

8. Eigenschaften von Netzwerken, Boolesche Zufallsnetzwerke, Artificial Chemistry

Netzwerktheorie: angewandte Graphentheorie
- untersucht Eigenschaften von (großen) Graphen, die in Anwendungsgebieten auftreten

Beispiele:

- metabolische Netzwerke
- soziale Netzwerke
- Verkehrsnetzwerke
- das WWW

zuerst in sozialen Netzwerken beobachtetes Phänomen:

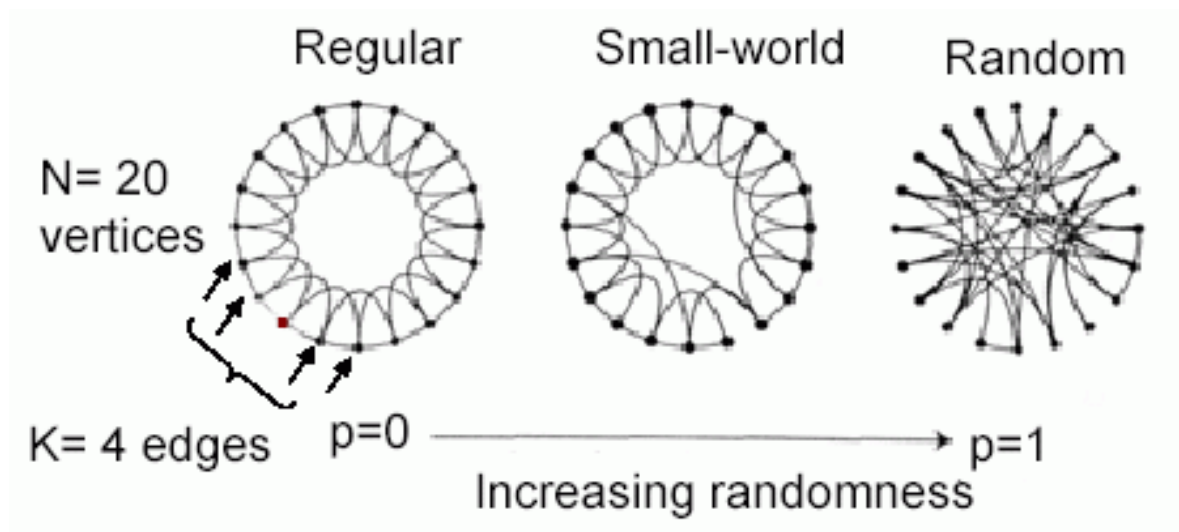
"Small World Effect"

bezieht sich (zunächst) auf kleine graphentheoretische Durchmesser:

- "six degrees of separation"
- Verknüpfung von 6 Bekanntschaften verbindet jeden mit jedem auf der Welt
- Experiment von Milgram 1967
- "die Welt ist doch klein!"
- dabei sind besonders die "schwachen Bindungen" wichtig ("The strength of weak ties"; Granovetter 1973)

graphentheor. Untersuchung von Watts & Strogatz (1998):

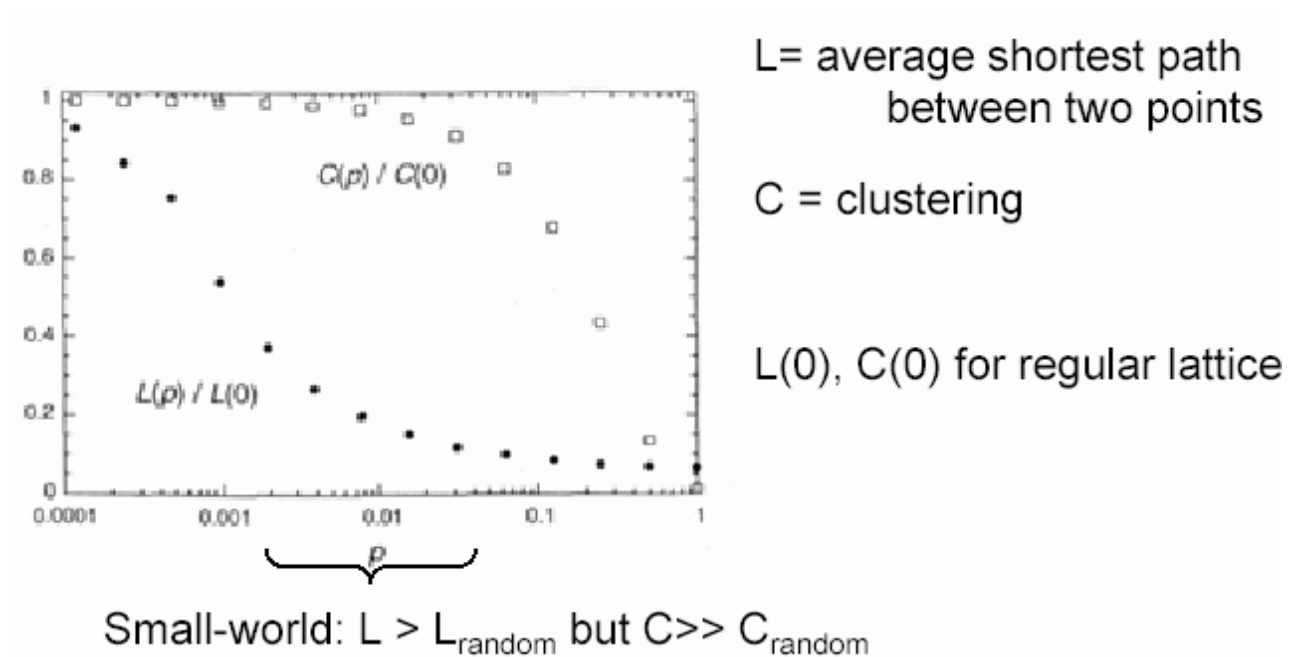
Zufallsgraphen, auf ringförmiger Knotenmenge definiert, einstellbarer Parameter p bestimmt, wieviele der zunächst regelmäßig zwischen eng benachbarten Knoten verlaufenden Kanten gelöst und zufällig neu verlegt werden



interessante Maße:

- *mittlere Pfadlänge* $L(p)$ (Länge des kürzesten Weges zwischen 2 Agenten im Netzwerk)
- *Clustering-Koeffizient* $C(p)$: misst, wieviele Nachbarn eines Agenten ihrerseits untereinander verbunden sind

Small World Effect:



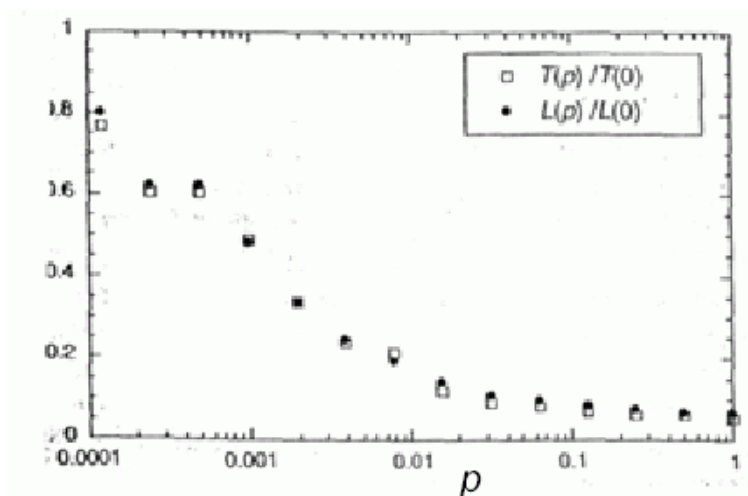
Messungen auch an realen Netzwerken

Beispiel: Netzwerk von Filmschauspielern – verbunden, wenn sie in einem Film zusammen gespielt haben

$$L = 3,6 > L_{\text{random}} = 3,0;$$

$$C = 0.79 \gg C_{\text{random}} = 0,00027 !$$

Auswirkungen z.B. auf die Seuchenausbreitung:



$T(p)$ = time required for global infection.

⇒ Gefahr der globalen Infektion deutlich größer in "Small World Netzwerken" als in reinen Zufallsnetzwerken

Skalenfreie Netzwerke

es sei $P(k)$ der Anteil der Knoten vom Grad k

wenn in großen Graphen gilt: $P(k) \sim k^{-c}$, so spricht man von einem skalenfreien (oder skaleninvarianten) Netzwerk

zusammenhängende, skalenfreie Netzwerke sind "unempfindlich" gegen zufälliges Löschen von Knoten (d.h. der Zusammenhang bleibt meistens erhalten), aber empfindlich gegen gezieltes Löschen von Knoten mit hohem Grad ("hubs").

Viele Small-World-Netzwerke sind auch skalenfrei.

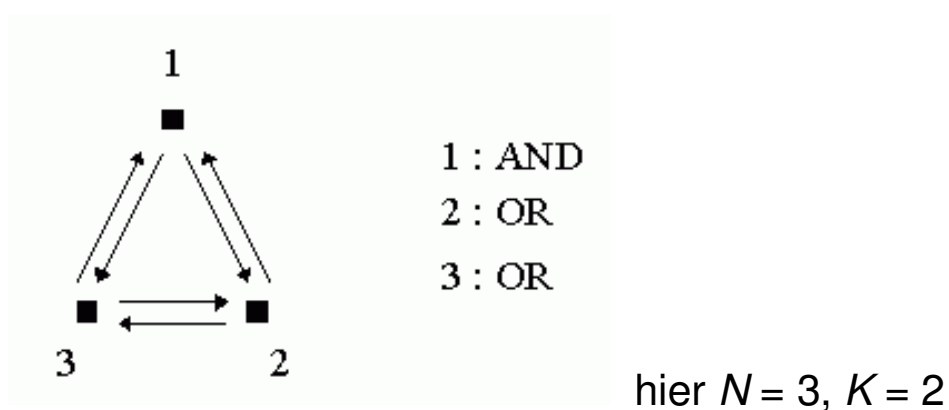
Gewöhnliche Zufallsgraphen (suche Knoten mit gleichverteilter Auswahl und verbinde diese) sind nicht skalenfrei!

Boolesche Zufallsnetzwerke

- nur 2 Enzymzustände: "aktiv" (1) und "inaktiv" (0)
- der Zustand jedes Enzyms hängt von K anderen Substanzen ab
- Abhängigkeit wird durch Boolesche Funktion beschrieben

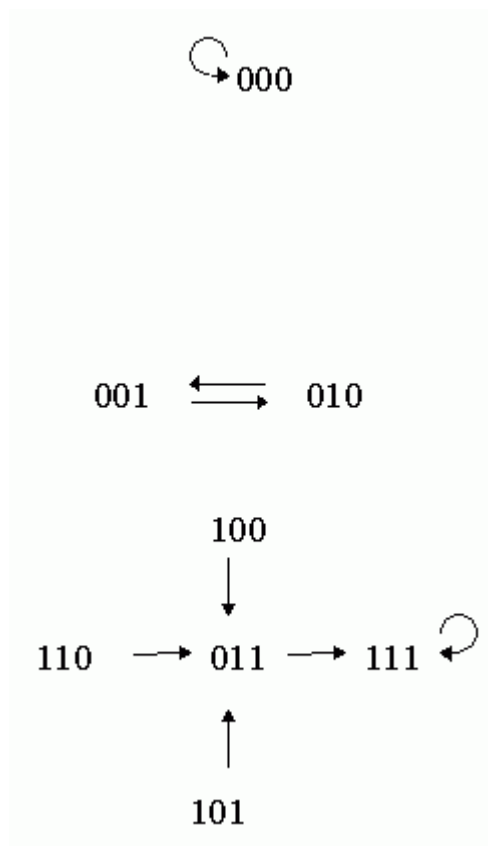
Beschreibung der Netzwerkstruktur durch Netzwerk von N zufällig verschalteten, zufällig ausgewählten Booleschen Funktionen mit K Inputs

Beispiel:



⇒ jeder Zustand (des gesamten Netzwerkes, d.h. durch bin. N -Tupel beschreibbar) hat eindeutig def. Nachfolgezustand

im obigen Beispiel:



(000) ist hier instabiler, stationärer Zustand

{(001), (010)} ist instabiler Zustandszyklus der Länge 2

(111) ist stabiler Attraktor (bei 1 Bit Mutation gelangt man nach endlich vielen Schritten wieder in diesen stationären Zustand)

- wieviele Attraktoren gibt es?
- wie groß sind die Attraktionsbereiche (Einzugsgebiete)?
- wieviele Attraktoren sind stabil gegen Mutationen / stabil gegen Änderung einzelner Boolescher Funktionen?

"homöostatische Stabilität" bedeutet:

kleiner Zustandszyklus mit riesigem Attraktionsbereich, der fast den gesamten Zustandsraum erfasst

Ergebnisse:

wichtigster Steuerparameter ist K (Zahl der Inputs)

$K = 1$: sehr kurze Zustandszyklen, schnelles "Einfrieren" des Netzwerks

$K = N$: Länge der Zustandszyklen $\approx \sqrt{2^N} = 2^{N/2}$
Zahl der Attraktoren $\approx N/e$
stark chaotisches Verhalten

$4 \leq K \leq N$: ebenfalls vorwiegend chaotisches Verhalten

$K = 2$: geordnetes Verhalten:

Länge der Zustandszyklen nur $\approx \sqrt{N}$
(das ist extrem kleiner Teil des ges. Zustandsraumes!)

nach wenigen Schritten entsteht riesige "eingefrorene Komponente", wo sich nichts mehr ändert

Benachbarte Zustände konvergieren im Zustandsraum

Stabilität gegenüber kleiner Änderung der Verdrahtung oder der Logik

zweiter Parameter P

kann verwendet werden, um auch in Zufallsnetzwerken mit $K > 2$ spontane Ordnung "zu erzwingen"

P mittelt für alle Booleschen Funktionen im Netzwerk den proportionalen Anteil des jeweils häufigeren Outputwertes (0 oder 1)

P bei 0,5: Netzwerk im chaotischen Regime

$P \rightarrow 1$: Netzwerk zunehmend geordnet

Durch Änderung von K und P Übergänge vom chaotischen ins geordnete Regime möglich

- im chaotischen Regime: großer Bereich, wo sich Zustände unregelmäßig ändern, darin kleine stabile Inseln
- im geordneten Regime: große gefrorene Komponente, darin kleine Inseln, wo noch Veränderungen stattfinden
- am Chaosrand: "Kontaktbahnen" zwischen den nicht-eingefrorenen Inseln; Signalkaskaden können durch das System wandern, ohne dass es völlig chaotisch wird

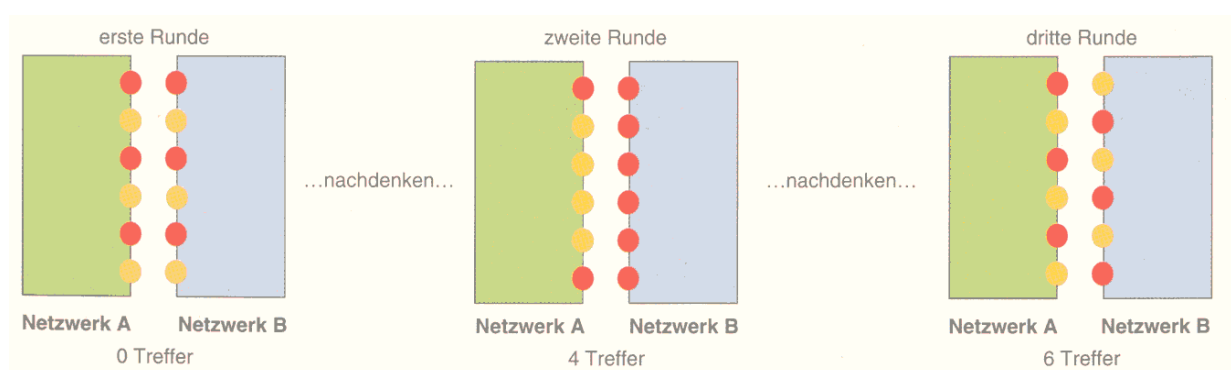
Hypothese:

Autokatalytische Netzwerke evolvieren durch Veränderung der Parameter K und P ins geordnete Regime nahe des Chaosrandes

Grund: Systeme, die zu komplexem Verhalten fähig sind, haben Überlebensvorteil

experimenteller Test an Booleschen Zufallsnetzwerken, die "gegeneinander spielen" (Anwendung eines "genetischen Algorithmus" (siehe später), Mutationen: Änderungen der Verdrahtung und der Booleschen Regeln) (Macready, Kauffman et al.)

→ Evolution zum geordneten Regime, aber unweit des Chaosrandes



(aus Kauffman 1991)

Weitere Möglichkeit, geordnetes Verhalten zu begünstigen:

Wahl von "kanalisierenden" Booleschen Funktionen (eine der Inputvariablen legt, wenn sie einen bestimmten Wert hat, bereits den Output fest – Beisp. "oder", "und", "nand"; nicht aber "xor").

Boolesche *NK*-Zufallsnetzwerke auch als Modelle für andere (verwandte) Arten von Netzwerken:

Netzwerke der Genregulation

hier: Knoten mit booleschen Funktionen = Gene,
Verknüpfung = epistatische Kopplung

Genome höherer Lebewesen:

- in erster Näherung durch *NK*-Zufallsnetzwerke mit $K = 2$ und "oder"-Funktionen beschreibbar
- die meisten Gen-Wechselwirkungen sind jedenfalls durch *kanalisierende* Funktionen charakterisiert

Stabilität gegen Mutationen empirisch:

- eine durch ein Hormon herbeigeführte Mutation eines Gens bei der Taufliege (*Drosophila*) führt zu einer Kaskade von 150 Veränderungen der Genaktivitäten der über 5000 Gene der Fliege.
- Bei entspr. Computersimulationen mit *NK*-Zufallsnetzwerken ($N = 5000$, $K = 2$) wurde ein Wert von 160 vorausgesagt.

(Schöneburg et al. 1994)

Genregelungsnetze für die Zelldifferenzierung bei Mehrzellern

Spezialfall der obigen Anwendung

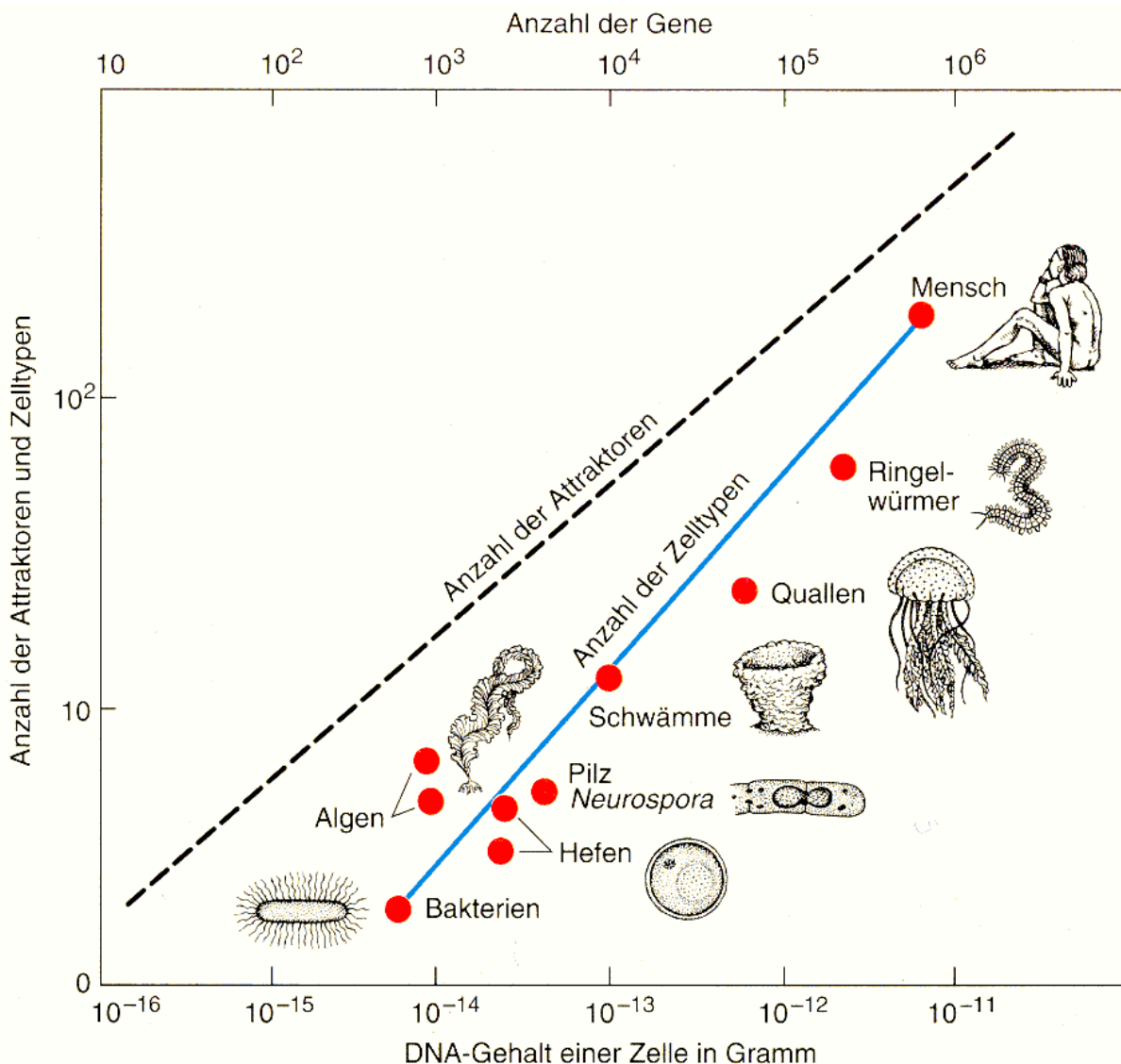
Knoten = Gene, stabile Aktivitätsmuster (Attraktoren) = stabile Muster der Genexpression, entsprechend den Zelltypen

- bei 100 000 Strukturgenen hätte man theoretisch $2^{100\,000}$ mögliche Aktivitätsmuster
- wenn das Netzwerk im geordneten Regime ist, gibt es jedoch im Mittel nur $\sqrt{100\,000} \approx 317$ Attraktoren, die kurze Zustandszyklen aufweisen

empirisch:

Drosophila: 15 000 Strukturgene, 60 Zelltypen

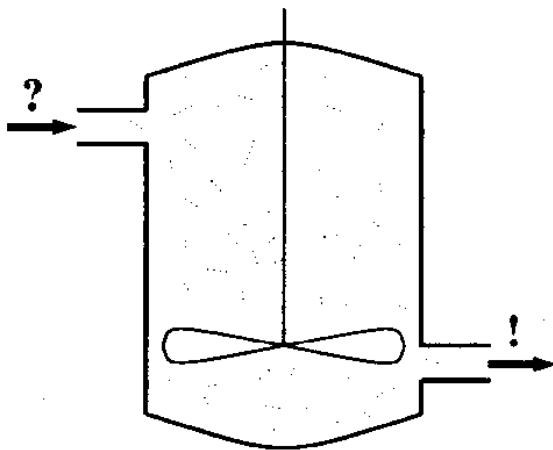
Mensch: ca. 100 000 Strukturgene, 256 Zelltypen



Artificial Chemistry

Die Art und Weise, wie Moleküle der "realen Chemie" reagieren, wird verallgemeinert auf beliebige Objekte

Modellvorstellung: gut gerührter chemischer Reaktor mit Zu- und Abfluss; darin befindet sich eine Population (Suppe) von Objekten, die jederzeit miteinander kollidieren und reagieren können.



(aus Skusa et al. 2000)

Formal besteht eine Künstliche Chemie aus:

- einer Objektmenge (z.B. Strings, Lambda-Terme, Zahlen, prädikatenlogische Formeln...)
- einer Regelmenge, die die Interaktionen zwischen den Objekten beschreibt (explizit: für jede Kombination von 2 Objekten muss das Reaktionsprodukt festgelegt werden, oder implizit durch Formeln, Algorithmen etc.)
- einer Kontrollstruktur, die Rahmen und Systemablauf festlegt (wann reagieren welche Objekte miteinander, welche Anfangs- und Randbedingungen werden eingehalten, hat der Reaktor eine Topologie...)

Beispiel: Zahlen-Divisions-Chemie ("rechnende Suppe")

Objekte: positive ganze Zahlen

Regeln:

$a + b \rightarrow c + b$ mit $c = a/b$ wenn a durch b teilbar,

$a + b \rightarrow c + a$ mit $c = b/a$ wenn b durch a teilbar,

$a + b \rightarrow a + b$ sonst.

(Der letzte Fall wird "elastische Kollision" genannt.)

Kontrollstruktur:

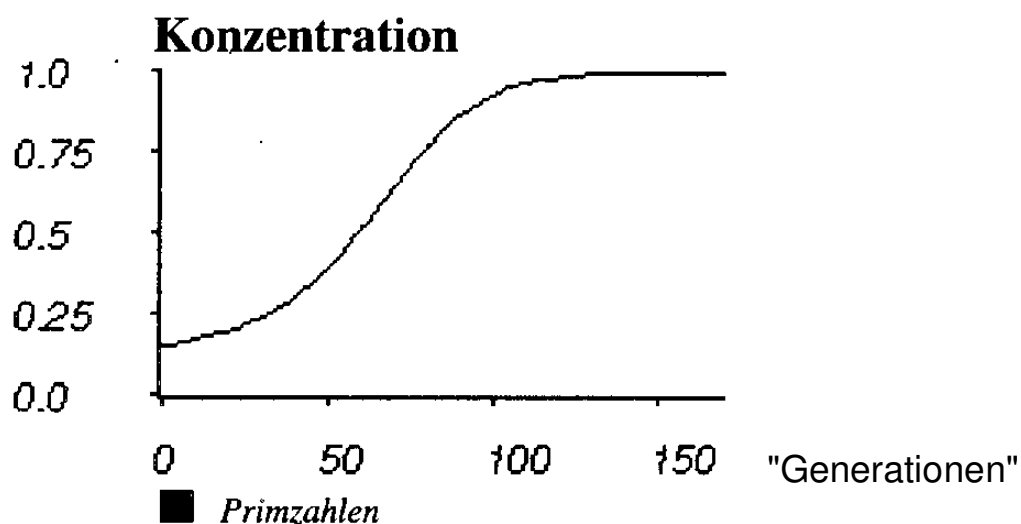
Initialisierung mit Zufallspopulation von M Zahlen

Reaktionsalgorithmus: zufälliges Ziehen zweier

Reaktionspartner

```
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
```

Ergebnis: die Population der Primzahlen nimmt mit der Zeit zu.

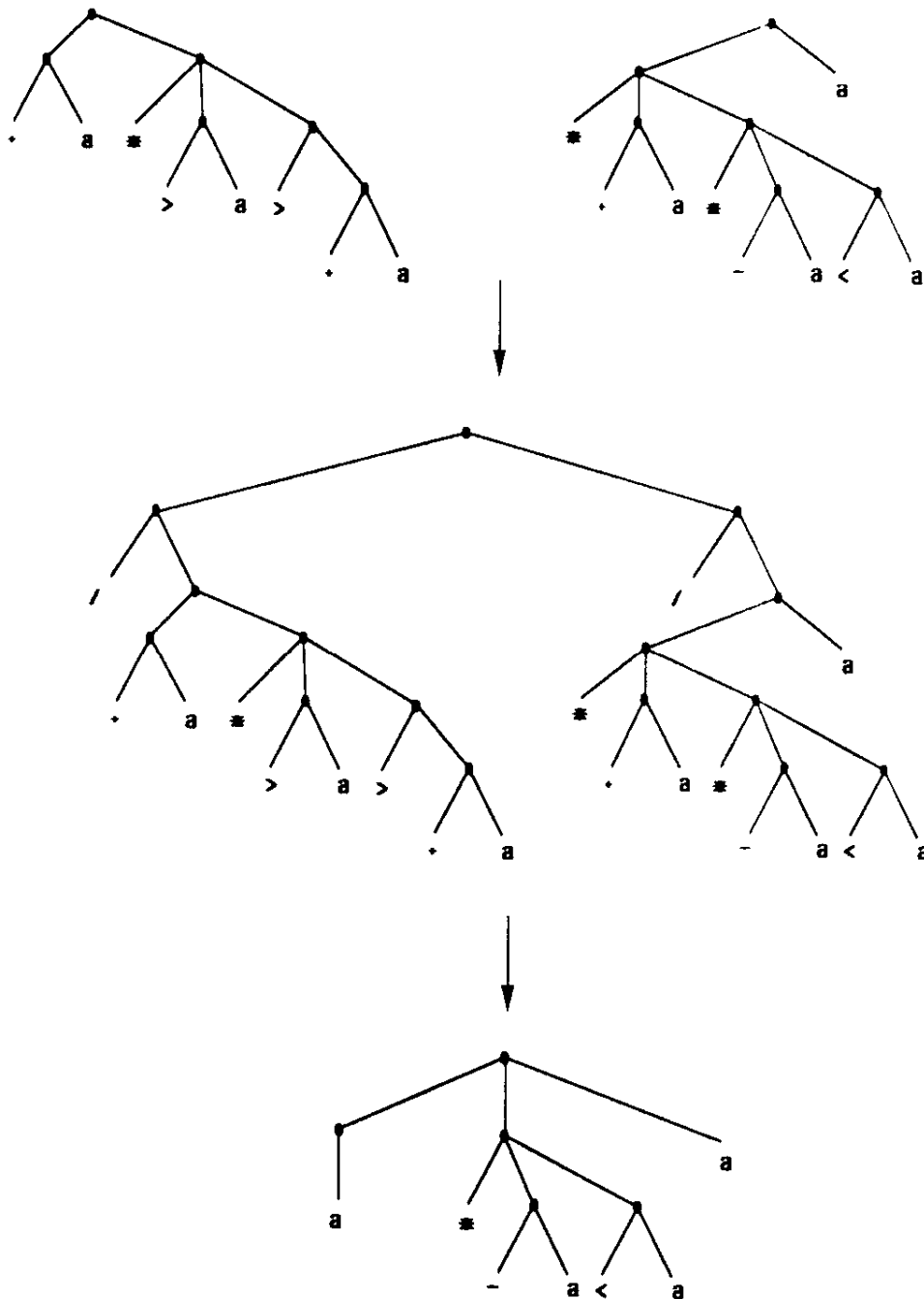


Anderes Beispiel: *AlChem*y (Fontana 1992)

die Objekte sind Terme des λ -Kalküls
(Variante von Lisp)

"function gas", "Turing gas"

Reaktionsergebnis zweier Terme: zusammengesetzter Term,
welcher ausgewertet wird



Start: Zufallsmenge von Funktionen

im Verlauf der Reaktionen verändert sich die Zusammensetzung des "Turing-Gases"

TABLE 2 State of an unperturbed Turing gas. The table lists the state of a Turing gas with $N = 1000$ particles after 3×10^5 collisions. First column: lexicographic order of the function. This number is the "name" used in the text to refer to a particular function. Second column: – marks indicate functions that disappear during the following 2×10^5 collisions. Third column: number of copies. Fourth column: function expression. See Figure 5 for the interaction graph and text for the details.

(# 1)	–	14	((('a)(a))('a)(a))
(# 2)		43	((('a)(a))('a)(a))((('a)(a))('a)(a))
(# 3)	–	16	(((((>'(>a)))(>'(>a))))(a))((>'(>a)))(>'(>a))) (a))(((>'(>a)))(>'(>a))))(a))((>'(>a)))(>'(>a))))))(a))
(# 4)		90	(((((>'(>a)))(>'(>a))))(a))((>'(>a)))(>'(>a))))(a))
(# 5)		33	((a(a))(((('a)(a))('a)(a))))
(# 6)		133	((a(a))(((('a)(a))('a)(a))((('a)(a))('a)(a))))
(# 7)	–	36	((a(a))((((>'(>a)))(>'(>a))))(a))((>'(>a)))(>'(>a)) '(>a))))(a))(((>'(>a)))(>'(>a))))(a))((>'(>a)))(>'(>a))))(a))
(# 8)		205	((a(a))((((>'(>a)))(>'(>a))))(a))((>'(>a)))(>'(>a)) '(>a))))(a))
(# 9)	–	10	(((*a)((*a)(*a)))
(# 10)		7	(((*a)(*a))
(# 11)		1	((>a)(>a))
(# 12)	–	37	((a)((*a)((*a)(*a))))
(# 13)		5	((a)((*a)(*a)))
(# 14)		7	((a)((>a)(>a)))
(# 15)		40	((a)(*a))
(# 16)		220	((a)(>a))
(# 17)		11	(*a)
(# 18)		92	(>a)

(aus Fontana 1992)

es entwickeln sich komplexe Reaktionsnetzwerke mit "Kooperationsbeziehungen" zwischen den Objekten:

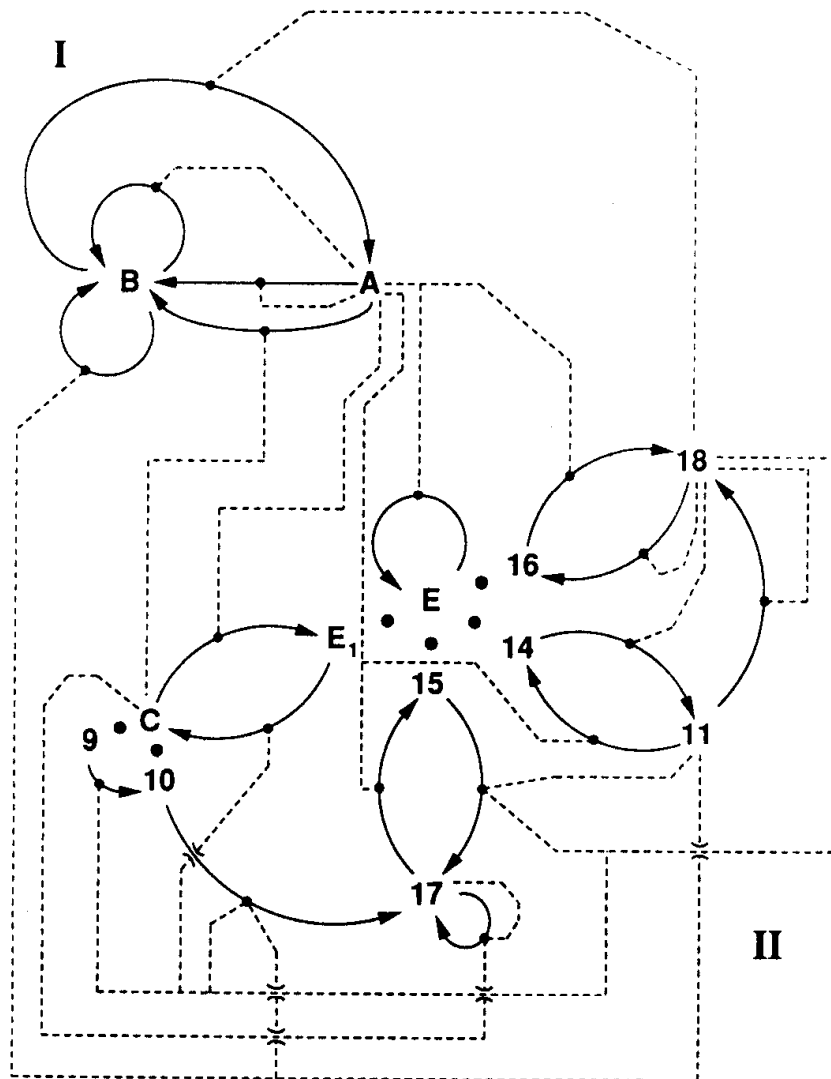


FIGURE 5 Interaction graph. The interaction graph of the functions listed in Table 2 is shown. The numbers denote the individual functions according to their ordering in Table 2. Capital letters denote sets, where $A = \{1, 2, 3, 4\}$, $B = \{5, 6, 7, 8\}$, $C = \{9, 10\}$, $E = \{12, 13, 14, 15, 16\}$, and $E_1 = \{12, 13\} \in E$. Solid arrows indicate transformations, dotted lines functional couplings. A dotted line originates in a function (or a set, see text), say, k , and connects (filled circle) to a solid arrow, whose head is j and whose tail is i . This is to be interpreted as $j = k(i)$. Large filled circles indicate membership in a particular set. Function 17 is an identity function. Note: all dotted lines and solid arrows that result from 17 copying everything else in addition to itself have been omitted. See text for details.

auch die "Morphologie" der Terme zeigt Anzeichen der
Strukturbildung; so bilden sich z.B. "selbstähnliche Terme":

