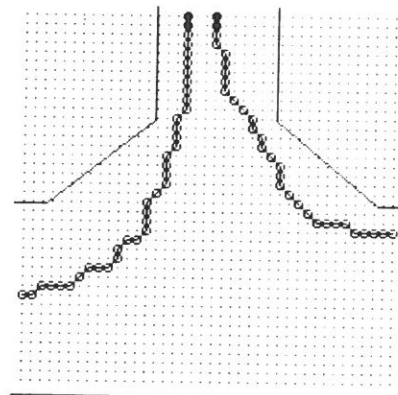


Abbildung 3b:  
Vorhersage bei vorgegebenen  
Startwerten

TAFs sollen allerdings in der Regel keine Einzeltrajektorien speichern und wiedergeben, sondern einen Erfahrungsschatz von sehr vielen Trajektorien auswerten. Ein TAF hat dann eher den Charakter

einer Häufigkeitsverteilung, und vorhergesagte Trajektorien stimmen nicht notwendigerweise mit einer beobachteten Trajektorie überein. Dies wird auch zwangsläufig der Fall sein, wenn Vorhersagen von Startwerten aus erfolgen sollen, für die überhaupt keine Erfahrungen vorliegen. Um dies zu ermöglichen, wird eine Verwaschungsoperation eingeführt, die Erfahrungen in bisher unbesetzte TAF-Bereiche propagiert.

### Verwaschung

Zu jeder Zelle  $(x,y,r,b)$  wird der gewichtete Mittelwert aller Nachbarzellen in Positionen  $x',y'$  orthogonal zur Bewegungsrichtung  $r$  addiert. Nachbarzellen in Bewegungsrichtung tragen entsprechend ihrer positiven Differenz zur Zelle bei. (Die leichte Unsymmetrie verhindert eine Erosion in Außenkurven und ist nicht wesentlich.) Die Verwaschung hat einen Generalisierungseffekt in dem Sinn, daß das Zählerfeld zusätzlich zu den ursprünglichen Trajektorien auch Nachbartrajektorien unterstützt, die ungefähr dazu parallel verlaufen. Die Abbildungen 4a und 4b zeigen Vorhersagen nach Verwaschen des in Abbildung 3a gezeigten TAF. Wegen der groben Richtungsdiskretisierung (45 Grad Schritte) wurde nur bezüglich  $x$ ,  $y$  und  $d$  verwaschen.

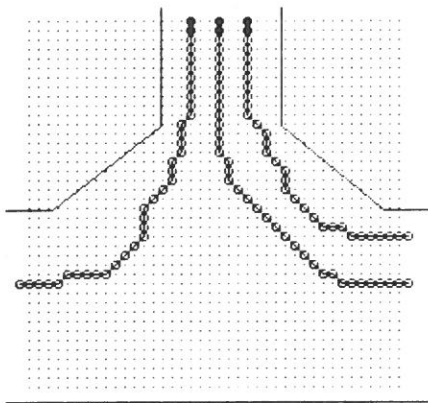


Abbildung 4a:

Vorhersagen nach einmaliger  
Verwaschung

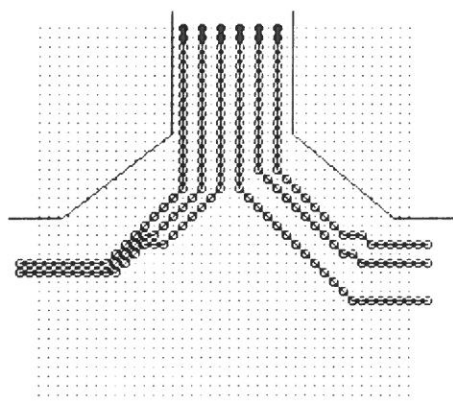


Abbildung 4b:

Vorhersagen nach dreimaliger  
Verwaschung

Es ist bemerkenswert, daß die Generalisierung auf Grund der analogen Repräsentationsform durch einfache, lokale Operationen erzielt werden kann. Bei einer propositionalen Repräsentation - z.B. bei einem der in Abschnitt 2 beschriebenen Ereignismodelle - wäre ein Prozeß, der die Gültigkeit des Modells erweitert, äußerst komplex.

Statt der expliziten Repräsentation einer einzigen "typischen" Trajektorie bietet ein TAF ein volles Spektrum von typischen bis weniger typischen Trajektorien. Dies hat den offensichtlichen Vorteil, daß differenzierte Prädiktionen für eine Vielzahl unterschiedlicher Trajektorienvorgaben generiert werden können. In Abwesenheit konkreter Vorgaben kann es andererseits nützlich sein, einen einzigen oder sehr wenige typische Verläufe als charakteristisch für den TAF herauszustellen, z.B. um typische Eigenschaften daraus ermitteln zu können. Wir bezeichnen solche typischen Verläufe als Ereignisprototypen. Sie können aus dem TAF durch Anwendung einer Konvergenzoperation gewonnen werden, die im wesentlichen komplementär zur Verwaschungsoperation ist.

#### Konvergenz

Die Konvergenzoperation besteht darin, daß jede Zelle die zu ihr hinführenden Nachbarzellen proportional zu ihrem eigenen Zählerstand verstärkt. Dadurch entsteht die Tendenz, daß Pfade von beliebigen Ausgangspunkten aus in einen stark ausgeprägten Pfad (einen Ereignisprototypen) einmünden, also konvergieren. Dies illustriert Abbildung 5.

Das linke Bild zeigt den Konvergenz-Effekt bei Prädiktionen, die von verschiedenen Startpunkten ausgehen. Im rechten Bild sind alle Pfade entlang lokaler TAF-Maxima markiert, also diejenigen Pfade, zu denen andere Pfade konvergieren. Sie charakterisieren "typisches Verhalten" und können z.B. herangezogen werden, wenn Prädiktionen ohne konkrete Vorgaben gemacht werden sollen. Die TAF-Maxima stellen eine reduzierte Form der im TAF enthaltenen Informationen dar, gewissermaßen sein "Skelett".

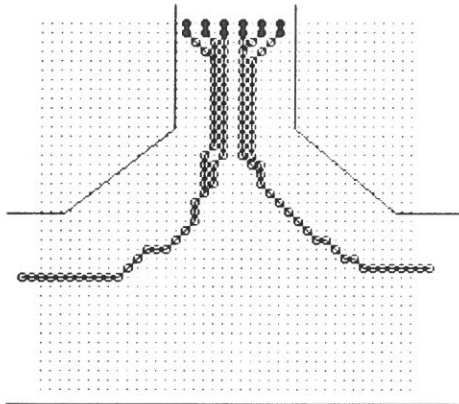


Abbildung 5a:  
Vorhersagen nach dreimaligem Ver-  
waschen und einmaliger Anwendung  
der Konvergenzoperation

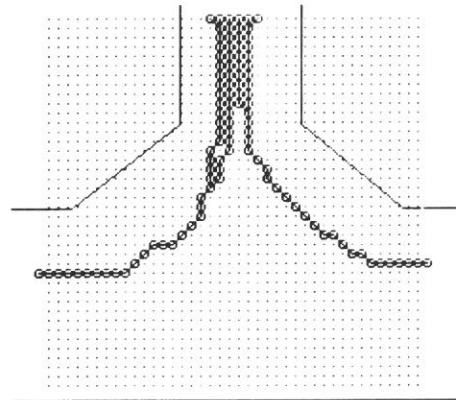


Abbildung 5b:  
Pfade entlang lokaler Maxima  
stellen Ereignisprototypen dar

Um ein TAF zur Visualisierung von Objekttrajektorien in verschiedenen Situationen einsetzen zu können, insbesondere auch in einer anderen Umgebung als in der Lernphase, müssen weitere Abstraktionen und Transformationen möglich sein. Beispielsweise sollte ein TAF für Abbiegeverhalten auf Straßenabzweigungen mit unterschiedlicher Geometrie anwendbar sein. Das kann in der Tat erreicht werden, indem das TAF um geeignete Merkmale erweitert und von irrelevanten Merkmalen abstrahiert wird. Diese Thematik ist allerdings noch aktueller Forschungsgegenstand.

#### 4. Zusammenfassung

Wir haben uns in diesem Beitrag mit der Repräsentation von Ereignissen befaßt. Ereignisse sind Ausschnitte aus dem räumlich-zeitlichen Kontinuum und erfordern deshalb in erster Linie die Repräsentation zeitabhängiger räumlicher Eigenschaften von Gegenständen. Ein Ereignis ist aber auch eine begriffliche Kategorie, zu deren Repräsentation generische Ereignismodelle erforderlich sind. Es wurden zwei Repräsentationsformen für Ereignismodelle gegenübergestellt, propositionale hierarchische Modelle und analoge Modelle. Die propositionalen Modelle sind aus qualitativen Prädikaten über zeitabhängigen beobachtbaren Größen aufgebaut und erlauben eine effektive Ereigniserkennung. Ereignisinstanzen erstrecken sich über Zeitintervalle, deren Grenzen in der Regel nicht genau festliegen, sondern durch Ungleichungssysteme eingeschränkt werden.

Als analoge Repräsentationsform für Ereignismodelle wurden TAFs vorgestellt, die beobachtete Trajektorien in akkumulierter und komprimierter Form enthalten. Im Gegensatz zu den propositionalen Modellen bieten TAFs die Möglichkeit, quantitative geometrische Trajektorienverläufe aus generischen Ereignismodellen abzuleiten, so daß Visualisierungen und Aussagen über typisches Verhalten generiert werden können. TAFs machen die räumlichen Ereignisdimensionen explizit und lassen sich - anders als propositionale Ereignismodelle - nicht unmittelbar auf abstrakte Situationen übertragen. Die jeweiligen Vor- und Nachteile der gegenübergestellten Repräsentationsformen legen es nahe, beide Formen komplementär zu verwenden. Dabei kommt den TAFs als Erfahrungsträgern eine primäre Rolle zu. Über das Zusammenspiel beider Repräsentationsformen können wohl erst dann weitergehende Einsichten gewonnen werden, wenn es KI-Systeme gibt, in denen heterogene kognitive Aufgaben integriert behandelt werden, z.B. in zukünftigen Robotersystemen.

## 5. Literaturverzeichnis

### **Allen 81**

A General Model of Action and Time

J.F. Allen

Technical Report 97, University of Rochester, Rochester/NY 1981

### **Allen und Hayes 85**

A Common-Sense Theory of Time

J.F. Allen, P.J. Hayes

Proceedings IJCAI-85, 1985, 528-531

### **Barrow et al. 72**

Some techniques for recognizing structures in pictures

H.G. Barrow, A.P. Ambler, R.M. Burstall

In: Watanabe, S. (Hrsg.) Frontiers of pattern recognition,

New York: Academic Press 1972, 1-29

### **Bruce 72**

A Model for Temporal References and Its Application in a Question  
Answering Program

B.C. Bruce

Artificial Intelligence 3, 1972, 1-25

### **Finke 85**

Theories Relating Mental Imagery to Perception

R.A. Finke

Psychological Bulletin 1985, Vol. 98, Nr. 2, 236-259

### **McDermott 82**

A temporal logic for reasoning about plans and actions

D. McDermott

Cognitive Science 6, 1982, 101-155



**Mohnhaupt und Neumann 87**

Szenenhafte Modelle für zeitabhängige Ereignisse

M. Mohnhaupt, B. Neumann

Universität Hamburg, Fachbereich Informatik, FBI-HH-B-127/87, 1987

**Miller and Johnson-Laird 76**

Language and Perception

G.A. Miller, P.N. Johnson-Laird

Cambridge University Press, Cambridge-London-Melbourne 1976

**Neumann und Novak 83**

Event Models for Recognition and Natural Language Description

B. Neumann, H.-J. Novak

Proceedings of the IJCAI-83, 1983, 724-726

**Neumann und Novak 86**

NAOS: Ein System zur natürlichsprachlichen Beschreibung  
zeitveränderlicher Szenen

B. Neumann, H.-J. Novak

Informatik Forschung und Entwicklung, Nr. 1,

Springer-Verlag, Berlin Heidelberg New York Tokyo, 1986, 83-92

**Novak 85**

Textgenerierung auf der Grundlage visueller Daten: Beschreibungen  
von Straßenszenen

H.-J. Novak

Dissertation, Fachbereich Informatik, Universität Hamburg, 1986

**Schank 73**

Identifikation of conceptualizations underlying natural language

R.C. Schank

in: Schank, R.C., Colby, K.M. (Hrsg.) Computer models of thought  
and language, pp. 187-247, New York: W.A. Freeman 1973

**Turner 84**

Logics for Artificial Intelligence

R. Turner

Ellis Horwood Ltd., 1984