

# ELiS\_e

[e'li:zə]

<Essener Linguistische Skripte\_elektronisch>

*Jürgen Klüver & Christina Stoica*

**Literarische Rationalität  
und Künstliche Intelligenz:  
Eine etwas andere Form der Analyse  
von Kriminalgeschichten**

*Faszination  
Sprache*

[elise@uni-essen.de](mailto:elise@uni-essen.de)

<http://www.elise.uni-essen.de>

# Literarische Rationalität und Künstliche Intelligenz: Eine etwas andere Form der Analyse von Kriminalgeschichten

Jürgen Klüver & Christina Stoica (Duisburg-Essen)

## Vorbemerkung

Karl-Dieter Bunting ist und war ein vielseitig interessierter Wissenschaftler, dessen vielfältige Aktivitäten wir hier weder aufzählen noch insgesamt würdigen können. Der Teil seiner Arbeiten, die uns als Nichtgermanisten am meisten interessierte, ist zweifellos die intensive Beschäftigung Buntings mit den so genannten Neuen Medien, deren pädagogische Möglichkeiten er sehr früh erkannte. Beispielsweise entwickelte er mit seinen Mitarbeitern bereits Anfang der neunziger Jahre ein Hypertextsystem zur Einführung in die Linguistik. Damit zählte er nicht nur in Deutschland zu den Pionieren auf diesem nach wie vor in der Entwicklung befindlichen Gebiet.

Wir hatten freilich nicht nur bei Bunting, sondern generell bei den Protagonisten der didaktischen Verwendung von speziellen Computerprogrammen häufig den Eindruck, dass den Möglichkeiten ganz neuer Lernziele zu wenig Aufmerksamkeit bei der Entwicklung von didaktisch verwendbaren Programmen gewidmet wurde. Damit meinen wir, dass neue Medien und insbesondere neuartige Computerprogramme nicht nur (aber vor allem) dazu verwendet werden sollten, solche Lernziele zu realisieren, die mit herkömmlichen Medien nicht oder nur unvollkommen angestrebt werden können. Mit anderen Worten, neue Medien sollen nicht einfach daran gemessen und entsprechend eingesetzt werden, die Vermittlung traditioneller Lernziele und Inhalte zu effektivieren – was sie häufig gar nicht können –, sondern es muss bei ihrer Entwicklung und Verwendung darum gehen, Themen und Probleme zu vermitteln, die sich herkömmlichen Lehr-/Lernmethoden entziehen. Dies haben wir mehrfach ausgeführt (Klüver, Stoica und Schmidt 2001). Nicht erst seit PISA sollte von daher die Parole lauten: Neue Lernziele braucht das Land und *deswegen* neue Medien.

Wir haben diese Programmatik in verschiedenen Lehrkontexten und mit verschiedenen Computerprogrammen realisiert. Das im Folgenden geschilderte Beispiel demonstriert die Verwendungsmöglichkeiten spezieller Programme, die aus dem Kontext der Forschungen zur so genannten Künstlichen Intelligenz (KI) stammen. Wir hoffen, dass Karl-Dieter Bunting Vergnügen daran findet, die Verwendung von Computerprogrammen einmal etwas anders vorgeführt zu bekommen.

## 0. Der Kontext

Die hier dargestellten Experimente mit verschiedenen Künstliche Intelligenz (KI)-Modellen, nämlich einem Expertensystem, einem interaktiven neuronalen Netzwerk und einer Kohonen-Karte, sind Ergebnisse eines Seminars „Rationalität literarischer Detektive und Modelle Künstlicher Intelligenz“ im Sommersemester 2003, das wir im Studiengang Kommunikationswissenschaft an der Universität Duisburg-Essen durchgeführt haben. Ziel des Seminar war es, die Parallelen zwischen dem rationalen Vorgehen literarischer Detektive einerseits und verschiedenen Model-

len der Künstlicher Intelligenz andererseits aufzuzeigen sowie die Studierenden in die formalen Denkweisen unterschiedlicher KI-Modelle einzuführen. Die Parallelen zwischen kognitionswissenschaftlichen Grundlagen spezieller KI-Modelle und literarischen Detektiven wurden schon einmal in einem Seminar thematisiert, das Klüver gemeinsam mit J. Reichertz 1997 durchführte. Dabei wurde ein Vorläufer eines der Programme eingesetzt, die hier dargestellt werden, nämlich ein Expertensystem.

Die Verwendung, speziell von Expertensystemen, im Kontext von Kriminalgeschichten ist nicht gänzlich neu. Bereits 1987 wurde in einem Einführungsbuch zur Künstlichen Intelligenz ein kleines Expertensystem dargestellt, das zur Unterstützung bei der Lösung eines fiktiven einfachen Kriminalfalls herangezogen wurde (Time Life 1987). Der Fall wurde speziell für das Expertensystem konstruiert. Ebenso verfuhr Favre-Bulle (2001), der einen kleinen selbst erdachten Fall durch ein Expertensystem lösen ließ. Die Verwendung von Expertensystemen bietet sich bei derartigen Geschichten auch an, da diese Geschichten häufig in der Art von Denksportaufgaben konstruiert sind und die auffallende Ähnlichkeit zwischen Expertensystemen und der Rationalität der klassischen literarischen Detektive – von Sherlock Holmes bis Hercule Poirot – uns schon vor einiger Zeit aufgefallen war. Neu an unserem Projekt waren die Verwendung einer klassischen Geschichte von Agatha Christie, also nicht eines speziell für die Programme konzipierten Falls, der didaktische Kontext und vor allem die parallele Verwendung unterschiedlicher KI-Programme.

Expertensysteme wurden übrigens mit unterschiedlichen Erfolgen auch in der realen Polizeiarbeit eingesetzt (Reichertz 1994). Die Polizei von Baltimore beispielsweise entwickelte ein Expertensystem zur Aufklärung von Einbruchsdelikten; das FBI versucht, Täterprofile von Serientätern mit Hilfe von Expertensystemen zu erstellen. Das freilich ist ein anderes Thema.

## **1. Das Seminar**

Im Studiengang Kommunikationswissenschaft in Essen wird in unterschiedliche Methoden des wissenschaftlichen Denkens eingeführt, darunter natürlich auch die Hermeneutik der Textinterpretation. Das besondere an diesem Studiengang ist die Möglichkeit, auch formale Denkweisen zu lernen und zwar anhand von Themen wie Künstliche Intelligenz, Modelle zur Konstruktion von Computersimulationen u. ä. m. Das Ziel unseres Seminars bestand darin, zu zeigen, wie die Interpretation von Kriminalgeschichten durch die Verwendung von Computerprogrammen unterstützt werden kann, die aus dem Forschungsbereich der Künstlichen Intelligenz stammen. Da es sich hier um die Modellierung kognitiver Prozesse handelt, bot sich das Genre der Kriminalgeschichte an, bei der es immer um das Lösen bestimmter Probleme geht. Damit war auch das Ziel der Interpretation der einzelnen Romane und Geschichten festgelegt: Es sollte nicht so sehr um eine Interpretation der Gesamtheit der Texte gehen, sondern es sollte die rationale Vorgehensweise der jeweiligen Detektive rekonstruiert werden.

Die Studierenden konnten Leistungsnachweise zum einen durch Referate zu ausgewählten Kriminalromanen erwerben, wobei es vor allem um die Rekonstruktion der von den Detektiven verwendeten rationalen Problemlösestrategien gehen sollte. Zum anderen sollten kleine Gruppen von Studierenden mit zwei von uns entwickelten Computerprogrammen arbeiten, nämlich mit einem Expertensystem und einem interaktiven neuronalen Netz (s. u.). Die Studierenden hatten dabei die Aufgabe, den Programmen sukzessive die Informationen zu einem Kriminalfall aus einer Geschichte von Agatha Christie einzugeben. Gleichzeitig sollten die Teilnehmer versuchen, sozusagen im Wettstreit mit den Programmen, den Mörder in der Geschichte selbst herauszufin-

den und zwar möglichst vor den Programmen. Praktisch sah das so aus, dass die Studierenden nicht die gesamte Geschichte auf einmal erhielten, sondern pro Woche ca. ein Fünftel. Damit wussten die Studierenden immer genauso viel über den Fall wie der literarische Detektiv – Hercule Poirot – und die beiden Programme. Die – für die eigentlichen Ziele des Seminars natürlich eher sekundäre – Frage war dann, ob die Programme und die Studierenden den Täter korrekt und zwar beweisbar identifizieren konnten, *bevor* Poirot in der Geschichte den Täter benennt und sein eigenes Vorgehen erklärt. Insbesondere ging es (auch) darum, ob die Studierenden oder die Programme schneller sein würden. Nach dem Ende des Seminars konstruierten wir zusätzlich eine Kohonen-Karte und gaben dieser ebenfalls die Informationen über den Fall ein.

Der Verlauf des Seminars bestand darin, dass in einer ersten Phase sämtliche Studierende einen Überblick über die logische Basis der verwendeten Modelle durch uns erhielten. Anschließend folgten Referate zu Agatha Christies „Mord im Orientexpress“, Chandlers „Little Sister“ und van de Weterings „The Mind Murder“ (deutsch: „Ketchup, Karate und die Folgen“). Während dieser Phase begannen die Studierenden auf die beschriebene Weise mit den Programmen zu arbeiten. In der letzten Seminarphase stellten dann diese Studierenden ihre Ergebnisse vor, erläuterten ihre Arbeitsweise und verglichen ihr subjektives Vorgehen mit dem der Programme.

## **2. Die Geschichte**

Der ausgewählte Text war „Die spanische Truhe“ von Agatha Christie. Die Geschichte folgt einem Muster, das Lesern der Queen of Crime gut vertraut ist:

Nach einer Party im Haus von Major Rich in London wird am nächsten Morgen die Leiche von Arnold Clayton, einem Freund von Rich, in einer spanischen Truhe im Wohnzimmer des Hauses gefunden und zwar vom Butler. Dieser sagt aus, dass Clayton vor der Party ins Haus gekommen war, um Rich zu sprechen. Clayton wollte sich für die Party entschuldigen, da er ein Telegramm erhalten hatte, das ihn nach Schottland rief. Da Rich noch nicht im Hause war, ließ der Butler Clayton alleine, der Rich die Nachricht schreiben wollte. Danach wurde Clayton nicht mehr gesehen und Rich sagt aus, dass er Clayton auch nicht gesehen habe. Die anderen Partygäste, die ebenfalls sowohl mit Rich als auch mit Clayton befreundet sind, haben Clayton nach dessen Eintreffen im Haus von Rich gemäß ihren Angaben auch nicht zu Gesicht bekommen. Die anderen Partygäste sind Ms. Clayton, die Frau des Opfers, Commander McLaren, ein Freund von Rich und den Claytons, sowie Mr. und Ms. Spence, ein mit allen anderen Personen befreundetes Ehepaar.

Der Polizei ist bekannt, dass Rich in Ms. Clayton verliebt ist, die als sehr attraktiv geschildert wird und die seine Gefühle erwidert. Demnach hat Rich ein Motiv, den Ehemann seiner Angebeteten umzubringen. Da alle Gäste erst nach Rich in dessen Haus eintrafen, haben sie anscheinend ein Alibi. Clayton muss vor der Party getötet worden sein, da Clayton von niemand mehr gesehen wurde. Rich hat ein Motiv und als Einziger – bis auf den Butler – kein Alibi. Da der Butler kein Motiv hat, muss Rich der Täter sein; er wird dementsprechend verhaftet.

Poirot glaubt dies nicht und sammelt folgende Fakten: a) McLaren ist ebenfalls in Ms. Clayton verliebt; b) das Telegramm, von dem Clayton sprach, ist eine Fälschung; c) McLaren und Clayton haben sich, kurz bevor Clayton zu Rich ging, in ihrem Club getroffen; d) in der Wand der spanischen Truhe sind frisch gebohrte Löcher; e) die Tatwaffe hat den Spences gehört; f) Clayton wusste von den Gefühlen zwischen Rich und seiner Frau und war – obwohl zu Unrecht – äußerst eifersüchtig. Konstitutiv für die Lösung dieses Falls sind demnach die Fragen nach

Alibis, Motiven, Zugangsmöglichkeiten zur Waffe, das Wissen darüber, wo genau Clayton zum Zeitpunkt des Mordes war und natürlich die tatsächliche Tatzeit (der Gerichtsmediziner kann die Tatzeit nur auf ca. drei Stunden genau datieren).

Entscheidend für die Bestimmung der Tatzeit – und damit für die Entscheidung, wer ein Alibi hat – ist die Entdeckung der Löcher in der Truhe. Da die Löcher offenbar frisch gebohrt worden sind, folgert Poirot, dass sie dazu dienten, das Atmen innerhalb der Truhe zu ermöglichen. Demnach muss jemand die Absicht gehabt haben, sich in die Truhe zu legen und von dort aus die Party zu beobachten. Dies konnte nur Clayton selbst gewesen sein, dessen Eifersucht (s. o.) bekannt war. Also hatte Clayton gar nicht mit Rich gesprochen, sondern hatte die Löcher gebohrt und sich selbst in die Truhe gelegt. Dies wiederum ergibt, dass Clayton nicht vor sondern während der Party ermordet worden war, womit keiner der Gäste mehr ein Alibi hatte. Die weitere Analyse der beteiligten Personen ergab, dass bis auf den Butler alle Zugangsmöglichkeiten zur Waffe hatten, dass jedoch nur Rich und McLaren ein Motiv hatten (Fakt a) und dass praktisch nur McLaren von den Plänen Claytons wissen konnte (Fakt c). Da auf den Mörder alle üblichen Aspekte eines Mordfalls zutreffen müssen – Motiv, Gelegenheit, Zugang zur Waffe, kein Alibi –, konnte durch Ausschluss der anderen Personen praktisch nur McLaren der Täter sein. Soweit die Geschichte.

### **3. Die Programme**

Die obige Rekonstruktion der Struktur der Geschichte ergab natürlich auch den Rahmen für die Konstruktion der Programme sowie die Einweisung der Studierenden. Da die beiden im Seminar verwendeten Programme auf einen ersten Blick und insbesondere in der Benutzung sehr unterschiedlich sind, mussten die entsprechenden Studentengruppen auch unterschiedlich eingewiesen werden.

#### **3.1. Das Expertensystem HERCULE 1**

Ein Expertensystem oder besser regelbasiertes System besteht zum einen aus „Fakten“, d. h. verschiedenen Informationen, und zum anderen aus Regeln, die die Fakten miteinander verknüpfen. Bei medizinischen Diagnosesystemen etwa werden bestimmte Symptome, die das System als Teil der Fakten hat, mit bestimmten Krankheiten und Therapien verknüpft – den anderen Faktenteilen –, so dass das System durch deduktive Schlussfolgerungen auf die Eingabe einzelner Symptome mit der Ausgabe wahrscheinlicher Krankheiten und vorgeschlagener Therapien reagieren kann.

Die Studierenden erhielten in unserem Fall streng genommen kein fertiges Expertensystem, da sie die entsprechenden Fakten selbst in die Wissensbasis eingeben sollten. Das Programm, in einer JAVA-Implementation von Eric Schmieders geschrieben, bestand demnach aus fallspezifischen Regeln, die insbesondere die eingegebenen Fakten bestimmten „Kategorien“ sowie den Personen zuordnen sollten. Als „Kategorien“ wurden die Begriffe „Motiv“, „Alibi“, „Zugang zur Waffe“ und „Wissen um die Tat“, d. h. den tatsächlichen Aufenthalt von Clayton zur Tatzeit, in das Programm eingegeben. Eine Eingabe erfolgte .z. B. in der Art „Butler, vor Party, Rich-Haus“ und „Tatzeit, vor Party“, da diese – nachträglich falsche – Tatzeit zu Beginn der Geschichte als Fakt eingegeben werden musste. Da das Programm eine kategoriale Zuordnung leisten kann, wendet es die Regeln an „wenn X vor Party, dann zur Tatzeit“ und „wenn X zur Tatzeit, dann kein Alibi“. Da die Regeln Zuordnungen zu den Kategorien *und* den Personen durchführen, konnte

das Programm schließen „Butler, kein Alibi“. Entsprechend führen Eingaben wie „McLaren, Liebe, Ms. Clayton“ zu den Zuordnungen „wenn Liebe, Ms. Clayton, dann Motiv“ sowie „McLaren, Motiv“. Wenn während der Arbeit mit HERCULE 1 andere Fakten über bestimmte Sachverhalte eingegeben werden, als bereits im System vorhanden, dann löscht das Programm die alten Fakten und wendet auf die neuen Fakten die Regeln an. So wird dann aus der ursprünglichen Folgerung „Mr Spence, Alibi“ (aufgrund der ursprünglichen Tatzeit) die spätere und endgültige Folgerung „Mr Spence, kein Alibi“.

Neben diesen Zuordnungsregeln gibt es noch zwei allgemeine Regeln:

- a) Da letztlich die Aufgabe des Programms darin besteht, aus dem Kreis der Verdächtigen den Mörder zu bestimmen, musste es das Konzept „Mörder“ bestimmen können. Dies geschieht durch die Regel „wenn Person X, Motiv, Zugang, Wissen, kein Alibi, dann Person X, Mörder“. Die Anwendung dieser Regel kann natürlich dazu führen, dass keine oder mehr als eine Person als Mörder identifiziert werden. Da wir die Geschichte vor Fertigstellung des Programms kannten, wussten wir, dass bei vollständiger Eingabe der Fakten das Programm eine eindeutige Lösung, d. h. genau eine Person, geben konnte. (Dies ist z. B. beim erwähnten Roman von Christie „Mord im Orientexpress“ anders, da dort 12 Personen gemeinsam den Mord begangen haben.) Voraussetzung ist natürlich die vollständige Eingabe *aller* Fakten.
- b) Die Lösung des Falles basiert insbesondere, wie skizziert, auf der Deutung der frisch gebohnten Löcher. Im Gegensatz zu den fallspezifischen Informationen wie „kein Alibi des Butlers“ muss hier eine Regel verwendet werden, die eine Form von „Weltwissen“ enthält. Insbesondere an dem allgemeinen Weltwissen von Menschen sind bekanntlich die Versuche gescheitert, Expertensysteme als *allgemeine* Modelle menschlichen Problemlöseverhaltens zu konstruieren. Das Programm enthält deswegen die Regeln „wenn Löcher, Truhe, dann Atmen“ und „wenn Atmen in Truhe, dann Clayton in Truhe“. Die Regel hätte komplexer, d. h. indirekter formuliert werden können, aber hier ging es nur um die Möglichkeitsgarantie, dass HERCULE 1 den Fall auch tatsächlich lösen kann. Wir testeten übrigens diese Regel an verschiedenen Studenten, die die Geschichte nicht kannten, mit der Frage, warum man Löcher in eine abgeschlossene Truhe bohrt; regelmäßig erhielten wir die Antwort, dass dies geschieht, um Luft zum Atmen zu haben – innerhalb der Truhe natürlich. Anscheinend gehört diese Regel tatsächlich zum Weltwissen von Menschen, so dass wir sie guten Gewissens in das Programm eingeben konnten. „Guten Gewissens“ bedeutet, dass wir dem Programm damit keinen unfairen Vorteil gegenüber den Studierenden einräumten.

Mit dieser Regelstruktur war die Entwicklung von HERCULE 1 abgeschlossen (eine genauere Beschreibung des Programms kann Interessenten gerne zugeschickt werden). Offensichtlich handelt es sich bei diesem System ebenfalls um ein klassisches Diagnosesystem (s. o.), die längst in verschiedenen praktischen Kontexten Verwendung finden. Bevor die Ergebnisse der Studentengruppe geschildert werden, die mit diesem Programm gearbeitet hat, sollen noch die beiden anderen Programme dargestellt werden.

### 3.2. Das interaktive neuronale Netz HERCULE 2

Interaktive Netze (IN) sind ein spezieller Typ neuronaler Netze: Sie bestehen aus Einheiten („Neuronen“), die miteinander verbunden sind. Diese Verbindungen werden „gewichtet“, d. h. sie werden mit bestimmten numerischen Werten versehen. Diese Werte modifizieren die Signale, die von einem Neuron zu anderen Neuronen gehen. Die Signale sind ebenfalls numerische

Werte, die von den Gewichtswerten der Verbindungen verstärkt oder verringert werden. Diese Gewichtswerte werden in der so genannten Gewichtsmatrix zusammengefasst, wobei eine Null in der Matrix bedeutet, dass es zwischen den entsprechenden Neuronen keine Verbindung gibt.

Prinzipiell können bei interaktiven Netzen alle Neuronen mit allen verbunden sein, d. h. die Gewichtsmatrix enthält dann keine Nullen. Das Lernen neuronaler Netze (s. u.) geschieht dadurch, dass aufgrund bestimmter „Lernregeln“ die Werte in der Gewichtsmatrix systematisch verändert werden. Interaktive Netze werden gewöhnlich nicht durch Lernregeln automatisch trainiert, so dass die korrekten Gewichtswerte vom Benutzer selbst eingegeben werden müssen. Es gibt natürlich die Möglichkeit, dies durch geeignete Optimierungsverfahren automatisch geschehen zu lassen – z. B. durch die Koppelung interaktiver Netze mit so genannten genetischen Algorithmen (Stoica 2000). Für die Ziele unseres Seminars verzichteten wir auf diese Möglichkeit, da die Studierenden selbst eine adäquate Gewichtsmatrix konstruieren sollten.

Netze dieser Art eignen sich u. a. dafür, logische Zusammenhänge zwischen einzelnen Aussagen darzustellen und damit Schlussfolgerungen zu ziehen (vgl. Waltz und Pollack 1985; Stoica 2000). Allerdings enthält – natürlich – ein interaktives Netz keine expliziten Regeln wie Expertensysteme. Die logischen Zusammenhänge zwischen verschiedenen Aussagen wie „McLaren liebt Ms Clayton“ und „McLaren hat ein Motiv“ müssen durch entsprechende Werte in der Gewichtsmatrix ausgedrückt werden, also z. B. durch „ $w(\text{McLaren, Ms Clayton}) = 1$ “, „ $w(\text{Mr Spence, Ms Clayton}) = 0$ “ und „ $w(\text{McLaren, Motiv}) = 1$ “ bzw. „ $w(\text{Mr Spence, Motiv}) = 0$ “.

Nach Erstellung einer Gewichtsmatrix werden eine oder mehrere – evtl. alle – Neuronen „extern aktiviert“. Dies bedeutet, dass dieser Teil der Neuronen als so genannte Eingabeschicht fungiert, diese werden mit einem bestimmten Aktivierungswert belegt, woraufhin sich die Dynamik des Netzwerks entfaltet. Aufgrund der Struktur eines interaktiven Netzes werden durch die externe Aktivierung alle Neuronen mehr oder weniger stark aktiviert; dies wird gewöhnlich in einer visuellen Darstellung des Endergebnisses vollständig angezeigt. Die Verbreitung eines Signals von der Eingabeschicht aus besteht demnach darin, dass die sendenden Neuronen ihren jeweiligen Zustand an andere Neuronen weitergeben, modifiziert durch die Gewichtswerte. Diese Aktivierungsausbreitung geschieht meistens gemäß der so genannten linearen Aktivierungsfunktion

$$A_i = \sum w_{ij} * A_j,$$

wobei  $A_i$  der Aktivierungswert des „empfangenden“ Neurons  $i$  ist,  $w_{ij}$  die Gewichtswerte zwischen dem empfangenden Neuron  $i$  und den „sendenden“ Neuronen  $j$  sind und  $A_j$  deren Aktivierungswerte. Zusätzliche technische Details können u. a. bei Stoica (2000) nachgelesen werden.

Zu Beginn einer IN-Simulation werden bis auf die extern aktivierten Neuronen alle Aktivierungswerte, d. h. die Zustandswerte der Neuronen, auf Null gesetzt. Das Ende der Simulation wird entweder durch ein Abbruchkriterium – z. B. Anzahl der Simulationsschritte, Approximation der Aktivierungswerte an ein vorgegebenes Ziel – oder dadurch definiert, dass das IN einen Attraktor, d. h. einen stationären Zustand erreicht, der die Aktivierungsdynamik auf bestimmte Werte fixiert. Dabei kann es durchaus geschehen, dass das IN keinen so genannten Punktattraktor erreicht, sondern einen Attraktor der Periode größer als 1. In dem Fall „oszilliert“ das IN um verschiedene Endzustände, es gibt demnach keine eindeutige Lösung des Problems aus. Da die Dynamik eines IN praktisch immer nichtlinear aufgrund der speziellen Topologie ist, lässt sich a priori nur selten prognostizieren, ob ein Punktattraktor erreicht wird und falls ja, welcher.

In dem Fall des Mordes an Clayton würde eine korrekte Lösung des Falls durch das IN bedeuten, dass ein bestimmtes Neuron, das McLaren repräsentiert bzw. McLaren als Mörder, signifikant am stärksten von allen Neuronen am Ende der Simulation aktiviert ist. Insbesondere sollte das IN einen Punktattraktor erreichen, d. h., es sollte sich ohne Abbruchkriterium stabilisieren.

Die Aufgabe der Studierenden, die mit diesem IN arbeiteten, bestand nun darin, aufgrund der ihnen zur Verfügung gestellten Textteile die Anzahl der Neuronen festzulegen – z. B. ein Neuron pro Person – und eine Gewichtsmatrix zu konstruieren. Die Studierenden hatten ein IN-Shell zur Verfügung, in das sie Anzahl und Bezeichnungen der Neuronen eingeben konnten sowie die verschiedenen numerischen Werte in der Gewichtsmatrix. Im Gegensatz zu der Gruppe, die mit HERCULE 1 arbeitete und die keine Vorinformation über die Geschichte und die logische Struktur des Programms erhalten hatte, bekam die 2. Gruppe den Hinweis, dass die Fakten auf die Personen und auf die genannten vier Kategorien „Motiv“, „Alibi“, „Zugang“ und „Wissen“ bezogen werden müssten. Daraus ergab sich für die Studierenden, dass sie mit mindestens 10 Neuronen zu arbeiten hatten, nämlich sechs für die Personen und vier weitere. Die Studierenden erhielten diese zusätzliche Information, da ihre Aufgabe deutlich schwerer war als die der ersten Gruppe. Diese brauchte ja nur die Fakten codiert in HERCULE 1 einzugeben, während die zweite Gruppe HERCULE 2 sozusagen auch logisch, nicht nur inhaltlich ergänzen musste. Es ist wohl kein Zufall, dass die Erfolge der beiden Gruppen bei den verschiedenen Aufgabenstellungen sehr unterschiedlich waren.

Es sei hier nur angemerkt, dass die logische Struktur von HERCULE 2 für die übrigen Seminarteilnehmer nur schwer oder gar nicht zu vermitteln war. Trotz einer generellen Einführung in die Operationsweise neuronaler Netze und insbesondere die von IN, und trotz einer nachträglichen Erklärung von uns wie HERCULE 2 auf den Täter gekommen ist, war an der Reaktion der primär geisteswissenschaftlich ausgebildeten Studenten zu erkennen, dass HERCULE 2 ihnen z. T. ein Rätsel geblieben war. Dies wurde sehr deutlich an Äußerungen wie „Ich begreife nicht, wie ein Programm durch Eingabe bestimmter Zahlen einen Täter identifizieren kann“. Für die gleichen Seminarteilnehmer war es dagegen kein *prinzipielles* Problem, die Funktionsweise von HERCULE 1 zu verstehen und aufgrund unserer nachträglichen Erläuterung nachzuvollziehen, wie HERCULE 1 den Täter ermitteln konnte. Wir werden auf diese verschiedenen Erfahrungen am Ende noch einmal zurückkommen.

### **3.3. Die Kohonen-Karte HERCULE 3**

Kohonen-Karten oder auch Self-Organizing-Maps (SOM) sind der schon fast klassische Prototyp neuronaler Netze, die „nicht überwacht“ lernen. Gemeint ist damit, dass im Gegensatz zu überwacht lernenden neuronalen Netzen ein SOM keine externen Vorgaben erhält, wie es seine Gewichtswerte zu justieren hat. Dies wird realisiert durch das so genannte „Winner-Take-All-Prinzip“, welches bedeutet, dass nur die Neuronen mit den aktuell höchsten Aktivierungswerten ihre Gewichtswerte mit einer bestimmten Lernrate und abhängig vom Output der sendenden Einheiten verändern können. Als Aktivierungsfunktion wird gewöhnlich eine sigmoide Funktion verwendet (vgl. dazu Stoica 2000). Ein Lernprozess ist dann beendet, wenn das Programm sich stabilisiert, also einen Punktattraktor erreicht.

Man kann sich die Operationsweise von SOM so verdeutlichen, dass diese gewissermaßen Cluster der verschiedenen Eingaben generieren, wobei die Clusterungen nach Musterähnlichkeiten erfolgen. Ritter und Kohonen (1989) illustrieren dies Operationsprinzip sehr anschaulich

an dem Beispiel der Ordnung von zoologischen Begriffen: Das SOM erhält Eingaben der Form „Löwen fressen Fleisch“, „Adler fliegen“, „Kühe sind Säugetiere“, „Adler legen Eier“, „Stare fliegen“, „Kühe sind Säugetiere“ etc. Das SOM bildet jetzt Cluster, die praktisch verschiedene Tierbegriffe in unterschiedliche Klassen einordnen – also z. B. Kühe und Löwen versus Adler und Stare. Eine visuelle Ausgabe eines SOM besteht dann darin, dass die logisch zusammengehörigen Begriffe einen räumlichen Cluster bilden und die anderen Begriffe davon entfernte Cluster. In gewisser Weise geht ein derartiges SOM demnach so vor, dass es die verschiedenen Begriffe um „Prototypen“ herum ordnet, wie es die bekannte Prototypentheorie von Rosch für das menschliche Denken postuliert (vgl. z. B. Lakoff 1987).

Entscheidend für die Verwendung von Kohonen-Karten ist, dass ein SOM nicht nur eine – ggf. variable – Matrix enthält, wie ein IN, sondern zwei: Zum einen hat ein SOM natürlich eine variable Gewichtsmatrix, wie alle lernenden Netzwerke. Zum anderen werden die Informationen, die die Kohonen-Karte nach Mustern ordnen soll, in einer „semantischen“ Matrix gespeichert. In dem obigen Beispiel zoologischer Begriffe besteht die Matrix aus den Tiernamen für die Zeilen und den Eigenschaften für die Spalten:

	Adler	Löwe	Kuh	etc.
fliegt	1	0	0	
frisst Fleisch	1	1	0	
Federn	1	0	0	
etc.				

Die Matrix ist also binär codiert und gibt an, ob bestimmte Eigenschaften auf einzelne Begriffe zutreffen oder nicht. Die semantische Matrix für das Kriminalproblem muss demnach in der einen Dimension die in der Geschichte auftretenden Personen kennzeichnen und in der anderen Dimension die kriminalistischen Kategorien wie Motiv, Alibi etc.

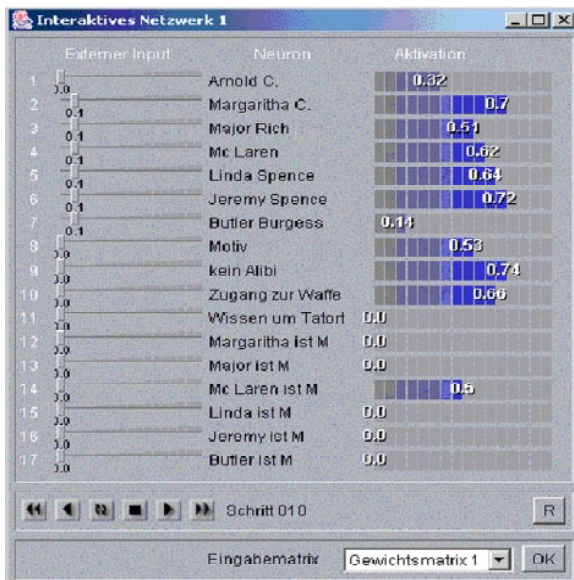
Wir verzichteten bewusst darauf, das SOM bereits während des Seminars einzusetzen, da nach unseren didaktischen Erfahrungen die Logik dieses Netzwerktyps Studierenden nur schwer zu vermitteln ist. Außerdem sollte das Seminar nicht zu sehr durch technische Probleme überfrachtet werden. Deswegen erwähnten wir diese zusätzliche Möglichkeit nur und zeigten interessierten Studenten die Ergebnisse.

#### 4. Experimentelle und didaktische Resultate

(a) Am schlechtesten schnitt die Gruppe ab, die sich mit HERCULE 1, also dem Expertensystem beschäftigt hatte. Die Gruppe von vier Studentinnen arbeitete gemeinsam an den Eingaben und versuchte auch durch gemeinsame Diskussionen den Täter möglichst früher als das Expertensystem zu identifizieren. Leider gelang es dieser Gruppe weder, den Täter durch eigene Schlussfolgerungen vor Bekanntgabe der endgültigen Lösung korrekt und vor allem *begründet* zu identifizieren, noch konnte sie HERCULE 1 alle notwendigen Informationen geben. Deswegen konnte auch das Programm keine definitive Lösung angeben. Vor allem war nicht die Tatsache der frisch gebohrten Löcher eingegeben worden, da die Gruppe dies für nicht relevant hielt. Eine Studentin aus der Gruppe charakterisierte ihren doppelten Misserfolg selbstironisch damit, dass „sie wohl mit typisch weiblicher Intuition vorgegangen wären“. Vorsichtshalber hatten wir für beide Programme korrekte Eingaben vorbereitet, so dass wir in der entsprechenden Sitzung demonstrieren konnten, dass HERCULE 1 sehr wohl in der Lage ist, den Fall korrekt zu lösen. Offenbar fiel es diesen Studentinnen schwer, sich von den eher lässigen Lesegewohnheiten zu

lösen, die bei der Lektüre von Kriminalgeschichten die übliche Einstellung sind, und sich statt dessen auf die logische Struktur des Textes zu konzentrieren.

(b) Deutlich besser waren die Ergebnisse der zweiten Gruppe, die es mit dem interaktiven Netzwerk HERCULE 2 zu tun hatte. Die Teilnehmer aus dieser Gruppe arbeiteten bis auf zwei Studentinnen einzeln und konnten fast alle die Gewichtsmatrix von HERCULE 2 so konstruieren, dass das Programm den Täter (fast) eindeutig identifizieren konnte. Eine typische Lösung von HERCULE 2 sieht folgendermaßen aus (Version von Steffi Cordes und Bianca Kierstein):



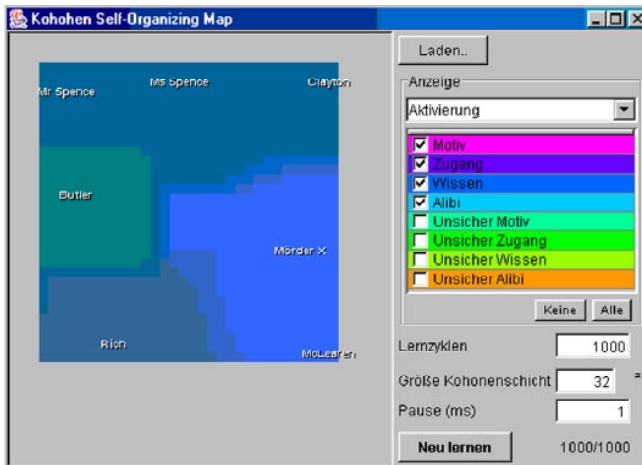
Man sieht, dass es sich um eine Liste von 17 Neuronen handelt, in der die beteiligten Personen gewissermaßen zweimal enthalten sind (bis auf das Opfer Clayton). Die zweite Repräsentation der Personen durch die Neuronen 12-17 „Margarita – der Vorname von Ms Clayton – ist M“ bedeutet, dass sie (potentielle) Mörder sind. Die ausschließliche Aktivierung des Neurons 14 „McLaren ist M“ zeigt, dass das Programm den Mörder eindeutig identifizieren konnte.

Von der Struktur der Geschichte her hätte man annehmen können, dass 10 Neuronen genügt hätten, nämlich die Personen und die Kategorien sowie evtl. noch ein Neuron „Mörder“, also 11 Neuronen. Das wurde auch mehrfach versucht; es zeigte sich jedoch, dass die entsprechende  $10 * 10$  Gewichtsmatrix zu klein war, um eine eindeutige Lösung zu erhalten. Deswegen verfielen die erfolgreichen Teilnehmer aus dieser Gruppe auf die Idee (bzw. wurden durch uns dazu angeregt), die Matrix u. a. auf die dargestellte Weise zu erweitern. Die Notwendigkeit, bei etwas komplizierteren Aufgaben mit möglichst großen Matrizen zu arbeiten, ist aus Theorie und Praxis neuronaler Netze her wohl bekannt. Generell lässt sich dieser Umstand so charakterisieren, dass neuronale Netze wie alle informationsverarbeitenden Systeme eine Mindestkomplexität haben müssen, um ihre Aufgaben zu lösen. Allerdings hätte es in diesem Fall auch gereicht, mit einer  $11 * 11$  Matrix zu arbeiten, wobei freilich die Gewichtswerte sich sehr deutlich hätten unterscheiden müssen (s. u.).

Die Teilnehmer(innen) aus dieser Gruppe konnten meistens auch den Mörder durch eigenes Nachdenken identifizieren, wenn auch etwas unsicher. Der Grund dafür, dass sie auch hier erfolgreicher waren als die Studentinnen der ersten Gruppe dürfte darin liegen, dass die zweite Gruppe von uns in die „kategoriale“ Struktur der Geschichte eingewiesen wurde, nämlich um die Zuordnung von Fakten zu den vier Kategorien, wodurch sie praktisch gezwungen waren, bei ihrem eigenen Vorgehen mit und ohne Programm sich ständig an dieser Struktur zu orientieren. Obwohl also die Aufgabenstellung dieser Gruppe schwieriger war als die der ersten Gruppe – die zweite Gruppe musste ja die Fakten sofort in ihrer Bedeutung für die kategoriale Struktur interpretieren –, hatte die zweite Gruppe vermutlich dadurch Erfolg, dass sie – wenn auch von uns angeleitet – von Anfang an strukturorientiert vorging.

(c) Die Kohonen-Karte wurde, wie bemerkt, nicht in dem Seminar eingesetzt, sondern von uns nur der modelltheoretischen Vollständigkeit halber konstruiert. (Die hier verwendete Kohonen-Karte wurde von Rouven Malecki implementiert.) Da bei einer Kohonen-Karte sowohl eine variable Gewichtsmatrix vorhanden sein muss als auch eine vom Benutzer zu konstruierende semantische Matrix, war unsere Aufgabe relativ einfach: Gemäß der „Prototypen-Logik“ von SOM (s. o.) definierten wir die Person „Mörder X“ als Prototyp, um den die anderen Personen zu clustern waren. Je näher eine Person dem Prototyp ist, desto höher ist die Wahrscheinlichkeit, dass diese Person der Täter ist. Gemäß den Informationen der Geschichte sieht dann die semantische Matrix für das SOM HERCULE 3 folgendermaßen aus:

	Motiv	kein Alibi	Zugang Waffe	Wissen Tat
Rich	1	1	1	0
McLaren	1	1	1	1
Mr Spence	0	1	1	0
Ms Spence	0	1	1	0
M. Clayton	0	1	1	0
Butler	0	1	0	0
Mörder X	1	1	1	1



Die darauf basierenden Durchläufe des SOM wurden mit 40 Neuronen durchgeführt (bei SOM kann man unabhängig von dem Problem die Anzahl der Neuronen bzw. die Größe der Gewichtsmatrix unterschiedlich festlegen). Ein typisches Ergebnis sieht dann folgendermaßen aus:

Man sieht, dass auch das SOM McLaren eindeutig als Täter identifiziert hat. Die anderen Personen sind relativ weit von X entfernt, auch wenn z. B. Rich nicht sehr nahe an X gesetzt wird. Allerdings ist Rich in einer räumlichen Zone, die farblich der von X entspricht – diese Version eines SOM clustert nämlich sowohl nach räumlicher Verteilung als auch nach Färbung der Raumsegmente.

(d) Praktisch gesehen ist ein SOM mit einem derart gering dimensionierten Problem eigentlich unterfordert. Wenn eine semantische Matrix der obigen Art vom Benutzer erst einmal konstruiert ist, braucht man auch eigentlich kein SOM mehr, da man aus den Vektoren sehr schnell selbst bestimmen kann, dass nur McLaren als X infrage kommt. Wünschenswert wäre es von daher, ein SOM-System zu haben, in das wie beim Expertensystem nur die Fakten eingegeben werden müssen, ohne diese wie beim IN und beim SOM schon selbst zu deuten, d. h. auf die Fallkategorien zu beziehen. M. a. W., gesucht ist ein zusätzliches Programm, das die semantische Matrix für das SOM automatisch generiert. Natürlich darf dies nicht durch explizite Regeln wie bei HERCULE 1 geschehen, da es sonst nur eine Variation des regelbasierten Systems wäre.

Wir lösten dies Problem, indem wir einen heteroassoziativen Speicher (HS) vor das SOM koppelten. Heteroassoziative Speicher sind so genannte feed forward Netze, die nach einem Trainingsprozess gelernten Muster auch bei fehlerhafter Eingabe wieder erkennen können und sie einem anderen Muster zuordnen (deswegen „hetero“). Der Trainingsprozess als überwachtes Lernen geschieht dadurch, dass man das Eingabemuster X und das Ausgabemuster Y als Vektoren mit numerischen Komponenten codiert und dann ein „Vektorprodukt“  $X * Y$  berechnen lässt, dessen Ergebnis die zu dem Musterpaar gehörige Gewichtsmatrix ist. Es handelt sich also hier weder um das skalare Vektorprodukt, dessen Ergebnis ein Skalar ist, noch um die Vektormultiplikation in einer Vektoralgebra, deren Ergebnis ein Vektor ist. (Technische Details können in jeder Einführung zu neuronalen Netzen nachgelesen werden.) Gibt man nach diesem Trainingsprozess X – ggf. auch fehlerhaft – ein, so generiert das Netzwerk Y. Der HS wurde implementiert von Marco Schütz, Christoph Carls und Martin Hoffmann.

Unser HS wurde nun so trainiert, dass wir das Netz die Paare (Liebe zu M. Clayton, Motiv), (Tatzeit Haus, Rich kein Alibi), (Freund von Spences, Zugang zur Waffe) und (Kontakt zu Clayton, Wissen um Tatzeit) lernen ließen. Wenn nun ein Benutzer die Information eingibt, „McLaren, Liebe zu Ms Clayton“, dann ordnet der HS dem semantischen Vektor von McLaren im SOM eine 1 an der Stelle „Motiv“ zu. Wird die Information „Butler, keine Liebe zu M. Clayton“ eingegeben, dann wird dem semantischen Vektor des Butlers an der Motiv-Stelle eine 0 zugeordnet. Der HS kann also in einfacher Weise generalisieren, da er *ohne dies explizit gelernt zu haben* „nicht Liebe“ eine 0 zuordnet.

Da der HS nach diesen Prinzipien ohne Probleme die obige semantische Matrix generieren kann, sind die anschließenden Simulationen durch das SOM natürlich identisch mit der oben dargestellten. Wir haben damit ein zweites System, in das nur die Informationen eingegeben werden müssen und das anschließend den Mörder korrekt identifizieren kann.

(e) Schließlich ist es auch möglich, durch Koppelung des heteroassoziativen Speichers mit dem IN dieses ebenfalls auf entsprechende Weise zu erweitern. M. a. W., man kann durch einen HS die Gewichtsmatrix eines IN generieren lassen, so dass auch das IN nur mit ungedeuteten In-

formationen auskommt. Dies geschieht so, dass der HS wie in (d) trainiert wird, allerdings mit der Maßgabe, dass der Information „Liebe“ wieder eine 1 zugeordnet wird, der Information „nicht Liebe“ dagegen eine -1 und entsprechend bei den anderen Informationen (eine bipolare Codierung). Alles andere ist identisch mit der Generierung der semantischen Matrix für das SOM. Nach der entsprechenden Trainingsphase legen wir fest, dass das IN aus 11 Neuronen bestehen sollte, nämlich wieder die sechs Verdächtigen, Mörder X sowie die vier Kategorien. Die entsprechende  $11 \times 11$  Matrix enthält dann eine 1, wenn eine Kategorie für eine Person in Frage kommt und sonst eine -1. Gewichtswerte, die z. B. Beziehungen zwischen den Personen entsprechen und für die es weder positive noch negative Informationen gibt, erhalten den Wert 0. Die entsprechende Gewichtsmatrix sieht dann so aus:

Eine Lösung des Mordfalls durch das erweiterte IN geschieht jetzt durch externe Aktivierung des Neurons Mörder X. Das Ergebnis ist, dass durch die Generierung einer ternär (1, 0, -1) codierten Gewichtsmatrix man nicht nur mit 11 Neuronen auskommt, sondern dass der Mörder sogar noch eindeutiger als im obigen Beispiel (b) identifiziert wurde. Damit kann auch das IN wie HERCULE 1 nur durch Informationseingabe benutzt werden.

## **5. Konsequenzen**

Natürlich wirkt die Anwendung dieser verschiedenen und z. T. gekoppelten KI-Programme auf einen eigentlich recht simplen fiktiven Mordfall etwas wie das berühmte Schießen auf Spatzen mit Kanonen. Wir hatten auch nicht primär die Absicht zu zeigen, dass mit derartigen Programmen Fälle wie dieser lösbar sind. Aus der erwähnten Geschichte von Expertensystemen ist dies hinlänglich bekannt. Deutlich gemacht werden an diesem Exempel einer „KI-Kriminalistik“ sollten vielmehr die folgenden Aspekte, die auf einige sehr grundlegende Probleme und deren mögliche Lösungen verweisen; für eine derartige Verdeutlichung eignet sich ein inhaltlich nicht sehr kompliziertes Problem besonders gut.

(a) Zum einen sind KI-Programme dieser Art ein didaktisch hervorragend geeignetes Medium, um Studierende sehr unterschiedlicher Studiengänge in KI-Modellierungen einzuführen. Natürlich brauchen Studierende der Geistes- und Sozialwissenschaften nicht alle technischen Details der hier dargestellten Programme zu kennen, und sie brauchen auch nicht selbst programmieren zu können. Sie können jedoch an derartigen Programmen nicht nur die prinzipielle Funktionsweise von KI-Modellen lernen, sondern auch (was eigentlich noch wichtiger ist) lernen, bestimmte Texte formal zu rekonstruieren und ihre spezifische Logik zu verstehen. Das war insbesondere bei der zweiten Gruppe, die mit einem IN arbeitete, sehr deutlich. Damit ergibt sich die Möglichkeit, Textanalyse auf eine formale Weise anzugehen, die die klassischen Verfahren nicht obsolet macht, sondern ergänzt. Dass dies möglich und für die Studierenden sehr fruchtbar ist, haben wir mit einem speziell für die formale Analyse sozialwissenschaftlicher theoretischer Texte konstruierten Expertensystem zeigen können (vgl. Klüver u. a. 2001).

Jedoch auch in Informatikstudiengängen, in denen wir ebenfalls lehren, können derartige Programme sehr sinnvolle didaktische Möglichkeiten bieten. Die Einführung in KI-Methoden wird nicht immer vermittelt, findet jedoch bei vielen Studierenden großes Interesse. Selbst zu einfachen Kriminalgeschichten verschiedene Programme des hier geschilderten Typs zu konstruieren ist ein didaktisches Vorgehen, das sich nach unseren Erfahrungen sehr lohnen kann. Auf die engagierte Motivation der Studierenden kann man sich praktisch immer verlassen.

Gerade für diese Studierenden ist jedoch ein weiterer bildungsökonomischer Aspekt äußerst

wichtig. Informatikabsolventen sehen sich in wachsendem Maße dem Problem konfrontiert, dass immer mehr Softwarefirmen dazu übergehen, Teile ihrer Entwicklungsarbeit in kostengünstigen Länder wie vor allem Indien durchführen zu lassen; z. B. geht die Softwarefirma SAP davon aus, dass bis zu 25 % ihrer Entwicklungsarbeit in den nächsten Jahren nach Indien verlagert werden wird (Der Spiegel Nr. 48, 2003). Für deutsche Informatikabsolventen bedeutet dies, dass sie zusätzliche Qualifikationen neben den Basisqualifikationen erwerben müssen, um konkurrenzfähig zu bleiben. Hierzu gehört vor allem die Fähigkeit, formale Modelle für die Bearbeitung äußerst unterschiedlicher Probleme entwickeln zu können, die anschließend in Computerprogramme umgesetzt werden. Für das Training derartiger Fähigkeiten sind so scheinbar abseitige Themen wie die hier dargestellten hervorragend geeignet; Aufgaben wie die Entwicklung von Programmen, die literarische Kriminalfälle lösen können, zwingen die Studierenden, sich einer allgemeinen Logik der Konstruktion formaler Modelle bewusst zu werden.

(b) Methodologisch gesehen bedeutet der Einsatz verschiedener Programme für das gleiche Problem die Möglichkeit, gewissermaßen die „Robustheit“ dieser Programmtypen zu testen. Damit ist gemeint, dass sichergestellt werden muss, dass derartige Programme keine Artefakte produzieren, sondern im methodologischen Sinne valide sind. Die Algorithmen unserer verschiedenen Programme sind offenbar z. T. extrem unterschiedlich; dennoch gelang es allen Programmen, die gestellte Aufgabe gleichartig zu lösen. Da es die richtige Lösung war, können wir uns offenbar auf die Validität der Programme verlassen.

Wir haben dies Prinzip der methodischen Überprüfung verschiedener Programme bereits in einem gänzlich anderen Kontext verwendet, nämlich bei der Simulation bestimmter gruppendynamischer Prozesse (Klüver und Stoica 2003). Ein Zellularautomat, eine Kohonen-Karte und ein genetischer Algorithmus sollten jeweils gleiche Differenzierungsprozesse von sozialen Gruppen simulieren; ihre Ergebnisse wurden z. T. auch empirisch anhand realer Gruppen überprüft. Die Tatsache, dass die verschiedenen Programme immer das im wesentlichen gleiche Ergebnis generierten und dass dies auch zum großen Teil mit den realen Prozessen übereinstimmte, ist ein deutliches Indiz dafür, dass derartige Programme tatsächlich „robust“ sind.

(c) Unter kognitionswissenschaftlichen Aspekten schließlich zeigen diese Programme, dass offensichtlich sehr unterschiedliche Arten von KI-Modellen auf einen zweiten Blick gar nicht so verschieden sind. Dies kann man sehr gut an einem Vergleich des Expertensystems mit dem erweiterten interaktiven Netz erkennen. Die Gewichtsmatrix des IN wird auf die oben beschriebene Weise durch den heteroassoziativen Speicher generiert, und zwar auf der Basis von dessen vorhergegangenem Training. Das Ergebnis ist dann eine Matrix, in der z. B. McLaren durch eine 1 ein Motiv zugesprochen wird – nach der eingegebenen Information „Liebe“ natürlich. Logisch gesehen bedeutet dies, dass das Gesamtsystem nach dem Training des HS *über das Äquivalent der Regel im Expertensystem* verfügte „wenn Liebe, dann Motiv“. Entsprechend konnte es anschließend mit der Information genauso umgehen wie das Expertensystem HERCULE 1.

Praktisch heißt das, dass das Gesamtsystem HS/IN damit das Expertensystem generieren kann, wenn die trainierten Zuordnungen in entsprechende IF-THEN Regeln übersetzt werden. Der häufig immer noch beschworene Gegensatz zwischen symbolischer KI – regelbasierte Systeme – und konnektionistischer KI – neuronale – Netze erweist sich damit in gewisser Weise als überflüssig: Es ist eine praktische Frage *und keine grundsätzliche*, ob für bestimmte Zwecke ein Expertensystem wie HERCULE 1 mit expliziten Regeln konstruiert wird oder ob man es vorzieht, durch Training bestimmter Netzwerke wie den HS die Zuordnungen an Beispielen lernen zu lassen

und dann durch Übertragung auf ein IN oder eine Kohonen-Karte das Problem mit Hilfe der eingegebenen Informationen lösen zu lassen. Dies ist das Thema der gegenwärtigen zahlreichen Versuche, durch neuronale Netz „rule extractions“ für Expertensysteme durchführen zu lassen; an unserem kleinen Beispiel lässt sich sehr gut zeigen, wie dies im Prinzip funktionieren kann.

Vermutlich ist dies auch der Weg, auf dem Menschen lernen, Probleme wie dies hier zu lösen. Entweder erhalten sie explizite Regelvorgaben, wie die zweite Studentengruppe in unserer Geschichte, oder sie müssen selbst Regeln entwickeln, nämlich anhand von Beispielen. Dies ist, nebenbei bemerkt, die Bedeutung des Prinzips des exemplarischen Lernens. Didaktisch ist längst bekannt, dass der zweite Weg eigentlich der günstigere für das Lernen ist, weil die Lernenden dadurch gezwungen sind, sich der Logik des jeweiligen Problembereichs selbst bewusst zu werden. Von daher zeigt unser Beispiel, dass KI-Modelle sehr realitätsadäquat sein können, wenn es um die Modellierung menschlicher Lern- und Problemlösungsprozesse geht. Dass diese Programme auch sehr leistungsfähig sind, ist ohnehin bekannt und die verschiedenen HERCULE-Typen haben es noch einmal bestätigt.

Selbstverständlich wollen wir mit diesem Beispiel neuartiger Verwendungen von Computerprogrammen in der Lehre nicht andere Möglichkeiten als irrelevant abtun. Dies wäre extrem ungerecht bei den vielen auch praktisch sehr nützlichen Arten, neue Medien in den Lehr- und Lernalltag einzubeziehen. Hier sollte es darum gehen, zu zeigen, dass man neue Medien auch ganz anders einsetzen kann, als es üblicherweise geschieht. Wenn die hier demonstrierte Kombination von Künstlicher Intelligenz und textwissenschaftlicher Hermeneutik andere Kolleginnen und Kollegen motivieren sollte, es auch einmal damit zu versuchen, dann sind wir gerne behilflich.

## **Literatur**

- Favre-Bulle, B. 2001: Information und Zusammenhang. Wien – New York: Springer
- Klüver, J., Stoica, C. und Schmidt, J. 2001: Soziologische Theorien, akademische Lehre und der Computer – ein Erfahrungsbericht. *Soziologie* 2/2001
- Klüver, J. und Stoica, C. 2003: Simulations of Group Dynamics with different Models. *Journal of Artificial Societies and Social Simulation* 6, 4
- Lakoff, G. 1987: *Women, fire and Dangerous Things. What Categories reveal about the Mind.* London – Chicago: The University of Chacago Press
- Reichertz, J., 1994: Polizeiliche Expertensysteme: Illusion oder Verheißung? In: Hitzler, R., Honer, A. und Maeder, C. (Hrsg.): *Expertenwissen.* Opladen: Leske+Buderich
- Ritter, H. und Kohonen, T. 1989: Self-organizing semantic maps. *Biological Cybernetics* 61
- Stoica, C. 2000: Die Vernetzung sozialer Einheiten. Hybride interaktive neuronale Netze in den Kommunikations- und Sozialwissenschaften. Wiesbaden: DUV
- Time Life 1987: Die Hilfe eines elektronischen Experten. In: Time Life Redaktion: *Computer Verstehen.* Amsterdam: Time Life