

# Künstliche Chemie

Andre Skusa, Wolfgang Banzhaf, Jens Busch, Peter Dittrich, Jens Ziegler

**Unter einer *Künstlichen Chemie* (Artificial Chemistry) versteht man ein System, das mit seinen Komponenten und Interaktionsregeln das Paradigma der natürlichen Chemie zum Vorbild hat, jedoch von den tatsächlichen physikalischen Randbedingungen abstrahiert. Die Dynamik einer Künstlichen Chemie, die durch die Interaktionen der Moleküle und den Reaktionsalgorithmus bestimmt wird, ist dabei der anderer komplexer Systeme ähnlich. Somit können unterschiedliche komplexe Systeme als *Realisierungen* der gleichen *Organisation* verstanden werden. Diese *Organisation* gilt es, abstrakt und realisierungsunabhängig zu beschreiben. Von besonderem Interesse sind deshalb solche für natürliche und komplexe Systeme kennzeichnenden Phänomene wie *Selbstorganisation*, *Strukturbildung*, *Evolution* und *Informationsverarbeitung*. Die Zielsetzungen bei der Untersuchung Künstlicher Chemien lassen sich vereinfachend in *Modellbildung*, *Informationsverarbeitung* und *Optimierung* unterteilen. Dies umfasst sowohl grundagentheoretische Untersuchungen zur *Evolution des Lebens*, als auch anwendungsorientierte Ansätze, die eine Künstliche Chemie z.B. zur metabolischen Steuerung von Robotern oder zum automatischen Beweisen einsetzen.**

## 1 Künstliches Leben und Künstliche Chemie

Eine der in vielen Facetten auftretenden Grundkonstanten wissenschaftlichen Nachdenkens ist die Frage nach dem, was *Leben* ist und wie man es systematisch erfassen und beschreiben kann. Obwohl erst in der jüngeren Wissenschaftsgeschichte diese Frage an Bedeutung gewann, zieht sie sich heute wie ein roter Faden durch ganze Forschungsbereiche. *Künstliches Leben* (Artificial Life, KL) ist nun ein interdisziplinärer Versuch, verschiedene Forschungsgebiete unter dieser Fragestellung zu vereinen. Doch nicht nur die gewachsene Bedeutung der Fragen nach Ursprung und Funktion des Lebens sind kennzeichnend für die KL-Forschung, sondern auch die Auseinandersetzung mit dem bislang vorherrschenden mechanistischen Weltbild.

Es ist eine Paradigmenverschiebung zu beobachten, die *komplexe Systeme*, zu welchen insbesondere Organismen zählen, im Allgemeinen betrifft. Charakteristisch für die neue Sichtweise ist die *Emergenz* globalen Verhaltens aus der Interaktion der Teilsysteme bzw. Systemkomponenten, die jeweils relativ *einfachen* und nur lokal wirksamen Regeln folgen. Es entsteht ein Verhalten, das sich einer Untersuchung durch Zerlegung des Systems in seine Einzelteile entzieht. Die Systembestandteile erklären sich nicht ausschließlich aus sich selbst heraus, sondern in entscheidender Weise durch ihre *Funktion* im Gesamtsystem und ihr Zusammenspiel mit den anderen Teilsystemen.

Ein weiteres Merkmal vieler durch KL geprägten Forschungsansätze ist die Unterscheidung zwischen den materiellen Eigenschaften und der Konfiguration der Systemkomponenten. Konkrete Systeme lassen sich auf einem Abstraktionsniveau beschreiben, mit dem die *Dynamik* der ablaufenden Prozesse erfasst wird. Unterschiedliche Systeme werden dadurch vergleichbar. Biologie wird unabhängig von ihrer konkreten materiellen Realisierung untersucht. Ein System ist *lebendig* nicht aufgrund besonderer Eigenschaften der Materie, aus der es besteht, sondern durch die *Organisation* dieser Materie, also die Beziehungen der Systemkomponenten oder ganzer Systeme zueinander.

Von besonderer Bedeutung für die Erforschung von lebendigen Systemen ist ein genaues Verständnis der von Darwin aufgestellten Evolutionstheorie. Das Prinzip von Veränderung und Auslese, entdeckt aufgrund der Beobachtung des Verhaltens von Populationen in einer begrenzten Umwelt, ist erklärungsmächtig, wirft aber auch neue Fragen auf. So scheint es auf anderen Ebenen, z.B. der molekularen Ebene, ebenso gültig zu sein und kann auch in höherstufigen (sozialen oder kulturellen) Systemen beobachtet werden. Trotzdem bleiben Fragen offen: Wie entstehen die Voraussetzungen, also die Entitäten, die variiert und selektiert werden? Wie entstehen grundlegende Wechselwirkungen, die sich vielleicht sogar gegenseitig bedingen? Mit einem abstrakten Modell, das die grundsätzlichen Eigenschaften von Systemen, in denen Evolution vorkommt, abbildet, ließen sich mit Hilfe von Simulationen die theoretischen Bedingungen und Gegebenheiten für eine Evolution des Lebens näher beschreiben.

Solch ein Modell zu erstellen ist die Grundmotivation einiger der Forschungsaktivitäten mit Künstlicher Chemie (Artificial Chemistry, KC). Unter Zuhilfenahme der Abstraktion vom natürlichen Vorbild der molekularen Prozesse werden meist mit Simulationen die dynamischen Vorgänge innerhalb eines komplexen Systems untersucht. Gesucht werden Strukturen, die in der Lage sind, sich und andere Strukturen, die im Laufe des Experimentes entstehen, zu verändern oder zu erhalten, sowie komplexere Strukturen auszubilden. Es handelt sich also um Formen der *Organisation*, *Selbstorganisation* und *Selbsterhaltung*, und die Bedingungen für die Entstehung solcher Strukturen.

Doch die biologisch motivierte KC-Forschung stellt nur einen Teil des Gebietes dar. Abbildung 1 gibt einen Überblick über das gesamte Spektrum der KC-Forschung.

Im Bereich der *Modellbildung* finden sich unter anderem die eben dargestellten Versuche, in biologischen Systemen oder der Darwinschen Evolutionstheorie eine theoretische Grundlage zu erarbeiten. Darüberhinaus sind auch andere Systeme, die prinzipielle Ähnlichkeiten mit chemischen Systemen aufweisen, von Interesse, so z.B. soziale Systeme oder parallele Prozesse. Die vermutete Verwandtschaft mit anderen Systemen führt über eine rein theoretische Betrachtung hinaus. In den

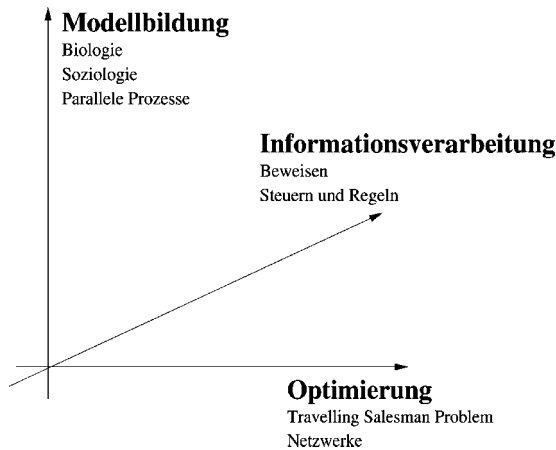


Abbildung 1: Schematische Darstellung der unterschiedlichen Einsatzgebiete von Künstlicher Chemie

Bereichen *Optimierung* und *Informationsverarbeitung* sollen die Erkenntnisse über die Funktionsweise dezentraler und paralleler Systeme praktisch nutzbar gemacht werden. So erhofft man sich mit einer nicht durch den Menschen beeinflussten Suche im Lösungsraum Fortschritte in klassischen Problemen der *Informationsverarbeitung*, wie z.B. dem *Automatischen Beweisen*. *Optimierungen* sind im Bereich von Netzwerken oder Parallelrechnern denkbar. Eine emergente Steuerung (*emergent computing*) könnte dezentral mit einem Minimum an Verwaltungsaufwand auskommen.

Die KC-Forschung in Dortmund beschäftigt sich sowohl mit den grundagentheoretischen und biologischen Fragestellungen, als auch mit praktischen Anwendungen. Im Folgenden wird zunächst anhand eines einfachen Beispiels die grundsätzliche Funktionsweise einer Künstlichen Chemie, so wie sie bei uns in allen Teilprojekten verwendet wird, vorgestellt. Anschließend werden Einblicke in die theoretische Untersuchung einer Binärstring-Chemie und in Anwendungen als automatische Beweiser und zur Robotersteuerung gegeben. Im Ausblick wird kurz gezeigt, was aus unserer Perspektive in Zukunft zu erwarten ist.

## 2 Allgemeine Form einer Künstlichen Chemie

Ausgangspunkt des hier verwendeten KC-Ansatzes ist die Modellvorstellung eines gut gerührten Reaktors mit einem Zu- und einem Abfluss (Abb. 2). Das Reaktionsgemisch in einem Reaktor wird als *Population* – auch der Begriff *Suppe* ist öfter zu finden – bezeichnet. Die Population besteht aus Molekülen bzw. Objekten, die jederzeit miteinander kollidieren und reagieren können. Eine einfache Form der zugrunde liegenden Reaktionsgleichung ist:

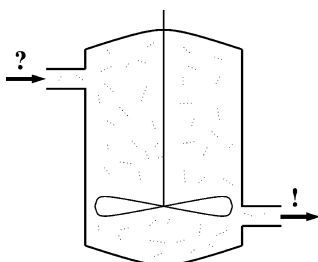


Abbildung 2: Gut gerührter Behälter mit Zu- und Abfluss als Modell

Hierbei entsteht aus der Kollision der Objekte  $s$  und  $s'$  das Reaktionsprodukt  $s''$ , die reagierenden Stoffe werden nicht aufgebraucht. Neue Moleküle entstehen entweder durch die Reaktionen im Inneren des Behälters oder werden von außen zugeführt. Wird die Kapazitätsgrenze überschritten, so erfolgt ein Abfluss durch Entfernung eines beliebigen Moleküls.

Es sind drei Komponenten einer Künstlichen Chemie, die festgelegt werden müssen: Die Repräsentation der Moleküle, das Reaktionsverhalten und das Verhalten des Reaktors. Etwas formaler betrachtet besteht eine Künstliche Chemie aus:

1. **Einer Objektmenge  $S$ :** Dabei kann es sich um abstrakte Symbole [21], Zeichenfolgen [4,16], Lambda-Ausdrücke [13], Binärstrings [5,11], Zahlen [7] oder Beweise [14] handeln. Ein Beispiel für eine implizite Definition sind die natürlichen Zahlen größer als zwei:  $S = \{2,3,4,\dots\}$ .
2. **Einer Regelmenge  $R$ , die die Interaktionen zwischen den Objekten beschreibt:** Diese die Reaktion bestimmenden Regeln hängen von der Wahl der Objekte und der Aufgabe, die das System erfüllen soll, ab. Es kann zwischen expliziten und impliziten Regeln unterschieden werden. Während bei einer expliziten Regel für jede Kombination zweier Objekte das Reaktionsprodukt festgelegt werden muss, bieten sich für eine implizite Reaktionsregel viele Möglichkeiten: Formale Automaten [9], Lambda-Kalküle [13], Turing-Maschinen [20], Matrixmultiplikation [5], einfache arithmetische Operationen [7] oder dergleichen mehr.
3. **Einem Algorithmus  $A$ , durch den der Systemablauf festgelegt wird:** Der Algorithmus bestimmt, wann welche Objekte miteinander reagieren, was mit dem Reaktionsprodukt geschieht, und welche Randbedingungen eingehalten werden müssen. Denkbar sind Simulationen eines gut gerührten Behälters ohne Topologie [4,11,13], mit einer euklidischen Topologie [21] (ähnlich den *zellären Automaten*), mit einem kontinuierlichen dreidimensionalen Raum [24] oder einer selbstorganisierenden Topologie [10].

Die prinzipielle Funktionsweise einer Künstlichen Chemie wird im Folgenden anhand eines einfachen Beispiels, einer *Zahlen-Divisions-Chemie*, erläutert.

Dabei handelt es sich um eine KC, die Berechnungen durchführt, welche aus einer zufälligen Anfangsbelegung mit natürlichen Zahlen Primzahlen erzeugt. Die *Objekte* sind also Zahlen eines festzulegenden Wertebereiches. Mit Hilfe der *Regeln* wird überprüft, ob zwei zufällig herausgegriffene Objekte  $a$  und  $b$  teilerfremd sind. Ist  $a$  durch  $b$  teilbar, sind das Divisionsergebnis  $c = a/b$  und der Teiler  $b$  das Reaktionsprodukt ( $a$  wird gelöscht) und werden zurück in die Population gegeben. Analog wird mit dem Fall, dass  $b$  durch  $a$  teilbar ist, verfahren. Wenn keine Teilung möglich ist, bezeichnet man dies als *elastische Kollision*, bei der beide Objekte unverändert in der Population bleiben. Als Reaktionsgleichungen geschrieben ergibt sich:

$$a + b \rightarrow c + b, \text{ mit } c = \frac{a}{b} \quad (2)$$

wenn  $a$  durch  $b$  teilbar ist.

$$a + b \rightarrow c + a, \text{ mit } c = \frac{b}{a} \quad (3)$$

wenn  $b$  durch  $a$  teilbar ist.

Andernfalls ist die Kollision von  $a$  und  $b$  elastisch, und es gilt:



Der einfachste *Reaktionsalgorithmus* wendet nach der Initialisierung die Regeln auf zwei zufällig gezogene Objekte an und stoppt bei gewünschtem Abbruch oder einer Primzahlkonzentration von 100%, wobei  $P$  eine Population über der Molekülmenge  $S$  der Populationsgröße  $M$  ist:

Reaktionsalgorithmus der Zahlen-Divisions-Chemie:

```

while  $\neg$  abbruch() do
  posA  $\leftarrow$  randomInt(1, M)
  posB  $\leftarrow$  randomInt(1, M)
  if  $P[\textit{posA}] \bmod P[\textit{posB}] = 0$ 
     $P[\textit{posA}] \leftarrow P[\textit{posA}] / P[\textit{posB}]$ 
  fi
  if  $P[\textit{posB}] \bmod P[\textit{posA}] = 0$ 
     $P[\textit{posB}] \leftarrow P[\textit{posB}] / P[\textit{posA}]$ 
  fi
od

```

Der Fortschritt in der Zeit wird in *Generationen* gemessen. In einer Generation werden  $M$  Kollisionen – entsprechend der Populationsgröße – durchgeführt. Ergänzungen des Algorithmus sind nötig, wenn z.B. eine bestimmte Topologie angewendet wird, und Reaktionen nur noch mit einigen Elementen in der Nachbarschaft stattfinden. In Abb. 3 ist das Ergebnis eines typischen Laufes zu sehen.

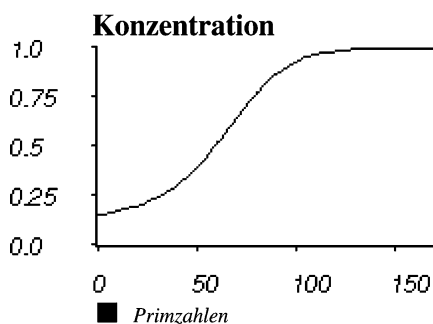


Abbildung 3: Ausschnitt aus dem Ausgabefeld des Primzahlengenerators, der im WWW unter <http://ls11-www.cs.uni-dortmund.de/alife/primegen> zu finden ist.

### 3 Selbstorganisation und Evolution in Binärstring-Reaktionssystemen

Eine Möglichkeit, mehr über die Eigenschaften und dynamischen Vorgänge innerhalb eines komplexen Systems interagierender Komponenten zu erfahren, ist eine KC mit *Binärstrings*, Ketten fester oder variabler Länge aus Nullen und Einsen. Diese Ketten können auf vielfältige Art und Weise miteinander verknüpft werden und daraus ein Reaktionsprodukt erzeugen.<sup>1</sup> Dazu verwenden wir Reaktionen, die auf Matrixmultiplikationen (*Matrixreaktion*) [5] oder formalen Automaten (*Automatenreaktion*) [9] basieren. Beide Methoden haben gemeinsam, dass bei einer Reaktion zweier Strings  $s, s'$  der eine zu einem aktiven Operator transformiert (gefaltet) wird, der dann auf den zweiten String  $s'$  angewendet wird, um das Reaktions-

produkt zu bilden. Im Vordergrund der Untersuchungen stehen die abstrakten Aspekte einer KC. Ziel ist es, Prozesse zu modellieren, die auch in anderen natürlichen Systemen (z.B. im Rahmen einer präbiotischen Evolution von RNA-Molekülen) vorkommen.

Im Falle der Matrixreaktion wird dazu  $s$  zu einer Matrix gefaltet, die multiplikativ auf Teilstücke von  $s'$  angewendet wird. Aus den Resultatstrings wird dann das Reaktionsprodukt  $s''$  zusammengesetzt.

In der Automatenreaktion wird aus  $s$  ein endlicher Automat abgeleitet. Der Reaktor ist gefüllt mit Binärstrings konstanter Länge  $N = 32$ , von denen jeweils zwei als Operator ( $s$ ) und Operand ( $s'$ ) zu einer Kollision  $s + s' \rightarrow s''$  zufällig ausgewählt werden. Für die Reaktion gibt es nun einen Automaten, der eine eindeutige Reaktion zwischen den beiden Binärstrings definiert und so konstruiert ist, dass er in jedem Fall hält. Der Reaktionsablauf entspricht dem Vorgehen bei Turing-Maschinen, allerdings sind hier die Bänder und Zustandsübergänge endlich. Der Operator ( $s$ ) wird in 4-Bit-Sequenzen unterteilt, die als Befehle interpretiert werden. Diese Befehle werden auf den Operandenstring ( $s'$ ) angewendet. Auf diesen Binärstring werden auch die Ergebnisse zurückgeschrieben. Alle nötigen Informationen zum Ablauf des Automaten, z.B. was die 4-Bit-Befehle bedeuten oder an welcher Stelle der Strings sich gerade die Ausführung befindet, sind in *Registern* des Automaten gespeichert.

Der algorithmische Ablauf ist bei beiden Reaktionstypen gleich: Zwei Binärstrings werden zufällig gezogen, die Kollision durchgeführt, und an zufällige Positionen in die Population zurückgegeben. Da die Populationsgröße konstant bleibt, werden dadurch zwei andere Objekte gelöscht.

Typische Experimente basieren auf Populationsgrößen zwischen  $M = 10^2$  und  $M = 10^6$  und laufen in einer Zeitspanne von bis zu 10000 Generationen. Bei einer solchen Datenmenge (bis zu  $10^{10}$  Kollisionen) ergeben sich neue Fragestellungen zu Interpretation und Betrachtungsebene der Ergebnisse. Lassen sich in KC-Systemen mit wenigen Molekülen die möglichen Reaktionen noch explizit angeben und die experimentellen Ergebnisse mit Differentialgleichungs-Modellen vergleichen (wie es für die 4-Bit-Matrixreaktion durchgeführt wurde [5]), liegt das Problem bei der Automatenreaktion in der großen Menge vieler im Detail vorliegenden Daten. Auf der *Mikroebene* sind zwar alle Daten vorhanden, die gebraucht werden, um die Entwicklung einzelner Moleküle und deren Einbindung in *Reaktionsnetzwerke* zu verfolgen. Allerdings steht nicht von vornherein fest, welche Daten wichtig für die Beobachtung sind, kommt es doch gerade auf Strukturen an, die erst im Laufe des Experimentes entstehen. Somit werden sowohl Methoden auf der *Makroebene* benötigt, die einen Anhaltspunkt für interessante Entwicklungen geben, als auch Tools, die auf einer *mesoskopischen Ebene* eine Verbindung zwischen genauester und größter Auflösung bieten [12].

Zu den Mitteln, auf der makroskopischen Ebene einen Überblick über das System zu bekommen, gehören Maßzahlen, wie die *Diversität* und die *Produktivität*. Die Diversität bezeichnet hier die absolute Anzahl unterschiedlicher Objekte. Finden elastische Reaktionen statt, so steht die Produktivität für den Anteil der nicht-elastischen Reaktionen. Produktiv sind diese Reaktionen, weil sie ein Reaktionsprodukt erzeugen.

Sucht man nach Entwicklungen, die Prinzipien der Selbstorganisation entsprechen oder eine Evolution von Strukturen aufweisen, sollten die Diversitäts- und Produktivitätskurven Sprünge oder Unregelmässigkeiten aufweisen. Oft aber verläuft z.B. die Diversitätskurve abnehmend und zeigt an, dass die

<sup>1</sup> Nähere Informationen zu KC mit Binärstrings gibt es auf der Homepage des von der DFG unter Ba 1042/2-1 und 2-2 geförderten BinSys-Projektes: <http://ls11-www.cs.uni-dortmund.de/alife/prj-binsys.html>

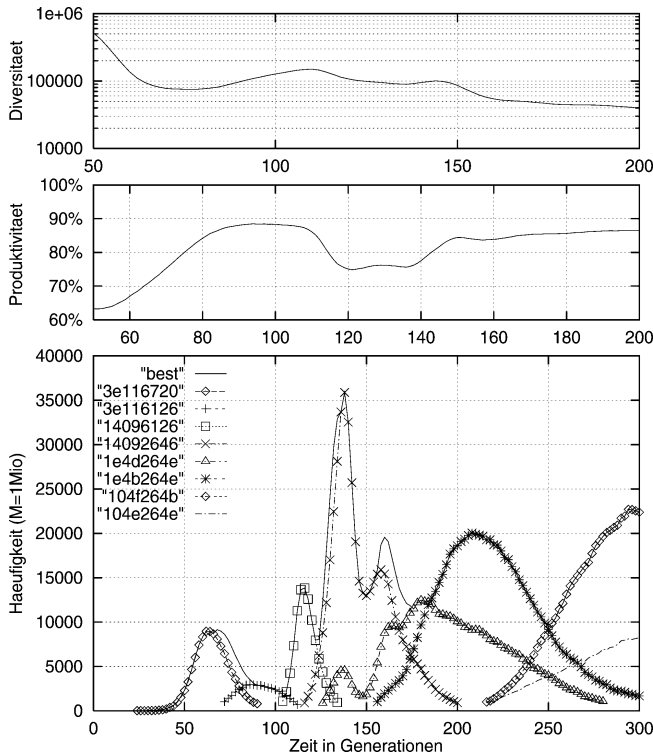


Abb. 4: Evolutive Phase in einem Experimentausschnitt. Unter bestimmten Randbedingungen wird die Systemdynamik aufrechterhalten und nicht von einigen wenigen Objekten bestimmt. In diesem Ausschnitt sind in den beiden oberen Kurven auf makroskopischer Ebene Indizien für die spontane Entstehung vorher nicht vorhandener Strukturen zu erkennen. Produktivität und Diversität nehmen eine Zeitlang zu und verhalten sich insgesamt unregelmäßig ohne auf einen Zielattraktor zuzulaufen. Einzelheiten der entstehenden und wieder verschwindenden Organisationen und Strukturen sind im unteren Bereich zu sehen.

Population, häufig schon nach kurzer Zeit, von einem oder wenigen Objekten beherrscht wird.

Unter bestimmten Bedingungen jedoch lässt sich ein Systemverhalten erzeugen, das *evolutive Phasen* beinhaltet. Ein Beispiel dafür ist in einem Ausschnitt aus einem Lauf in Abb. 4 zu sehen. Die dargestellten Objekte stehen zusammenfassend für ganze Gruppen von Objekten, die erst im Laufe des Experimentes mit hohen Konzentrationen entstehen und nach einiger Zeit wieder verschwinden. Die steigende Produktivitätsrate und das gebremste Absinken der Diversitätskurve auf der makroskopischen Ebene sind Indizien für diese Entwicklungen.

Solche Experimente zeigen, dass ein Binärstringsystem zu komplexem Verhalten und zur Selbstorganisation fähig ist. Es bilden sich nicht-triviale Organisationsstrukturen, die zum Startzeitpunkt noch nicht vorhanden waren. Eine „natürliche“ Evolution findet ohne die Anwendung externer künstlicher Operatoren - beispielsweise Mutation und Selektion - in Bezug auf eine *Fitness* statt. Es sind die Objekte selbst, die ihre Variationen durch Veränderung in den Reaktionsprodukten erzeugen. Bemerkenswert ist, dass die Objekte dadurch ihren *eigenen Veränderungsmechanismus* modifizieren. Binärstringsysteme ermöglichen aufgrund ihrer Handhabbarkeit und Effizienz das Studium derartiger *metaevolutiver Vorgänge*.

## 4 Automatischer Theorembeweiser

In diesem Abschnitt wird eine Anwendung der KC gezeigt, bei der es weniger um den Nachweis des Aufbaus oder der Selbsterhaltung von Organisationsstrukturen [6, 14, 13] geht. Vielmehr wird hier das Potential der KC in Bezug auf reale Problemlösung von – aus der Sicht der Informatik – traditionellen Aufgabenstellungen demonstriert. Das automatische Beweisen kann als der älteste Zweig der Künstlichen Intelligenz angesehen werden (siehe [19], vgl. [22]). Dabei geht es um die folgende Fragestellung:

Gegeben ist eine Menge  $\mathcal{A}$  von Aussagen in wohlgeformter Struktur sowie eine weitere wohlgeformte Aussage  $T$ . Es soll festgestellt werden, ob  $T$  eine logische Konsequenz aus  $\mathcal{A}$  ist (ob  $T$  also aus  $\mathcal{A}$  folgt,  $\mathcal{A} \models T$ ). Dies ist genau dann der Fall, wenn  $T$  der Aussagenmenge  $\mathcal{A}$  unter keinen Umständen widerspricht.<sup>2</sup> Sei  $\mathcal{A} = \{A_1, \dots, A_n\}$ . Dann werden die  $A_i$  *Axiome*, die Gesamtheit  $\mathcal{A}$  *Theorie* genannt. Umformungen, die gültige Aussagen in neue, ebenfalls bezüglich der Theorie gültige Aussagen überführen (abbilden), werden *Inferenzregeln* genannt. Jede durch eine solche Regel abgeleitete Aussage ist daher konsistent zur zugrunde liegenden Theorie.

Lässt sich eine Aussage  $T$  durch (wiederholte) Anwendung von Inferenzregeln auf  $\mathcal{A}$  sowie auf bisher von  $\mathcal{A}$  abgeleitete Folgerungen ableiten, so heißt  $T$  *Theorem* des Axiomensystems  $\mathcal{A}$ . Das Protokoll der einzelnen Ableitungen ist der zugehörige Beweis.

Als Formalismus wird die Prädikatenlogik 1. Ordnung eingesetzt. Sie bietet die Möglichkeit, Objekte mit ihren Eigenschaften und ihren Beziehungen untereinander zu repräsentieren. Die logischen Aussagen sind strukturiert in *Klauseln* als Disjunktion von *Literals*, die ihrerseits atomare Ausdrücke oder deren Negation sind. Literale können entweder erfüllt (wahr) oder unerfüllt (falsch) sein. Es handelt sich daher um eine 2-wertige Logik. Im vorliegenden System kommen zwei Inferenzregeln zum Einsatz. Die *Schnittregel* (cut-rule) basiert auf der Schlussweise des verallgemeinerten Modus Ponens:

$$\frac{A \vee B \quad \bar{A} \vee C}{B \vee C}$$

Es bezeichnen  $A, \bar{A}, B, C$  Literale,  $\bar{A}$  die Negation von  $A$ . Abb. 5 verdeutlicht die Wirkungsweise. Die zweite Inferenzregel ist die Faktorregel, die Klauseln unter Beibehaltung der Aussage komprimiert.

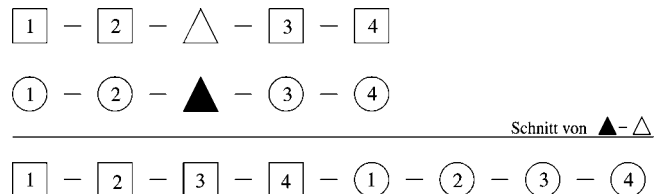


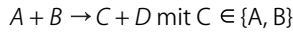
Abbildung 5: Wirkungsweise der Schnittregel. Die geometrischen Objekte symbolisieren Literale, deren Verbindung Klauseln. Falls zwei Literale  $L_1$  und  $L_2$  mit entgegengesetzter Polarität unifizierbar sind (hier die beiden Dreiecke), liefert die Schnittregel eine Klausel, die alle Literale außer  $L_1$  und  $L_2$  enthält.

Folgt  $T$  aus Axiomensystem  $\mathcal{A}$ , so kann das bewiesen werden durch die Widersprüchlichkeit der Aussagenmenge  $(\mathcal{A} \cup \bar{T})$ . Insgesamt wird also das Resolutionskalkül mit der Widerlegungsstrategie benutzt. Dieses Beweissystem ist konsistent und vollständig [18, 23].

<sup>2</sup> Mit Begriffen aus der Prädikatenlogik formuliert: Es gibt kein Modell, in dem  $\mathcal{A}$  gültig ist,  $T$  jedoch nicht.

#### 4.1 Aufbau des Reaktors

Jedes Molekül repräsentiert eine Klausel begrenzter Länge. Folgendes implizite Reaktionsschema zwischen Molekülen  $A$  und  $B$  liegt dem Reaktor zugrunde:



Produkt  $C$  ist dabei zufällig einer der Reaktionspartner  $A$  und  $B$ , während Produkt  $D$  die Klausel ist, die durch Resolution<sup>3</sup> oder Faktorisierung der durch  $A$  und  $B$  repräsentierten Klauseln entsteht. Ist jedoch die Anwendung dieser beiden Inferenzregeln nicht möglich, so wird die Reaktion *elastisch* genannt:  $A$  und  $B$  reagieren nicht miteinander.

Die das Problem kodierenden Klauseln bilden die Menge der *Startklauseln*. Der Reaktor wird mit  $M$  Ausfertigungen einer jeden befüllt.  $M$  bezeichnet dabei die *Multiplizität* und bestimmt die Reaktorgröße. Im Falle einer elastischen Reaktion erfolgt ein Zufluss von genau einer – zufällig ausgewählten – Startklausel, die einen der beteiligten Reaktionspartner ersetzt, so dass die Größe des Reaktors konstant bleibt.

Der Reaktionsprozess wird gestoppt, sobald ein Produkt die leere Klausel beschreibt. In diesem Fall wurde ein Widerspruch erreicht.

#### 4.2 Resultate

Wie alle automatischen Theorembeweiser, die auf dem Resolutionsverfahren mit Widerlegungsstrategie beruhen, modelliert der Reaktor einen semi-entscheidbaren Prozess: Wenn die zu beweisende Aussage  $T$  tatsächlich aus der Theorie gefolgert werden kann, ist mit einem Ende der Reaktionen zu rechnen. Ist jedoch eine bestimmte Anzahl von Reaktionen überschritten, so ist von der Unmöglichkeit einer Beweisfindung auszugehen und der Reaktor zu stoppen.

In allen durchgeführten Untersuchungen wurde ein Beweis gefunden, sofern dieser theoretisch überhaupt möglich war. Die Anzahl der dazu benötigten Reaktionen ist abhängig von der Komplexität des konkreten Problems sowie der Größe des Reaktors (Abb. 6). Ein Reaktor mit einer hohen Zahl von Molekülen benötigt mehr Reaktionen, andererseits verhindert ein Reaktor mit zuwenig Molekülen das Finden eines Beweises.

In Abb. 7 ist die Zahl der für die Lösung des Problems aus Tafel 1 benötigten Reaktionen zu 30000 Läufen für  $M=20$  klassifiziert zusammengestellt. Man kann zeigen, dass eine Poisson-Verteilung vorliegt. Ein komplexes Beispiel, der Beweis der Existenz des Rechtsinversen in der Gruppentheorie [18], führt dagegen bei gleicher Multiplizität zu einem höheren Verteilungsmittelwert.

- (1) Alle Hunde jaulen nachts.
- (2) Unruhige Schläfer besitzen nichts, was nachts jault.
- (3) Jeder, der eine Katze hat, besitzt keine Mäuse.
- (4) John hat entweder eine Katze oder einen Hund.

(Beweise:) Wenn John ein unruhiger Schläfer ist, dann besitzt er keine Mäuse.

Tafel 1: Beweisproblem aus M. Gini, <http://www-users.cs.umn.edu/~gini/5511/logic>, letzter Zugriff 19.10.99

<sup>3</sup> Das zu resolvierende Literalpaar  $(a, b) - a$  ist dabei Literal von  $A$  und  $b$  Literal von  $B$  – wird zufällig aus der Menge aller resolvierbaren Literal-Paarungen ausgewählt.

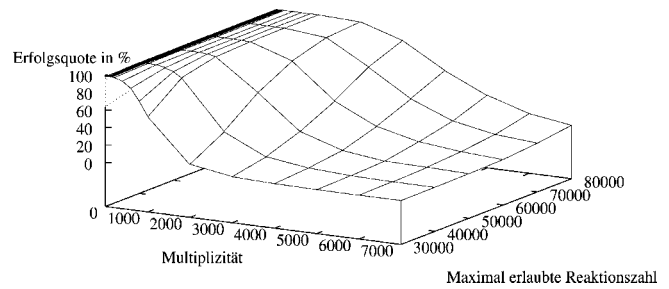


Abbildung 6: Erfolgswahrscheinlichkeit zu geg. Reaktorgröße und Reaktionslimit. Beweisproblem aus Tafel 1. Jeder Punkt repräsentiert das Mittel aus 50 Läufen.

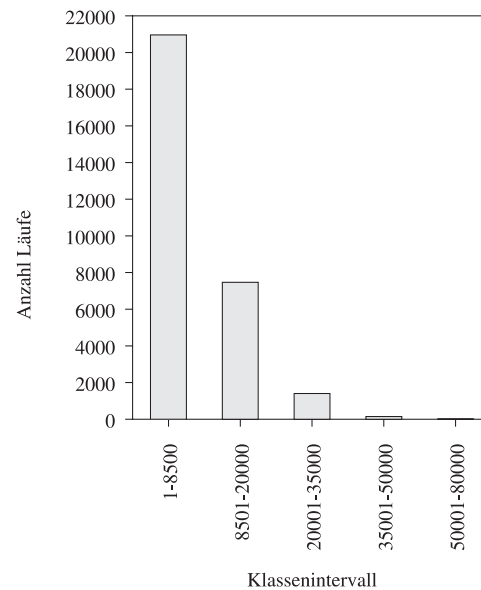


Abbildung 7: Klassifizierung der jeweils benötigten Reaktionen zur Beweisfindung. Beweisproblem aus Tafel 1,  $M = 20$ , 30000 Läufe.

Durch ihren inhärenten Parallelismus ersetzt die KC in diesem Beispiel herkömmliche Suchstrategien in großen Lösungsräumen.

## 5 Metabolische Steuerung eines Roboters

Ein weiteres Anwendungsbeispiel ist die Steuerung eines autonomen Roboters mit einer Künstlichen Chemie. Hierbei dient ein kleines Bakterium als natürliches Vorbild: *E. Coli*. Dieser Einzeller kann als ein natürliches Informationsverarbeitendes System von hoher Komplexität betrachtet werden [17]. Ein besonders interessanter Teil dieses Informationssystems ist die Bewegungssteuerung. Hierbei werden Informationen in Form von aufgenommenen Molekülen verarbeitet wobei die Informationen aus der Umwelt des Bakteriums in Bewegungsmuster transformiert werden. Diese Fortbewegungsart heißt *Chemotaxis*. *E. coli* verfährt hierbei nach dem simplen Prinzip, sich immer in die Richtung zu bewegen, in der die höchste Konzentration an Nahrung herrscht. Die Entscheidung, die aktuelle Bewegungsrichtung zu ändern, hängt dabei von der Konzentration bestimmter Metabolite ab und ist ein Beispiel für ein chemisches „Gedächtnis“.

Der Metabolismus des Bakteriums kann also als ein robustes, paralleles und verteiltes Informationssystem angesehen werden, das zu einem Input (einer Umweltsituation) einen spe-

zifischen Output (eine Bewegung) generiert. Dies ist exakt die Aufgabe, die ein autonom operierender Roboter zu lösen hat, und darum liegt es auf der Hand, diese Aufgabe mit Hilfe eines künstlichen Metabolismus, also einer Künstlichen Chemie, zu lösen.

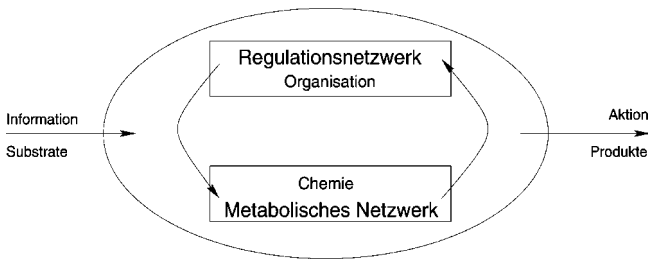


Abbildung 8: Abstrakte Sicht der Organisation des Informationssystems. Stoffwechsel- und Regulationsnetzwerk sind über Stoffflüsse bzw. Informationsflüsse miteinander gekoppelt.

Eine abstrakte Darstellung des bakteriellen Informationssystems ist in Abb. 8 zu sehen. Die Reaktionspfade bilden das metabolische Netzwerk, die Gesamtheit aller Pfade und Ihre Verknüpfungen bilden die regulatorische Einheit.

Die räumliche Verteilung von Substanzen in einem Reaktor soll in erster Näherung mit Hilfe von partiellen Differentialgleichungen dargestellt werden (obwohl das Innere eines Bakteriums alles andere als „ein Sack voller Enzyme“ ist [17]). Durch Kompartimentierung gelangt man zu einem Modell, das aus einem System gekoppelter gewöhnlicher Differentialgleichungen besteht, wobei jedes Kompartiment durch einen Konzentrationsvektor dargestellt werden kann. Um die Simulation des Gleichungssystems nicht zu aufwendig werden zu lassen, ist in den nachfolgend beschriebenen Experimenten nur ein Kompartiment verwendet worden. Die Reaktionen zwischen den im Reaktor vorhandenen Substanzen müssen nun noch so definiert werden, dass die Steuerung des Roboters anhand der Konzentration bestimmter Stoffe erfolgt. Die Konzentration dieser Substanzen (im Folgenden Motorsubstanzen genannt) muss das Ergebnis eines Reaktionsmechanismus sein, der Informationen über den Zustand der Umwelt in Form von Sensorsubstanzen in Motorsubstanzen transformiert. Die Veränderung der Konzentrationen wird über eine numerische Integration des Differentialgleichungssystems erreicht. Der in unserem Labor verwendete Miniaturroboter Khepera verfügt über acht Entfernungssensoren und acht Helligkeitssensoren, die in regelmäßigen Abständen auf dem Roboter montiert sind. Die Konzentration der Sensorsubstanzen wird nun durch einen zum aktuellen Sensorwert proportionalen Zu- bzw. Abfluss geregelt. Dadurch wird innerhalb des Reaktors die aktuelle Umweltsituation durch den Konzentrationsvektor der Sensorsubstanzen repräsentiert. Dadurch, dass die Drehzahl der zwei Motoren proportional zur Konzentration der Motorsubstanzen ist, bewegt sich der Roboter in die gewünschte Richtung.

Aufgabe des künstlichen Metabolismus ist es nun, die Konzentration der Motorsubstanzen mittels geeigneter Reaktionswege in einer Art und Weise zu ändern, dass der Roboter entsprechend einer vorher definierten Verhaltensweise auf Umweltsituationen reagiert. In diesem Experiment sollte der Roboter sowohl Hindernissen ausweichen als auch einem positiven Helligkeitsgradienten folgen. Diese beiden Aufgaben simultan zu erfüllen, bedeutet zum einen, für jede einzelne Aufgabe einen Reaktionsmechanismus bereitzustellen, der die Transformation der Sensorinformationen in eine entsprechende Kon-

zentration der Motorsubstanzen sicherstellt, und zum anderen, die beiden Reaktionswege in einer Art und Weise miteinander zu verknüpfen, dass in Konfliktsituationen kein Deadlock auftreten kann. Dies bedeutet eine „chemische“ Lösung des Problems, mehrere parallele und adaptive Prozesse zu einem kohärenten Verhalten zusammen zu fassen (siehe hierzu auch z.B. [8]).

Das eigentliche Problem aber besteht nun darin, eine geeignete Menge von Reaktionen und Reaktionspfaden zu finden, die so miteinander verknüpft sind, dass das gesamte Reaktionsnetzwerk sich den Anforderungen entsprechend verhält. Das bedeutet, dass die Künstliche Chemie in der Lage sein soll, auf verschiedene, a priori unbekannte Umweltsituationen so zu reagieren, dass das Verhalten des von ihr gesteuerten Roboters den definierten Zielen entspricht.

### 5.1 Das Design der Künstlichen Chemie

Um nun diesen Metabolismus zu erzeugen, bieten sich mehrere Wege: (i) die Struktur des Metabolismus wird von Hand erstellt, (ii) das Reaktionsnetzwerk mit den geforderten Eigenschaften wird mit evolutionären Methoden erzeugt oder (iii) ein durch Selbstorganisationsvorgänge emergierender Metabolismus besitzt geeignete Eigenschaften.

Die erste Möglichkeit bedeutet, einen Metabolismus (d.h. ein Netzwerk von Reaktionsmechanismen) in einer Art und

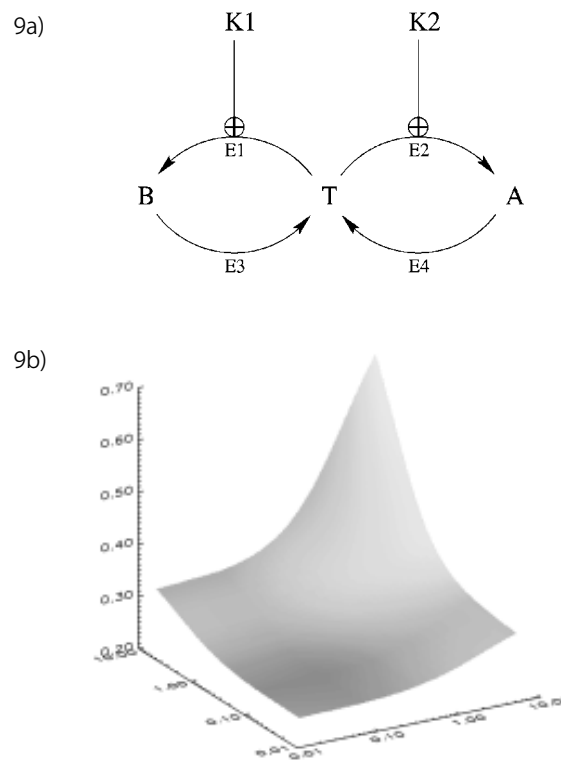


Abbildung 9: Schematische Darstellung eines Reaktionsmechanismus, dessen Funktionsweise einem logischen „AND“ gleichkommt. A; B und T sind die Substrate bzw. Produkte der durch die Enzyme E1 - E4 katalysierten Reaktionen. Die beiden Katalysatoren K1 und K2 bilden die Eingabewerte des Gatters, der Ausgang ist T. Die Abbildung 9b zeigt den Verlauf der Konzentration von T in Abhängigkeit von K1 und K2.

Weise zu „verdrahten“, dass die geforderten Eigenschaften erreicht werden. Die Möglichkeit (ii) realisiert eine strukturierte Suche nach geeigneten Metabolismen, bei der explizite Reaktionsmechanismen dahingehend bewertet werden, ob sie das

Verhalten des Roboters zufriedenstellend erzeugen. Mittels Evolutionärer Algorithmen (in unserem Falle Genetisches Programmieren) werden dann immer bessere Reaktionsnetzwerke evolviert bis schließlich das gewünschte Verhalten erreicht ist. Dies ist eine Weiterentwicklung bzw. Anwendung der in [4] beschriebenen *Metadynamics*. Experimente hierzu werden im Moment durchgeführt. Durch Selbstorganisationsvorgänge emergierende Metabolismen (Methode (iii)) entstehen durch evolutionäre Vorgänge im Reaktor selbst. Dies steht im Gegensatz zu den Evolutionären Algorithmen, bei denen Selektion, Rekombination und Fitness extern definiert werden. Das Problem, einen Metabolismus durch Selbstevolution zu erzeugen ist, dass bislang nicht geklärt ist, auf welche Weise der Fortschritt der Evolution von außen zu steuern ist, damit einerseits das Ziel erreicht wird und gleichzeitig noch von Selbstevolution gesprochen werden darf. Dies bedeutet jedoch, dass die Suche nach einem geeigneten Metabolismus einem *random search* in sehr großen Suchräumen gleichkommt und deshalb in endlicher Zeit wenig erfolgversprechend ist, denn die Möglichkeiten einer zufälligen Chemie sind – sogar bei einer Beschränkung auf eine endliche Objektmenge  $S$  – enorm.

Der tatsächlich implementierte Metabolismus wurde von Hand entworfen und wird in [25] im Detail präsentiert. Er realisiert ein Netzwerk von Reaktionen, die der Enzym-Substrat-Reaktion nachempfunden sind. Mit diesen Reaktionen lassen sich logische Funktionen chemisch realisieren, wie u.a. [2, 15], gezeigt haben. Abbildung 9 zeigt z.B. ein chemisches AND-Gatter.

Durch eine Verknüpfung dieser logischen Bausteine zu einem Reaktionsmechanismus kann ein bestimmtes Verhalten chemisch implementiert werden. Wenn mehrere verschiedene und z.T. auch gegenläufige Mechanismen existieren (man denke nur an ein Hindernis, welches den Weg zu einer Lichtquelle versperrt), so kann auch der *action selection*-Mechanismus auf diese Weise realisiert werden. Der Roboter wird so in die Lage versetzt, mit einer Künstlichen Chemie Informationen in Echtzeit zu verarbeiten und die gestellte Aufgabe zu erfüllen.

## 6 Zusammenfassung und Ausblick

In dem vorliegenden Beitrag haben wir versucht zu motivieren, dass ein wesentlicher Aspekt des interdisziplinären Forschungsgebietes Künstliches Leben die zugrundeliegenden Objekte und ihre Wechselwirkungen sind. Wir haben argumentiert, dass die sich hieraus ergebenden Fragestellungen oft im Bereich einer abstrahierten oder künstlichen Chemie untersucht werden können. Zugleich haben wir Anwendungen aufzuzeigen versucht, die von diesen Modellvorstellungen ausgehen.

Neben den hier erwähnten Anwendungen einer Künstlichen Chemie ist auch das in jüngerer Zeit mit großem Interesse studierte DNA-Computing ein Gebiet, das innerhalb von KC betrachtet werden kann. Im DNA-Computing [1] kommt es darauf an, Algorithmen so zu formulieren, dass sie in einem aus DNA-Molekülen bestehenden realen Reaktionssystem zu einem gewünschten Ergebnis führen. Man muss daher die Wechselwirkungen so aufbauen, dass die Lösungen stabile Produkte der vielen möglichen Reaktionen sind. Es werden also im DNA-Computing im Grunde dieselben Fragen gestellt wie in der KC. In der Simulation wird hierbei mit Vorteil auf KC-Methoden zurückgegriffen werden.

Auch die schon vor vielen Jahren formulierten Evolutionären Algorithmen [3] können als Spezialfall einer Künstlichen

Chemie betrachtet werden. Die Lösungen reagieren hierbei mit Operatoren zu neuen Lösungen, die, abhängig von gewissen Filterbedingungen, in die Population entlassen werden.

Interessanterweise ist das Paradigma der Künstlichen Chemie allgemein genug, auch andere Anwendungen zuzulassen. So ist es zum Beispiel denkbar, das Handelsgeschehen an der Börse mit einem solchen System zu modellieren. Hier werden dann Käufer und Verkäufer zu Reaktionsteilnehmern, der Handel (mit dem dazugehörigen Preis) zum Reaktionsprodukt. Selbstorganisation tritt dadurch auf, dass der Preis einer einzelnen Reaktion das Reaktionsverhalten anderer Teilnehmer beeinflusst.

Grundsätzlich lassen sich alle „Zweipunktwechselwirkungen“ unter den hier erwähnten Aspekten untersuchen. Besonders interessant sind diejenigen, bei denen eine Beeinflussung der Reaktionsdynamik durch die Ergebnisse vorliegt. Hierdurch wird eine Rückkopplung in das System eingeführt, die zu einer *autonomen Evolution* des Systems führt. Diese Eigenschaft bietet unserer Meinung nach ein hohes Potential zur Erklärung biologischer Evolutionsvorgänge, das es in Zukunft auszu-schöpfen gilt. Das Studium der Evolution solcher komplexen Systeme gehört zu den aufregendsten Gebieten, die die Forschung am Beginn des 21. Jahrhundert zu bieten hat.

## References

- [1] L. M. Adleman. Molecular Computation of Solutions to Combinatorial Problems. *Science*, 266: 1021-1024, 1994.
- [2] A. Arkin and J. Ross. Computational Functions in Biochemical Reaction Networks. *Biophysical Journal*, 67: 560-578, 1994.
- [3] T. Bäck, D. B. Fogel and Z. Michalewicz, editors. *Handbook of Evolutionary Computation*. IOP Publishing and Oxford University Press. Bristol, UK, 1997.
- [4] R. J. Bagley and J. D. Farmer. Spontaneous Emergence of a Metabolism, in *Artificial Life II, Proceedings*, ed. C. G. Langton MIT Press, Cambridge, MA, 1992.
- [5] W. Banzhaf. Self-replicating sequences of binary numbers - Foundations I and II: General and Strings of Length  $n = 4$ . *Biological Cybernetics*, 69, pages 269 - 281, 1993.
- [6] W. Banzhaf. Self-Organisation in a System of Binary Strings, in *Proceedings of Artificial Life IV*, ed. R. Brooks and P. Maes, MIT Press, Cambridge. Pages 109 - 118, 1994.
- [7] W. Banzhaf, P. Dittrich and H. Rauhe. Emergent computation by catalytic reactions. *Nanotechnology*, 7, pages 307 - 314, 1996.
- [8] R. A. Brooks. Coherent Behavior of Many Adaptive Processes, in *From Animals to Animats 3*, ed. D. Cliff and P. Husbands and J. Meyer and S. Wilson MIT Press, Cambridge, MA, 1994.
- [9] P. Dittrich. *Selbstorganisation in einem Binärstringsystem mit algorithmischen Sekundärstrukturen*. Universität Dortmund, Informatik, 1995.
- [10] P. Dittrich and W. Banzhaf. A topological structure based on hashing - emergence of a "spatial" organization, in *Fourth European Conference on Artificial Life (ECAL97)*, 1997.
- [11] P. Dittrich and W. Banzhaf. Self-Evolution in a Constructive Binary String System. *Artificial Life*, 4, pages 203 - 220, 1998.
- [12] P. Dittrich, J. Ziegler and W. Banzhaf. Mesoscopic Analysis of Self-Evolution in an Artificial Chemistry, in *Artificial Life VI*, ed. C. Adami and R.K. Belew and H. Kitano and C.E. Taylor MIT Press, Cambridge, MA. Pages 95 - 103, 1998.
- [13] W. Fontana. Algorithmic Chemistry, in *Artificial Life II: Proceedings of the Second Artificial Life Workshop*, ed. C. G. Langton and et al. Addison-Wesley, Reading MA. Pages 159 - 209, 1991.
- [14] W. Fontana and L. W. Buss. The barrier of objects: From dynamical systems to bounded organizations, in *Boundaries and Barriers*, ed. J. Casti and A. Karlqvist Addison-Wesley. Pages 56 - 116, 1996.
- [15] A. Hjelmfelt, E. D. Weinberger and J. Ross. Chemical Implementation of Neural Networks and Turing Machines. *Proc. Natl. Acad. Sci. USA*, 88, pages 10983 - 10987, 1991.

- [16] S. A. Kauffman. *The Origins of Order*. Oxford University Press, New York, 1993.
- [17] J. W. Lengeler. Die biologische Zelle als komplexes Informationssystem. *Biologie in unserer Zeit*, 426, pages 65 - 70, 1996.
- [18] D. W. Loveland. *Automated Theorem Proving: A Logical Basis*. North-Holland. Amsterdam, 1978.
- [19] G. F. Luger and W. A. Stubblefield. *Artificial intelligence and the design of expert systems*. The Benjamin/Cummings Publishing Company, 1989.
- [20] M. Thürk. *Ein Modell zur Selbstorganisation von Automatenalgorithmen zum Studium molekularer Evolution*. PhD thesis. Universität Jena, Naturwissenschaftliche Fakultät, 1993.
- [21] F. J. Varela, H. R. Maturana and R. Uribe. Autopoiesis: The organization of living systems. *BioSystems*, 5, pages 187 - 196, 1974.
- [22] A. N. Whitehead and B. Russel. *Principia Mathematica*. Cambridge University Press. London, 1950.
- [23] L. Wos, R. Overbeek, E. Lusk and J. Boyle. *Automated Reasoning: Introduction and Applications*. Prentice-Hall. Englewood Cliffs, NJ, 1984.
- [24] K.-P. Zauner and M. Conrad. Conformation-Driven Computing: Simulating the Context-Conformation-Action Loop. *Supramolecular Science*, 5, pages 791 - 794, 1998.
- [25] J. Ziegler, P. Dittrich and W. Banzhaf. Towards a Metabolic Robot Control System, in *Information Processing in Cells and Tissues.*, ed. M. Holcombe and R. Paton Plenum Press, 1998.

#### Kontaktadresse:

Universität Dortmund  
 Lehrstuhl für Systemanalyse  
 Fachbereich Informatik  
 D-44221 Dortmund  
 Tel.: +49-231-9700-951 (Sekretariat)  
 Email: <nachname>@LS11.cs.uni-dortmund.de  
 Internet: <http://ls11-www.cs.uni-dortmund.de>



**Wolfgang Banzhaf** vertritt seit 1993 das Fachgebiet „Grundlagen und Anwendungen der Informatik in den Ingenieurwissenschaften“ am FB Informatik der Universität Dortmund. Er studierte Physik an den Universitäten Stuttgart, München und Karlsruhe. Nach einer dreijährigen Assistentenzeit am Lehrstuhl für Synergetik der Univ. Stuttgart war er mehrere Jahre für die Firma Mitsubishi Electric in der Forschung tätig, 1989 bis 1992 in Japan und 1992 bis 1993 in den USA. Seine Forschungsinteressen liegen auf den Gebieten „Computational Intelligence“ und „Emergente Systeme“.



**Jens Busch** ist seit 1999 als wissenschaftlicher Mitarbeiter an der Universität Dortmund beschäftigt. Forschungsschwerpunkte liegen in den Gebieten Molekulares Rechnen, Multiagentensysteme, Genetisches Programmieren.



**Peter Dittrich** ist seit 1995 als wissenschaftlicher Mitarbeiter an der Universität Dortmund beschäftigt. Forschungsschwerpunkte liegen in den Gebieten Künstliche Chemie, selbstmodifizierende Maschinen, sowie den Anwendungen der Informatik in den Geisteswissenschaften.



**Andre Skusa** ist Student der Informatik und Philosophie und ist seit 1998 als studentischer Mitarbeiter an der Universität Dortmund beschäftigt. Interessenschwerpunkte liegen in den Gebieten Evolution, Selbstorganisation, komplexe adaptive Systeme.



**Jens Ziegler** ist seit 1997 als wissenschaftlicher Mitarbeiter an der Universität Dortmund beschäftigt. Forschungsschwerpunkte liegen in den Gebieten Künstliche Chemie, Metabolische Informationsverarbeitung, Genetisches Programmieren und Robotik.