

# Die Modellierung von Raum und Struktur in molekularbiologischen Systemen

Mathias John  
Graduiertenkolleg dIEM oSiRiS  
Universität Rostock  
mj069@informatik.uni-rostock.de

## 1 Aufgabenstellung

Die Entwicklung eines Formalismus zur Modellierung von Raum und Struktur in molekularbiologischen Systemen.

## 2 Modellierung von Raum und Struktur in molekularbiologischen Systemen

Das wissenschaftliche Interesse der Medizin gilt zunehmend der Untersuchung intrazellulärer Systeme, da dort die Ursprünge vieler Krankheiten gefunden werden können. Die Abläufe auf dieser Ebene stellen sich jedoch als hoch komplex dar. Zum einen formen die Interaktionen von Proteinen Netzwerke, in denen einzelne Komponenten viele verschiedene Rollen einnehmen können. Zum anderen ist es trotz neuer Technologien nicht möglich, jede Komponente eines Netzwerks zu entdecken und zu observieren. Motiviert durch diese Probleme versuchen Biologen in Zusammenarbeit mit Informatikern, Modelle von molekularbiologischen Systemen zu entwerfen. Diese sollen helfen, das aus Experimenten entstandene Wissen zu strukturieren, neue Einsichten in die Systeme zu gewinnen und neue Theorien zu entwickeln.

Bei der Modellierung von molekularbiologischen Systemen müssen verschiedene räumliche Effekte betrachtet werden, die mit der starken Unterteilung von Zellen durch verschiedene Strukturen zusammenhängen. Zum einen existieren Membranen, die als natürliche Barrieren die Bewegung von Molekülen beeinträchtigen, so dass Konzentrationsgefälle entstehen. Abhängig vom Zellbereich sind somit Reaktionen zu beobachten, die sich stark in ihrer Art und ihrem Ablauf voneinander unterscheiden. Zum anderen bilden Mikrotubuli ein Zellgrundgerüst (Zellskelett), an dem aktive Transportprozesse stattfinden, d.h. es werden Moleküle unter Energieverbrauch zumeist entgegen bestehender Konzentrationsgefälle befördert. Neben der räumlichen Struktur von Zellen, spielt auch die Größe und Struktur von Molekülen eine tragende Rolle. Zum Beispiel beeinflusst die sekundäre und tertiäre Struktur von Proteinen maßgeblich ihre Funktionalität (z.B. s. [LWT05]). Des

weiteren können große Moleküle im begrenzten Raum zwischen den Mikrotubuli Molekülstaus verursachen (siehe [TNAT05]). Durch neue Analysetechniken im Bereich der Molekularbiologie, wie z.B. die Konfokalmikroskopie oder das High Throughput Screening, können solche Effekte deutlich besser untersucht werden. Die dabei gewonnenen Informationen sollen auf Grund ihrer Wichtigkeit in die Modelle einfließen. Für diesen Anwendungskontext werden spezielle Modellierungsansätze benötigt.

### 3 Verwandte Arbeiten

Klassische Methoden zur Modellierung und Simulation von molekularbiologischen Systemen basieren auf Differentialgleichungen, die die gemittelten Konzentrationen der chemischen Elemente eines Modells kontinuierlich beschreiben (s. z.B. [LSK<sup>+</sup>03]). Es ist jedoch bekannt, dass in einigen Fällen, wie z.B. der Gentranskription von Tryptophan [TT03] nur sehr wenige Moleküle an den ablaufenden Prozessen beteiligt sind. In diesem Rahmen sind Methoden von Vorteil, die anstelle von Konzentrationen mit Teilchenanzahlen arbeiten, wie z.B. der Stochastische Simulations Algorithmus (SSA) [Gil77]. Der SSA eignet sich gerade im Bereich der Molekularbiologie sehr gut, da er die Entwicklung von Systemzuständen stochastisch betrachtet und es somit ermöglicht, sowohl Unschärfen im Wissen über das beobachtete System als auch natürlich zufällige Vorgänge, wie die Brownsche Bewegung, in die Simulation zu integrieren. Zur Modellierung von Systemen für SSA-Simulationen wird dem Vorschlag von Regev und Shapiro [RS04] folgend eine stochastische Variante des  $\pi$ -Kalküls [Mil99], der stochastische  $\pi$ -Kalkül [Pri96], genutzt. Die Modellierung im stochastischen  $\pi$ -Kalkül sieht jedoch bis auf wenige Ausnahmen, die sich auf die Abbildung von Membranen beschränken (z.B. BioAmbients [RPS<sup>+</sup>04]), keine Betrachtung von räumlichen Effekten oder strukturellen Eigenschaften vor. Es existiert ein auf dem SSA basierender Ansatz, die Next Subvolume Methode [EE04], welcher die Möglichkeit bietet, Systeme räumlich zu unterteilen. Jedoch können dabei komplexere Effekte, wie z.B. aktiver Transport, nicht betrachtet werden.

### 4 Vorarbeiten

Es wurde eine Idee zur Erweiterung des  $\pi$ -Kalküls entwickelt, mit deren Hilfe komplexe räumliche und strukturelle Informationen in die Modellierung integriert werden kann. Diese soll nun näher erläutert werden. Zum besseren Verständnis wird zunächst jedoch kurz die Modellierung biochemischer Systeme mit Hilfe des  $\pi$ -Kalküls beschrieben.

Die  $\pi$ -Kalkül-Syntax sieht den Term Prozess vor, der sich aus parallelen Kompositionen,  $\nu$ -Operatoren und Summen, mit den Summanden Senden und Empfangen, zusammensetzt (s. Tab. 1). Um ein biochemisches System mit Hilfe des  $\pi$ -Kalküls zu modellieren, wird jede chemische Reaktion auf einen Kanal abgebildet. Zusätzlich existiert für jedes Molekül eine Summe, mit Summanden für jede Reaktion, an der das Molekül teilnehmen kann. Im Rahmen des stochastischen  $\pi$ -Kalküls erhält jeder Kanal als Annotation eine stochastische

Rate, die sich aus der kinetischen Rate der zugeordneten Reaktion ergibt. Auf Grund der stochastischen Semantik, die der stochastische  $\pi$ -Kalkül bietet, können so beschriebene Systeme direkt mit Hilfe des SSA simuliert werden.

---

Prozess	$P ::=$	$P_1 \parallel P_2 \mid$ $(\nu c).P \mid$ $\sum_i S_i$	Parallele Komposition $\nu$ Operator Summation
Summation	$S ::=$	$x!(y).P \mid$ $x?(y).P$	Senden Empfangen

---

Tabelle 1: Die  $\pi$ -Kalkül Syntax

Zur Integration komplexer räumlicher und struktureller Informationen in die Modellierung von biomolekularen Systemen soll der  $\pi$ -Kalkül mit Prozessattributen beliebigen Typs erweitert werden. Die Attribute und deren Werte sollen dabei innerhalb der Modellierungssprache definiert werden können. Dies ermöglicht die Verwendung von Attributen zur Beschreibung von Raum (z.B. Positionen) als auch von Struktur (z.B. sekundäre Proteinstruktur). Damit die Attribute nicht nur als zusätzliche Annotationen von Prozessen fungieren, sondern auch Einfluss auf die Abläufe des modellierten Systems haben, sollen flexibel Konstrukte definiert werden können, die beschreiben, wie die Attribute die Kommunikation zwischen Prozessen beeinflussen. Zu beachten ist dabei, dass sich die Erweiterungen möglichst nahtlos in die  $\pi$ -Kalkül Syntax einfügen und auch intuitiv zu handhaben sind, da davon maßgeblich der Erfolg des Ansatzes abhängt. Zusätzlich müssen die neuen Konstrukte ein gewisses Maß an Expressivität bieten, da auch Funktionen als Attribute verwendet werden sollen, um z.B. komplexe Bewegung von Molekülen zu beschreiben.

Die Integration von Attributen bezieht sich zunächst nur auf den  $\pi$ -Kalkül und ist mit keiner speziellen Ausführungssemantik gekoppelt. Grundsätzlich kann der  $\pi$ -Kalkül zur Beschreibung von zwei Arten von Systemen genutzt werden, für diskrete, d.h. Systeme deren Zustand sich ausschließlich diskontinuierlich über die Zeit ändert, und für hybride, d.h. Systeme deren Zustand sich sowohl diskontinuierlich als kontinuierlich über die Zeit ändert. Eine rein-kontinuierliche Ausführungssemantik ist für den  $\pi$ -Kalkül auf Grund seiner Kommunikationssemantik ungeeignet. Bisher waren alle  $\pi$ -Kalkül basierten Sprachen, wie stochastic  $\pi$ , mit einer diskreten Ausführungssemantik verbunden. Da jedoch auch hybride Ansätze in der Systembiologie vielfach Verwendung finden, z.B.: [dJGH<sup>+</sup>03], wurde in einer Vorarbeit zunächst eine  $\pi$ -Kalkül-basierte Sprache mit räumlichen Attributen und einer hybriden Ausführungssemantik entwickelt [JEU07].

## 5 Geplante Arbeiten

In [JEU07] wurde ein Ansatz vorgestellt, der zunächst ausschließlich vorgegebene räumliche Attribute zulässt. Dieser ist in oben genannter Weise zu erweitern. Daraufhin soll die vorgestellte hybride Ausführungssemantik verfeinert und zusätzlich eine passende stochasti-

sche Ausführungssemantik entwickelt werden. Der zweite Punkt ist von sehr großer Wichtigkeit, da auf diese Weise die Sprache an den SSA gekoppelt werden kann, der in der Systembiologie wegen der oben genannten Gründen bereits vielfach Einsatz findet. Für die so entstandene Sprache soll untersucht werden, inwiefern sich statische Analyse und Simulation zur Modelluntersuchung eignen. Dafür sind in Kooperation mit den Wet-Lab-Partnern des GRK, Modelle zu entwerfen. In diesem Zusammenhang soll auch geklärt werden, wie sich räumliche und strukturelle Informationen aus den in Experimenten produzierten Daten zur Integration in die Modelle extrahieren lassen.

## Literatur

- [dJGH<sup>+</sup>03] Hidde de Jong, Jean-Luc Gouzé, Céline Hernandez, Michel Page, Tewfik Sari und Johannes Geiselmann. Hybrid Modeling and Simulation of Genetic Regulatory Networks: A Qualitative Approach. In *HSCC*, Seiten 267–282, 2003.
- [EE04] J. Elf und M. Ehrenberg. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Syst Biol (Stevenage)*, 1(2):230–236, Dec 2004.
- [Gil77] D. T. Gillespie. Exact Stochastic Simulation of Coupled Chemical Reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [JEU07] Mathias John, Roland Ewald und Adelinde Uhrmacher. A Spatial Extension to the Pi Calculus. In *Proceedings of the FBTC Workshop 2007*, (to appear) 2007.
- [LSK<sup>+</sup>03] E. Lee, A. Salic, R. Krüger, R. Heinrich und M. W. Kirschner. The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol*, 1(1), October 2003.
- [LWT05] Roman A. Laskowski, James D. Watson und Janet M. Thornton. ProFunc: a server for predicting protein function from 3D structure. *Nucl. Acids Res.*, 33(suppl2):W89–93, 2005.
- [Mil99] Robin Milner. *Communicating and Mobile Systems: the Pi-Calculus*. Cambridge University Press, June 1999.
- [Pri96] Corrado Priami. Stochastic pi-calculus with General Distributions. In P. Clu, Hrsg., *Proceedings of PAPM '96*, 1996.
- [RPS<sup>+</sup>04] Aviv Regev, Ekaterina M. Panina, William Silverman, Luca Cardelli und Ehud Shapiro. BioAmbients: an abstraction for biological compartments. *Theor. Comput. Sci.*, 325(1):141–167, September 2004.
- [RS04] Aviv Regev und Ehud Shapiro. The  $\pi$ -Calculus as an Abstraction for Biomolecular Systems. In Grzegorz Ciobanu, Gabriel Rozenberg, Hrsg., *Modeling in Molecular Biology*. Springer, 2004.
- [TNAT05] Kouichi Takahashi, Satya Nanda, Vel Arjunan und Masaru Tomita. Space in systems biology of signaling pathways : towards intracellular molecular crowding in silico. *FEBS letters*, 579(8):1783–1788, 2005.
- [TT03] Nancy Trun und Janine Trempy. *Fundamental bacterial genetics*. Blackwell, 2003.