

2 Modellierungsformalismen: S-Systeme und Boolesche Netze

2.1 Potenzgesetze

In der Natur sind allometrische Beziehungen der Form

$$X_i = b_{ij} X_j^{a_{ij}}$$

weit verbreitet. Eine solche Beziehung nennt man auch *Potenzgesetz*.

Geometrische Größen (Flächen A , Volumina V) zeigen im Allgemeinen geometrische Skalierung (proportional zum Quadrat oder zur dritten Potenz der Länge L):

$$L \rightarrow \lambda L \quad \Rightarrow \quad A \rightarrow \lambda^2 A, \quad V \rightarrow \lambda^3 V$$

Der Exponent gibt die räumliche Dimension der Größe an. Ausnahme: Fraktale sind geometrische Figuren, deren Inhalt oder Rand eine nicht-ganzzahlige Dimension hat (Beispiel Sierpinski Dreieck). Viele biologische Größen skalieren mit dem Körpergewicht, aber der Exponent ist kein Vielfaches von $1/3$ (wie bei geometrischer Skalierung zu erwarten wäre): dies nennt man allometrische Skalierung.

Beispiele für allometrische Skalierung

- Kleibers Gesetz, welches den Zusammenhang zwischen Masse und Stoffwechsel von Tieren beschreibt. Dieser Gesetzmäßigkeit zufolge skaliert der Metabolismus I mit der Potenz von $3/4$ mit der Körpermasse m $I = I_0 \cdot m^{3/4}$ mit einer körpergrößenunabhängigen Normierungskonstanten I_0 .
- Eine ähnliches Gesetz beschreibt den Zusammenhang zwischen Masse und Lebenszeit von Tieren. Dieser Gesetzmäßigkeit zufolge skaliert die Lebenszeit t mit der Potenz $1/4$ mit der Körpermasse m

$$t = m^{1/4},$$

was bedeutet, daß größere Tiere länger leben.

- Baumgröße G und dessen Durchmesser D

$$D = G^b$$

2.2 S-Systeme

S -Systeme (S für Sättigung (von Saturation) oder Synergismus) dienen zur Beschreibung und Simulation biologischer und biochemischer Systeme die einem Grenz- oder Sättigungszustand zustreben.

Änderungen biochemischer Größen können meistens durch Bilanzgleichungen beschrieben werden

$$\begin{aligned} d/dt X_i t &= X_i' = V^+(X_1, \dots, X_n) - V^-(X_1, \dots, X_n), \quad i = 1, \dots, n \\ &= \text{Produktion} - \text{Abbau} \\ &= \text{Einfluß} - \text{Ausfluß} \\ &= \text{Synthese} - \text{Komplexbildung} \end{aligned}$$

wobei V^+, V^- unbekannt und in der Regel nicht meßbar sind.

Beispiel $X_1 \rightarrow X_2 \rightarrow X_3$ Abbau von $X_1 =$ Produktion von X_2 , etc. Die Differentialgleichungen lauten

$$X_i'(t) = \alpha_i \prod_{j=1}^n X_j^{g_{ij}} - \beta_i \prod_{j=1}^n X_j^{-h_{ij}} \quad i = 1, \dots, n$$

mit $\alpha_i = \beta_i e^{b_i}$ und $g_{ij} = h_{ij} + \tilde{a}_{ij}$.

Dies ist die allgemeine Beschreibung von S -Systemen mit n Variablen $X_i, i = 1, \dots, n$.

Anmerkungen zu S-Systemen

- $\alpha_i, \beta_i \geq 0$ entsprechen den Ratenkonstanten der klassischen Kinetik
- Das Modell berücksichtigt nicht die tatsächlichen Reaktionsstöchiometrien, d.h., es kann Materie verschwinden!
- $g_{ij}, h_{ij} \in \mathbb{R}$ entsprechen der Reaktionsordnung.
Klassisch: $g_{ij}, h_{ij} \in I$ entsprechen der Anzahl der beteiligten Moleküle



Man kann aber ableiten, daß Reaktionsordnungen nicht ganzzahlig sein müssen, z.B. bei ineffizienten Prozessen (z.B. ATP-Erzeugung in der oxidativen Phosphorylierung)

- *S*-Systeme werden aus zwei Annahmen abgeleitet (siehe unten)
 1. Potenzgesetz
 2. stationärer Zustand
- flexibel: alle biochemischen Prozesse können in der Form eines *S*-Systems beschrieben werden
- einfach zu generieren: aus der Netzwerkstruktur kann man direkt das Differentialgleichungssystem ableiten
- Parameter haben definierte Bedeutung

Aber: *S*-Systeme sind nur in der Nähe des stationären Zustandes gerechtfertigt.

2.3 Boolesche Netze als Modelle für Genaktivitäten

2.3.1 Boolesche Netze

Viele biologische Prozesse können näherungsweise als binäre Funktionen beschreiben werden.

Beispiel Diskretisierung einer sigmoiden Signal-Antwort-Kurve

- ein Gen kann entweder transkribiert werden oder nicht, d.h. an oder aus sein.
- Proteine können aktiv sein oder nicht, etc. ...

In kinetischen Modellen wird häufig aus einer mikroskopischen ja-nein-Entscheidung eine makroskopische kontinuierliche Steuerungsfunktion (beschreibt die Konzentrationen oder Wahrscheinlichkeiten der mikroskopischen Zustände). Man kann aber auch diese Funktion wieder diskretisieren (Näherung!).

Formal Ein biologisches System besteht aus N Komponenten ("Knoten"), die jeweils den Zustand "an" (=1 oder "wahr") "aus" (0 oder "falsch") oder haben können.

$$\Rightarrow \left| \{0, 1\}^N \right| = 2^N \text{ mögliche Zustände}$$

Geometrisch liegen die Zustände auf den Ecken eines Hyperwürfels.

Bei einer Zustandsänderung wird der neue Zustand jeder Komponente durch K andere Knoten mittels einer logischen ("booleschen") Funktion bestimmt:

Def.: Seien $\{i_1, \dots, i_{K_i}\}$ die K_i Knoten, die den Zustand σ_i des Knoten i bestimmen. Dann wird $\sigma_i(t+1)$ durch eine boolesche Funktion f_i ermittelt, mit

$$f_i : \{0, 1\}^{K_i} \xrightarrow{P} \{0, 1\},$$

die jede Konfiguration der K_i Knoten auf 1 oder 0 abbildet.

Es gilt also

$$\sigma_i(t+1) = f_i(\sigma_{i1}(t), \dots, \sigma_{iK_i}(t))$$

Beispiel 3-Knoten System, jeder ist mit jedem verknüpft. Die booleschen Funktionen: jedem Zustand $\{0, 1\}^3$ wird ein neuer zugeordnet, z.B.:

σ_1	σ_2	σ_3		f_1	f_2	f_3
0	0	0		0	1	1
0	0	1		1	0	1
0	1	0		1	1	1
0	1	1	→	0	1	1
1	0	0		1	1	1
1	0	1		1	1	1
1	1	0		1	0	1
1	1	1		1	1	0

Hier gilt z.B. $f_1 = \sigma_1 \vee (\neg\sigma_1 \wedge (\sigma_2 \text{ XOR } \sigma_3))$.

Man kann die booleschen Funktion in einem Transitionsgraphen darstellen (Hyperwürfel).

Umschaltchemata Man kann boolesche Netze auch nach ihren Schaltschemata klassifizieren. Es gibt

- synchronisierte deterministische Netze (siehe Beispiel), wo jeder Zustand genau einen Nachfolger hat.
- deterministische asynchrone Netze, etc.
- zufällig asynchrone Netze
 - zufällig ausgewählter Knoten wird geschaltet
 - ein Zustand kann mehrere Nachfolger haben (nicht-deterministisch)
- Stochastische Regeln: Die Ergebnisse 1 oder 0 sind nicht strikt festgelegt, sondern erscheinen mit der Wahrscheinlichkeit P (für 1) und $(1 - P)$ (für 0).

Topologien Man kann boolesche Netze mit verschiedenen Zufallsstrukturen erzeugen

- **Zufällige Netze (random networks)**

Diese tauchen auf, wenn man Gen-Netzwerke mit der Hypothese untersucht, daß sich diese zufällig gebildet haben. Im Erdős-Renyi-Zufallsgraphen wird jede mögliche Kante unabhängig und mit gleicher Wahrscheinlichkeit realisiert. Für die Anzahl der Eingänge K_i zum Knoten i (Eingangsgrad) ergibt sich die Poissonverteilung

$$P_{in}(K_i) = e^{-K} \frac{K^{K_i}}{K_i!}, \quad K > 0$$

Den mittleren Eingangsgrad $\langle K_i \rangle = K$ nennt man auch Konnektivität; sie charakterisiert das boolesche Netz.

- **Skalenfreie Netzwerke**

Beim Studium biologischer Netze, z.B. Protein-Netzwerke oder Gen-Netzwerke, stellte man allerdings fest, daß die Wahrscheinlichkeit $P(k_i)$, daß ein zufällig gewählter Knoten i k_i Eingänge hat, nicht Poisson-verteilt ist, sondern dem Potenzgesetz gehorcht:

$$P(k_i) \sim c k_i^{-\gamma}, \quad c, \gamma > 0$$

D.h. viele Knoten haben wenig Eingänge, wenig Knoten haben viele Eingänge. Die letzteren werden "Hubs" genannt.

Bei diesen Netzwerken macht eine mittlere Konnektivität wenig Sinn, daher werden sie skalenfrei genannt.

Skalenfreie Netze können durch die folgende Konstruktion entstehen ("bevorzugtes Verbinden"). Zu einem bestehenden Graphen werden neue Knoten hinzugefügt. Jeder neue Knoten wird mit einem der bestehenden Knoten verbunden; dieser Knoten wird zufällig ausgewählt, mit Wahrscheinlichkeiten proportional zur Anzahl der bereits bestehenden Kanten. Knoten, die bereits viele Nachbarn haben, erhalten daher umso mehr neue Nachbarn ("Die Reichen werden noch reicher").

Eigenschaften Der Zustandsraum $\{0, 1\}^N$ ist endlich.

⇒ Zustände werden wieder erreicht. Diese heißen *Attraktoren* und können einelementig sein (z.B. 011 im obigen Beispiel) oder mehrelementig (z.B. $101 \rightarrow a111 \rightarrow 110 \rightarrow 101$) im obigen Beispiel). Sie heißen dann *Punkt-* oder *Kreisattraktoren*. Manchmal zielt die Modellierung darauf ab, die Punktattraktoren als Differenzierungszustände der Zelle zu interpretieren.

Dynamik Man kann bei booleschen Netzen drei Arten von dynamischem Verhalten unterscheiden

1. **Ordnung**

Der Initialzustand findet schnell seine Attraktoren, kleine Perturbation dieses stationären Zustandes sterben schnell aus.

2. **Chaos**

Kleine Perturbationen von den Attraktoren breiten sich über das gesamte Netz aus und es gibt über längere Zeit keine stabilen Zustände.

3. **Zwischen Ordnung und Chaos (die Grenze)**

Teilweise finden Perturbationen schnell wieder in Attraktoren, teilweise aber breiten sie sich über das gesamte Netz aus.

Unter Umständen kann man voraussagen, welches dynamische Verhalten in einem Netz vorherrscht.

2.3.2 Zufällige N-K-Netze (Kauffman 1969)

Annahmen:

1. Jeder Knoten wird von exakt K anderen Knoten kontrolliert.
2. Die K Knoten die jedes Gen kontrollieren sind zufällig gewählt aus einer Gleichverteilung.
3. Die booleschen Funktionen werden zufällig festgelegt (aber dann festgehalten)

$$\sigma_i(t+1) = f_i(\sigma_{i1}, \dots, \sigma_{iK}) = \begin{cases} 1 & \text{mit Wahrscheinlichkeit } P \\ 0 & \text{mit Wahrscheinlichkeit } 1 - P \end{cases}$$

Dann ist das dynamische Verhalten im wesentlichen durch die Parameter P und K bestimmt. Es gilt: Für alle P existiert ein $K_c(P) = [2P(1-P)]^{-1}$, so daß für $K < K_c(P)$ alle Perturbationen in Attraktoren (Ordnung) enden und für $K > K_c(P)$ alle Perturbationen durch das gesamte Netz propagiert werden (Chaos).

Für $P = 1/2$ ergibt sich:

Chaos ($K \rightarrow N$): Attraktorlänge $L \sim 2^{N/2}$, Attraktorzahl $M \approx N$

Ordnung ($K \approx 2$): Attraktorlänge $L \sim N^{1/2}$, Attraktorzahl $M \approx N^{1/2}$

Kauffman machte die Beobachtung, daß biologische (genetische) Netze nahe am Chaos sind ("living on the edge"). Argumente: (i) Eine mittlere Zahl $K = 2$ von Regulatoren pro Gen ist realistisch. (ii) Differenzierungszustände müssen teils stabil, teils flexibel sein. (iii) Die Zahl der Zelltypen (ca. 250 bei Säugern) entspricht ungefähr der erwarteten Zahl der Attraktorzustände (ca. 330 bei einem Netzwerk mit 10^5 Knoten).

Für skalenfreie Netzwerke konnte eine ähnliche Beziehung abgeleitet werden:

Für alle p existiert ein γ_c , so daß für $\gamma > \gamma_c$ das System in geordnetem Zustand und für $\gamma < \gamma_c$ in chaotischem Zustand ist.

Bei biologischen scale free networks, z.B. Proteinnetzwerken von Hefe, fand man $\gamma \in [2; 2,5]$, d.h. auch hier ist das Netz an der Grenze zum Chaos.

Software zu booleschen Netzen

- Matlab RBN Toolbox
- Random Boolean Networks [RBN]
- Discrete Dynamics Lab [DDLab]

2.4 Übersetzung boolescher Netze in einfache kinetische Modelle

Linearer Abbau **Einfacher Abbau** $S \xrightarrow{v=k \cdot s}$ mit Massenwirkungskinetik $v = k \cdot s$.

Differentialgleichung $\frac{ds}{dt} = -k \cdot s$ mit Anfangsbedingung $s(0) = s_0$.

Lösung: $s(t) = s_0 e^{-k \cdot t}$

Beweis durch Einsetzen oder durch Trennung der Variablen:

$$\int_0^t \frac{1}{s(t')} \frac{ds}{dt'} dt' = \int_0^t -k dt' \quad \Rightarrow \quad \int_0^t \frac{d \ln s}{dt'} dt' = -k t$$

Lösungskurvenschar mit einem Parameter (k) und einer Anfangsbedingung (s_0).

Charakteristische Zeit $\tau = 1/k$. Es gilt $s(\tau) = \frac{1}{e} s(0)$, bzw allgemein $s(t + \tau) = \frac{1}{e} s(t)$.

Halbwertszeit $t_{1/2} = \ln 2 \cdot \tau$; es gilt $s(t + t_{1/2}) = \frac{1}{2} s(t)$

Produktion und linearer Abbau **Produktion und linearer Abbau** $v_0 \rightarrow S \xrightarrow{v=k \cdot s}$.

Differentialgleichung $\frac{ds}{dt} = v_0 - k \cdot s$ mit $s(0) = s_0$.

Stationärer Zustand mit Konzentration \bar{s} : Bedingung $\frac{ds}{dt} = 0$ ergibt $\bar{s} = \frac{v_0}{k}$

Stimmen die Einheiten?

Allgemeine Lösung: Ersetze $s(t) = \bar{s} + x(t)$ mit Abweichung $x(t)$. Es gilt

$$\frac{dx}{dt} = \frac{ds}{dt} = v_0 - k(x + \bar{s}) = -kx \quad (1)$$

mit Lösung $x(t) = x(0) e^{-kt} = (s(0) - \frac{v_0}{k}) e^{-kt}$ wie oben, also insgesamt:

$$s(t) = \frac{v_0}{k} + (s_0 - \frac{v_0}{k}) e^{-kt}$$

Exponentielle Annäherung an den stationären Zustand \Rightarrow stabiler Fixpunkt. (A propos: Wie lauten die Eigenwerte der Jakobimatrix?)

Steuerung mit diskreten Schwellwerten (“stückweis konstante Produktion”)

(a) Die Produktionsrate wechselt zwischen den Werten $v_0 = 0$ und $v_0 = v_{\max}$ hin und her. Jeder Puls der Produktion führt zu einer exponentiellen Annäherung an den stationären Zustand, gefolgt von einem exponentiellen Abklingen.

(b) Netzwerk mit diskreten Steuerungsfunktionen: jeder Knoten hat genau einen Eingangsknoten. Steigt dessen Wert über einen Schwellwert, wird die Produktion angeschaltet. Somit wandern die Pulse durch das Netzwerk. Genauso kann ein Knoten mehrere Eingangsknoten haben, deren Niveaus mittels Schwellwerten in binäre Signale übersetzt werden. Eine boolesche Funktion bestimmt dann, ob die Produktion mit der Rate 0 oder v_{\max} geschieht.