

# Mathematische Modellierung in den Lebenswissenschaften – Stochastische Modelle und Simulationen mit dem Softwarepaket Ants<sub>InFields</sub>

Mathematical modelling in the life sciences – stochastic models and simulations with the software package Ants<sub>InFields</sub>

Institut für Biomathematik und Biometrie

Felix Friedrich

Die Modellierung von Fragestellungen in den Lebenswissenschaften verlangt komplexe Ansätze, die einer analytischen Untersuchung im klassischen Sinne meist nicht zugänglich sind. Simulationen sind deshalb fester Bestandteil des mathematischen Handwerks geworden; sie helfen, analytisch nicht zugängliche Größen zu ermitteln, Modelle zu validieren und komplexe Zusammenhänge zu veranschaulichen. Das von uns entwickelte Softwarepaket Ants stellt einen Rahmen für die Implementierung und Visualisierung stochastischer Systeme bereit und ist speziell für die Simulation einer bestimmten Klasse von stochastischen Modellen – so genannten Markov-Feldern – konzipiert. Einige Beispiele für Anwendungen in den Lebenswissenschaften sollen sowohl die Bandbreite stochastischer Modelle als auch die Vielseitigkeit dieser Simulationsbasis demonstrieren. Auch Beispiele aus der Lehre werden präsentiert.

## Das Softwarepaket Ants

Das Softwarepaket Ants wurde von Felix Friedrich an der Universität Heidelberg und am Institut für Biomathematik und Biometrie

Modelling of problems in the life sciences requires complex approaches, which are generally not accessible to analytical investigation in the classic sense. Simulations have thus become a regular component of the mathematical approach; they permit to solve analytically intractable problems and allow the validation of models. Complex dependencies can be illustrated by means of simulation. The Ants software package that we have developed provides a framework for the implementation and visualisation of stochastic systems; it is conceived especially for simulations of a particular class of stochastic model – the so-called Markov fields. Some examples of applications in the life sciences illustrate both the wide range of stochastic models and also the versatility of this basis for simulation. Some examples from teaching situations are also presented.

(IBB) der GSF entwickelt und in Teilen beim Springer Verlag auf CD veröffentlicht. Geschrieben ist es in der Programmiersprache Oberon und benutzt das Heidelberger Simulationspaket Voyager. Der Name Ants

deutet auf die Beobachtung hin, dass selbst einfache Interaktionsmuster im Kleinen komplexe Muster im Großen erzeugen können. Ants ist eine Software zur Simulation stochastischer Prozesse und zur statistischen Inferenz auf Gibbs-Markov-Feldern. Dabei stehen stochastische Algorithmen zur Lösung hochdimensionaler Minimierungsprobleme und Simulationen stochastischer Abläufe im Vordergrund. Ziel ist einerseits, Software-Komponenten zur Untersuchung stochastischer Systeme bereit zu stellen, andererseits den Transport wissenschaftlicher Erkenntnisse mit interaktiven Demonstrationswerkzeugen zu unterstützen.

Ants enthält zahlreiche Visualisierungen von ein-, zwei- und höher dimensionalen Daten, Gittern und Nachbarschaftsstrukturen auf diesen. Es stellt einen generischen Rahmen für Markov-Ketten-Simulationen (Markov Chain Monte Carlo Methods, MCMC) und für Markov-Felder bereit. Höchste Priorität haben auf zweidimensionalen Daten basierende Simulationen, wie sie in der Bayes-Bildanalyse vorkommen. Obwohl Oberon ein Benutzerinterface mit eigenen Stärken mitbringt, wurde Ants mit einer Vielzahl von (nach außen kaum sichtbaren) Tools ausgestattet, die eine Einbettung in das umgebende Betriebssystem erleichtern und dem unerfahrenen Anwender eine sofortige intuitive Benutzung erlauben.

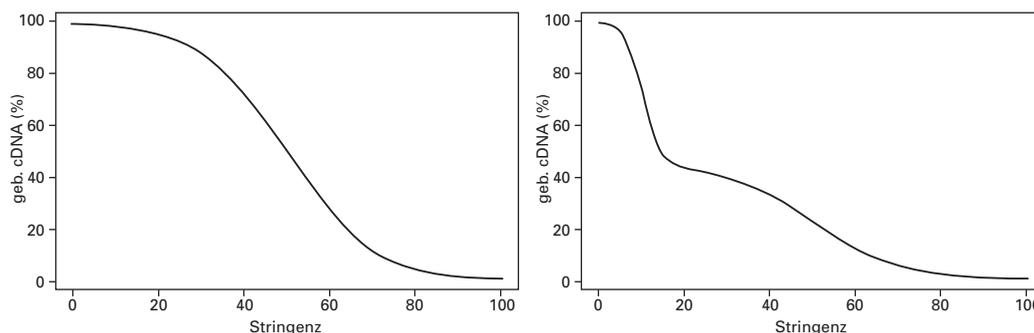
### Ants in den Lebenswissenschaften

In diesem Abschnitt werden einige Anwendungen auf Projekte in der GSF vorgestellt,

speziell Beispiele aus der Genomanalyse und der medizinischen Bildanalyse.

#### Qualitätskontrolle von Microarrays

Ein Microarray dient, für eine große Anzahl von Genen gleichzeitig, der quantitativen Bestimmung von mRNA in speziellen Zellarten oder Geweben. Einsträngige Teilstücke komplementärer DNA (cDNA) werden in den bis zu 16 000 Mikropunkten (Spots) eines Chips immobilisiert. Die cDNA wird anhand von Vorinformationen über die zu erwartenden bzw. interessanten mRNA-Sequenzen zusammengestellt. Auf jeden Spot gibt man nur eine bestimmte Art von cDNA und bringt anschließend eine kleine Menge der Probeflüssigkeit auf, etwa aus Tumorgewebe extrahierte Zellflüssigkeit. Idealerweise binden sich nun diejenigen mRNA-Sequenzen aus der Probe, die zu der cDNA des Spots komplementär sind, an diese cDNA – man spricht hier von Hybridisieren – und werden somit ebenfalls immobilisiert. Nach Entfernung aller nicht derart gebundenen mRNA-Stränge vom Chip kann man die Anzahl der gebundenen Stränge pro Spot mit Hilfe eines Fluoreszenzmarkers bestimmen. Die Intensität des Fluoreszenzsignals ist dabei idealerweise proportional zur Anzahl der an dem Spot immobilisierten mRNA-Stücke. Da die anfangs aufgebrachten cDNA-Stränge bekannt sind, lässt dies Rückschlüsse auf die Menge der komplementären mRNA in der Zelle zu. Da mRNA ein Zwischenprodukt bei der Biosynthese von Proteinen ist, die zum Beispiel als Enzyme den Metabolismus steuern, erhält man somit wichtige Informationen etwa über die



**Abb. 1:** Zu erwartende Verlaufskurven der Signalintensitäten in Abhängigkeit von der Stringenz (Fraktionierungskurven, schematisiert). Erwartungen für einen Spot ohne (links) bzw. mit Kreuzhybridisierung (rechts).

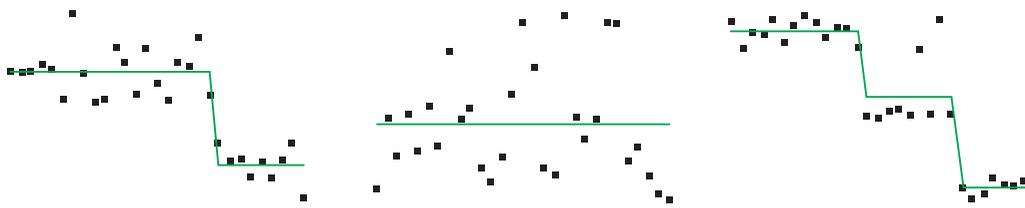


Abb. 2: Drei reale Beispielkurven für die Analyse auf Kreuzhybridisierung mit ihren Auswertungen, die auf einem Bayes-Ansatz beruhen. Links: keine Kreuzhybridisierung; Mitte: Zuordnung nicht möglich; rechts: Kreuzhybridisierung.

Besonderheiten der biochemischen Prozesse in einem Tumor.

Microarray-Experimente bieten aber nicht nur große Vorteile, sondern sind auch mit verschiedenen Fehlerquellen behaftet, die ihre genaue Auswertung erschweren. Eine dieser Fehlerquellen, die „Kreuzhybridisierung“, wurde im IBB untersucht. Kreuzhybridisierung bedeutet, dass eben nicht nur genau die komplementäre mRNA an einem Spot bindet, sondern auch noch andere, was die Intensitätsmessung erheblich verfälschen kann. Ob es zu Kreuzhybridisierung gekommen ist, lässt sich experimentell feststellen: Ein hybridisierter Microarray wird mit sukzessive ansteigenden Konzentrationen einer Formamid-Lösung behandelt und in jedem Schritt die Intensität des Fluoreszenzsignals gemessen. Bei einem Spot, an dem nur spezifische cDNA gebunden ist, erwartet man einen scharfen Abfall der Signalintensität bei einer Waschmittelkonzentration (Stringenz), die der Bindungsenergie von mRNA und Sonde entspricht. Bei Kreuzhybridisierung rechnet man mit mehreren solchen Intensitätsabfällen bei verschiedenen Konzentrationen, wobei der letzte dieser Punkte der Ablösung der spezifischen mRNA (höchste Bindungsenergie) vom Chip entspricht (s. Abb. 1).

Der Beitrag der Software Ants ist hier ein automatisiertes Erkennen der Anzahl der Intensitätsabfälle für reale Daten, deren Messwerte oftmals stark verrauscht sind (s. Abb. 2). Um Kreuzhybridisierung zu detektieren, ist offenbar ein Verfahren nötig, das aus diesen Daten das hier wesentliche Merkmal extrahiert, nämlich die Anzahl der Sprünge. Da der Abfall der Signalintensität sehr scharf ist, unterstellen wir den Daten, dass sie aus einem stückweise konstanten Signal – einer Treppenfunktion – entstanden

sind und versuchen, diese Funktion aus den Daten zu schätzen.

Konkret wird zu beobachteten Daten  $y$  ein Funktional  $H_\gamma(x, y) = \gamma U(x) + D(x, y)$  – das so genannte Potts-Funktional – angegeben, das jedem Kandidaten  $x$  eine Zahl zuordnet und einerseits umso größer ist, je mehr der Kandidat  $x$  von den Daten  $y$  abweicht ( $D(x, y)$ ), andererseits aber auch sehr groß ist, wenn  $x$  viele Sprünge hat ( $U(x)$ ). Der so genannte Hyperparameter  $\gamma$  (eine reelle Zahl) steuert dabei das Gleichgewicht zwischen diesen beiden Termen. Ist  $\gamma = 0$ , so wird nur der Abstand zu den Daten bewertet, ist  $\gamma$  hingegen sehr groß, so fällt die Bewertung der Stufen deutlich höher aus. Minimiert man nun dieses Funktional in  $x$ , so bekommt man ein Signal, das – im Sinne des durch  $\gamma$  eingestellten Gleichgewichts – den besten Kompromiss zwischen Datennähe und geringer Sprungzahl darstellt. Die beiden Terme  $U(x)$  und  $D(x, y)$  werden – im Kontext der Minimierung des Funktionals – „Strafterme“ genannt.

Am IBB wurde ein auf dynamischer Programmierung basierender Algorithmus entwickelt, der diese Minimierung des Funktionals mit hoher Geschwindigkeit zulässt. Da aber die Ergebnisse dieser Analyse wesentlich davon abhängen, welches  $\gamma$  man wählt, besteht noch das Problem, gleichzeitig für eine sehr große Anzahl an Stringenzkurven diejenigen Werte von  $\gamma$  zu bestimmen, die dem Problem angemessen sind. Hierfür wurden verschiedene Methoden, so genannte Modellwahlkriterien, erprobt; bei Simulationen und visueller Inspektion mit Ants erwies sich eine davon als geeignet: das „längste monotone Intervall“. Ergebnis: In etwa 2 Prozent der betrachteten Spots spricht die Datenlage für die Anwesenheit von Kreuzhybridisierungen, in etwa 58 Prozent der Fälle



Abb. 3: Simulationspanel für Fraktionierungskurven

dagegen. Für einen großen Teil der Spots ist keine eindeutige Aussage möglich, wozu auch die Struktur des Experiments beiträgt.

Dieses Projekt wird seit 2004 im Rahmen des Sonderforschungsbereichs 386, „Statistische Analyse diskreter Strukturen“, von der Deutschen Forschungsgemeinschaft gefördert. Für weitergehende Untersuchungen der Fraktionierungskurven, bei denen es um die Auswertung unter Variation von Datentreue- und Bestrafungsterm gehen soll, ist Ants ein wichtiges Hilfsmittel: Die Verfahren lassen sich damit an realen Daten schnell durchführen und visualisieren. Dafür wurde ein umfangreiches Simulationstool für Zeitreihen aus Microarrays erstellt (Abb. 3).

### Stochastische Modellierungen von Genexpressionsmustern

Ein wichtiger und gut studierter Signalweg in der Entwicklungsbiologie ist der Delta-Notch-Signalweg. Hauptakteure sind die beiden Proteine Delta und Notch, wobei Notch die Rolle eines Rezeptors und Delta die eines Liganden spielt. Beide Proteine sind Transmembranproteine: Sie übertragen Signale nur zwischen benachbarten Zellen.

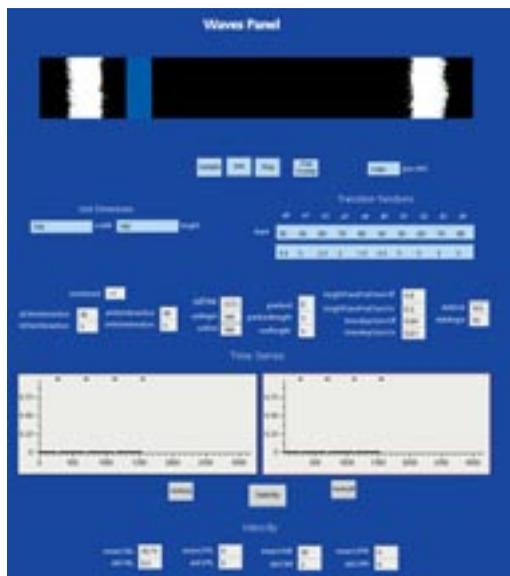
#### Laterale Inhibition

Ein über den Delta-Notch-Signalweg induzierter Prozess ist die laterale Inhibition, die in der Embryonalentwicklung vieler Orga-

nismen eine wesentliche Rolle spielt. Da Zellen mit unterschiedlichen Aufgaben ihren Ursprung in Gewebe aus anfangs äquivalenten Zellen haben, muss es einen Faktor geben, der die spätere Differenzierung steuert. Ein solcher Faktor ist die laterale Inhibition: Zellen hindern ihre Nachbarn daran, das gleiche Schicksal anzunehmen wie sie selbst. Am IBB wurde untersucht, ob Wechselwirkungen zwischen benachbarten Zellen ausreichen, um ein stabiles Verhältnis zwischen zwei verschiedenen Zellarten herzustellen. Außerdem sollten qualitative Merkmale charakterisiert werden, unter denen sich bestimmte biologische Muster einstellen. Das Gewebe wird modelliert durch ein endliches Gitter  $G$ , versehen mit einer Nachbarschaftsstruktur: Jeder Gitterpunkt in  $G$  entspricht einer von acht Nachbarn umgebenen Zelle. Eine Zelle kann dabei drei Zustände annehmen – Delta-dominant, neutral oder Notch-dominant. Die zeitliche Entwicklung des Gewebes wird anhand eines interagierenden Teilchensystems beschrieben: Zu jedem Zeitpunkt  $t \geq 0$  wird das Zellgewebe durch eine Konfiguration auf dem Gitter dargestellt. Die zeitliche Dynamik wird durch einen Markov-Prozess modelliert. Die Übergangswahrscheinlichkeiten dieses Prozesses haben wir wie folgt gewählt: Je mehr Notch-dominante Zellen in der Nachbarschaft einer Zelle liegen, desto größer ist die Wahrscheinlichkeit, dass diese sich zu einer Delta-dominanten Zelle entwickelt. Liegt andererseits eine Delta-dominante Zelle in der Nachbarschaft, so wird diese Wahrscheinlichkeit sehr klein. Die experimentelle Validierung des Modells an Simulationen mit hoher Fallzahl hat die Stabilität des Prozesses bestätigt.

#### Genexpressionswellen

Wichtige Teile der Wirbelsäule, wie Wirbelkörper, Skelettmuskeln und Nerven, entwickeln sich aus den Somiten: rundliche Zellhaufen, die in periodischen Zeiteinheiten vom unsegmentierten präsomitischen Mesoderm (PSM) abgetrennt werden (das Mesoderm ist das mittlere der drei embryonalen Keimblätter, aus denen bestimmte Organsysteme hervorgehen). Experimente haben gezeigt, dass eine Expression von Genen wellenförmig vom hinteren zum vorderen Ende des PSM läuft und dort die



**Abb. 4: Simulationspanel für Expressionswellen.** Einzelne Konfigurationen können in Echtzeit beobachtet werden. Weiße Punkte entsprechen Zellen, die ein Gen exprimieren, schwarze Punkte entsprechen Zellen, deren Expression abgeschaltet ist. Der blaue Balken in der linken Gitterhälfte simuliert ein aus dem Gitter herausgeschnittenes Segment, was ein reales Experiment zur Autonomieuntersuchung simulieren soll. In der unteren Hälfte werden Zeitreihen zu Stabilitätsuntersuchungen und weitere Parameter angezeigt.

Abtrennung der einzelnen Somiten auslöst. Ursache dieser Expressionswelle sind zyklische Gene – Gene, deren Expression periodisch an- und abgeschaltet wird. Dabei rufen nicht Zellbewegung oder diffundierende Signale die Expression hervor, sondern diese ist eine autonome Eigenschaft der PSM-Zellen. In der Entstehung der Somiten spielt der Delta-Notch-Signalweg eine wich-

tige Rolle. Eine mögliche, aber noch nicht ausreichend untersuchte Aufgabe dieses Weges ist es, die Genexpression zwischen zwei benachbarten Zellen zu synchronisieren. Anhand eines zeitdiskreten stochastischen Prozesses, ähnlich dem der lateralen Inhibition, wird untersucht, ob eine solche Welle – allein durch eine innere Uhr und durch lokale Wechselwirkungen gesteuert – stabil über einen längeren Zeitraum laufen kann (Abb. 4).

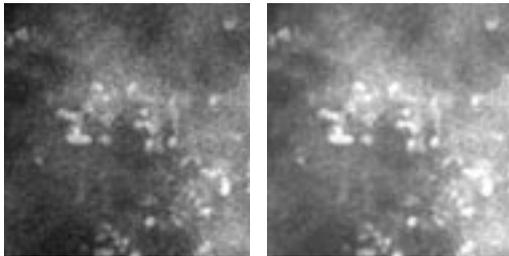
### Niedrig-Dosis-Röntgen-Mammographie

In der Analyse digitaler Bilddaten aus Röntgen-Mammographien der weiblichen Brust – zur (Früh-)Erkennung von Krebs – geht es darum, die relevanten Daten vom Rauschen zu trennen. Als Hinweise auf Krebs haben sich „Mikrokalke“ etabliert; solche Kalkeinschlüsse erscheinen auf Röntgenaufnahmen als kleine helle Flecken mit Durchmessern von 0,2 – 1 mm. Die meisten maschinellen Werkzeuge zur Nachbearbeitung bzw. Analyse digitaler Bilder versuchen, das Auffinden dieser Intensitätsmaxima zu vereinfachen oder sogar vollständig zu automatisieren. Bestehende Verfahren zielen darauf, die Qualität der Bilder zu steigern und einen „second reader“ bereitzustellen, um die Diagnose durch den Radiologen zu unterstützen.

Um die Daten für die weitere Analyse mit anderen Methoden vorzubereiten, haben wir eine möglichst unspezifische Vereinfachung vorgeschlagen: Die Daten sollen einerseits geglättet werden, andererseits aber sollen Kanten erhalten bleiben. Wie bei den Stringenzkurven im Beispiel Microarray bietet sich auch hier das Potts-Modell an, allerdings angewendet auf Bilder, also zweidimensionale Daten. Auch hier besteht das



**Abb. 5: Zweidimensionales Pseudopotts-Modell zu Interaktionsparameter  $\gamma = 1$ ; Original, optimale Zeilen, optimale Spalten und Mittelwertbild über Zeilen und Spalten.**



**Abb. 6: Mammographie; links: Originalbild, rechts: Pseudopotts-Glättung zu  $\gamma = 1$ .**

Potts-Modell aus konkurrierenden Termen für Datentreue und Anzahl der Kanten (Sprünge im Zweidimensionalen sind Kanten). Da es aber für das zweidimensionale Potts-Modell keinen exakten und schnellen Minimierungsalgorithmus gibt, wurde ein abgeändertes Funktional vorgeschlagen, für das sich ein solcher angeben lässt. Dabei werden die Bilder optimal in Zeilen und Spalten unterteilt (Abb. 6).

Wie die visuelle Inspektion der Mammographien zeigt, unterdrückt der Algorithmus das Rauschen bei Erhaltung der wesentlichen Strukturen. Soll das Verfahren jedoch vollautomatisiert ablaufen, müssen die Glättungsparameter an die in Röntgenmammographien stark variierenden Rauschcharakteristiken adaptiert werden. Hier besteht noch erheblicher Forschungsbedarf.

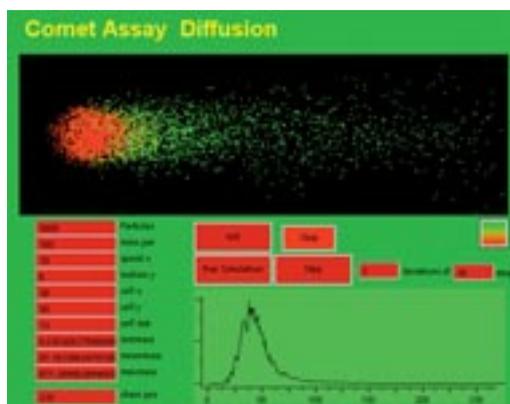
Zeilen- und spaltenweises Vorgehen erzeugt Artefakte im entrauschten Bild und erschwert die Adaption an diagonale Kanten. Dieser Missstand ließe sich möglicherweise durch hierarchische Unterteilung in Rechtecke und Dreiecke verschiedener Winkel beheben. Ein solches Alternativmodell wurde bereits entwickelt und implementiert, jedoch noch nicht zur Analyse von Mammographien herangezogen (Abb. 7).

### Comet Assay

Mit dem 1984 entwickelten und erst seit rund zehn Jahren üblichen Verfahren „Comet Assay“ wird die Degradierung von Zellkernen bzw. DNS-Molekülen gemessen. Eine nicht degradierte DNS ist eine lange lineare Kette von Molekülen, die bei Bestrahlung in viele kleine Stücke zerfällt; die Masse jedes Stücks ist proportional zu seiner Länge. Mittels Gelelektrophorese kann man auf die Massenverteilung dieser Stücke schließen. In einer Studie soll die Rekombinationsfähigkeit bestimmter Zellen quantifiziert werden. Dazu vergleicht man gesunde und bestrahlte Zellen nach unterschiedlich langer Ruhezeit. Für dieses Experiment wurde in einer am IBB angefertigten Diplomarbeit ein stochastisches Modell entwickelt, um die statistische Analyse abzusichern (Abb. 8). Simulationen, die mit Ants interaktiv ausge-



**Abb. 7: Pseudopotts-Glättung unter Einbeziehung verschiedener Winkel zu Interaktionsparametern  $\beta = 0, 500, 1000$  und  $5000$  (v.l.n.r.).**



**Abb. 8: Simulationspanel zum Comet-Assay.** Die Masse der DNS-Stücke ist farbig codiert (rot: schwer, grün: leicht).

führt werden können, erlauben, die Gültigkeit des jeweils gewählten Modells intuitiv zu beurteilen und geeignete Parameter zu wählen.

### fMRI-Brain-Mapping

Ziel des Brain-Mappings ist es, Regionen des menschlichen Gehirns zu identifizieren, die durch Stimulation aktiviert wurden. Typische Stimulationen sind Box-Car-Signale, wie Licht oder Ton an-aus, oder auch taktile Reize. Als Maß für die Aktivität des betreffenden Hirnareals gilt der Sauerstoffgehalt des Blutes, indirekt gemessen über funktionelle Magnetresonanztomographie (fMRI). Da die eingehenden Signale stückweise konstante Zeitreihen sind, werden sie mit einem Potts-Funktional (einem Modell für

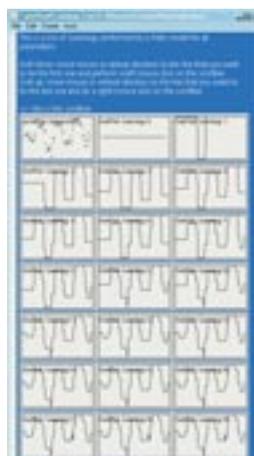
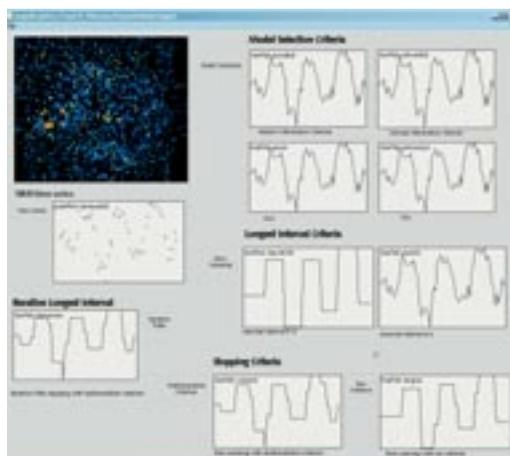
stückweise konstante Daten) adäquat beschrieben. Das Signal wird auf seinem Weg Auge – Gehirn – Messinstrument aber auf äußerst komplizierte Weise verändert. Entsprechend hoch komplex sind Verfahren, die viele dieser Einflussfaktoren in die Beurteilung einbeziehen, ob ein Hirnareal durch das gegebene äußere Signal aktiviert wurde.

Es lag die Frage nahe, ob eine stückweise konstante Glättung der ausgehenden Signale nicht ein einfaches Verfahren wäre, Aktivitätsregionen schnell zu detektieren. Auch hier bedarf es einer automatisierten Wahl des Interaktionsparameters im Potts-Funktional. Mit Hilfe von Ants entstand eine leicht zugängliche Plattform zur Analyse der Hirnzeitreihen. Aktivierte Areale werden durch Auszählen der Glättungs-Maxima visualisiert, was – zusammen mit der Möglichkeit, interaktiv Voxel und deren Zeitreihen auszuwählen – zu einem tieferen Verständnis der Methode führt (Abb. 9).

Auch wenn sich die hier eingesetzte Methode nicht mit ausgefeilten Techniken aus dem Bereich der Hirnsegmentierung messen will, so besticht sie doch durch ihre Robustheit und hohe Geschwindigkeit. Aktivierte Hirnareale konnten erkannt werden.

### Ants in der Lehre

Die Vermittlung komplexer Zusammenhänge in der Stochastik ist ein weiteres Anwendungsfeld von Ants. Beispielhaft wurde das Paket als Demonstrationssoftware zum Buch „Image Analysis, Random Fields and Mar-



**Abb. 9: Panel zur Zeitreihenanalyse auf Hirndaten.** Rechts ist ein komplettes  $\gamma$ -Scanning zu einem Voxel zu sehen, die eingefärbten Plots korrespondieren zu verschiedenen Modellselektionsverfahren.

kov Chain Monte Carlo Methods“ von G. Winkler eingesetzt (Abb. 10).

Die Beispiele sind auf Panels organisiert, die sich dem Nutzer als interaktive Fenster präsentieren. Um das Verständnis für stochastische Verfahren zu vertiefen, sind in Ants Beispiele implementiert, die beim interaktiven Betrachten sowohl Eigenschaften wie auch Grenzen der Methoden erkennen lassen (Abb. 11).

Im Folgenden wird eine Auswahl an Einsatzmöglichkeiten in der Vermittlung wissenschaftlicher Inhalte der Stochastik vorgestellt.

### Der Gibbs-Sampler auf dem Potts-Modell

Das Potts-Modell ist das einfachste stochastische Zufallsfeld-Modell mit lokalen Wechselwirkungen auf einem Gitter, das mathematisch nichttrivial und in realen Anwendungen von Bedeutung ist. Das Ising-Modell ist ein auf zwei Zustände (Schwarz-weiß-Bilder) beschränktes Potts-Modell. Der in den Abschnitten „Mikroarray“ und „Brain-Mapping“ beschriebene eindimensionale Fall nimmt insofern eine Sonderstellung ein, als dort für das Potts-Funktional ein schneller, deterministischer Minimierungsalgo-

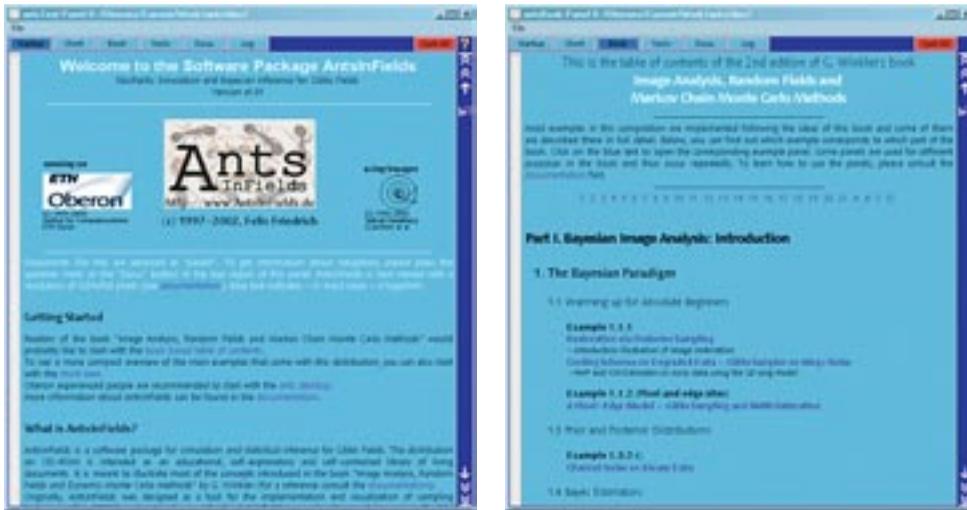


Abb. 10: Startoberfläche und interaktives Inhaltsverzeichnis als Zugang zu vielen Beispielen



Abb. 11: Typisches Beispielpanel und interaktiver Hilfetext

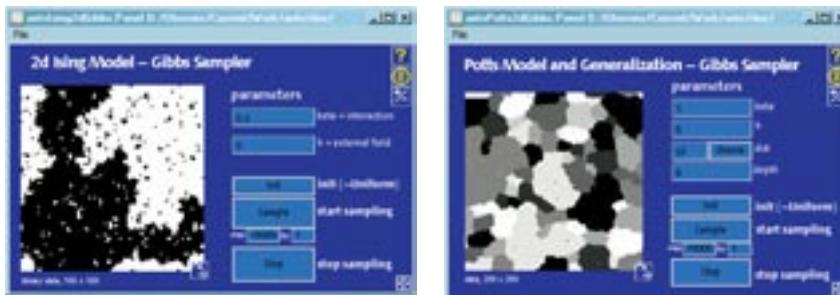


Abb. 12:  
Panels zur  
Simulation des  
Ising- und des  
Potts-Modells

rithmus existiert. Für höhere Dimensionen gibt es einen solchen Algorithmus nicht. Die Komplexität des Minimierungsproblems wächst sogar exponentiell mit der Anzahl der Pixel. Daher benötigt man stochastische Algorithmen (Markov Chain Monte Carlo Methoden, MCMC), mit denen solche Minimierungsprobleme zu lösen sind. Ein solcher Algorithmus, der es erlaubt, sehr viele Zufallsfelder auszuwürfeln – also zu simulieren – ist der Gibbs-Sampler (Abb. 12).

Zum Verständnis dieser Methoden und der Modelle trägt es wesentlich bei, einen laufenden Simulationsprozess interaktiv zu beobachten. So kann man beim Experimentieren mit dem Gibbs-Sampler mit verschiedenen Parametern ein Gefühl für die Stär-

ken und Schwächen des jeweiligen Modells entwickeln.

### Markov-Ketten

Einfache Modelle, wie der eindimensionale Random Walk, veranschaulichen grundlegende mathematische Eigenschaften der Markov-Ketten-Methoden (hier: Konvergenz zur stationären Verteilung). Anwendungen auf klassische Probleme wie das berühmte Handelsreisenden-Problem verdeutlichen Grenzen der Sampler (hier: die geringe Konvergenzgeschwindigkeit) und rufen zum kritischen Umgang mit solchen Algorithmen auf. Beispiele aus der Bayes-Bildanalyse hingegen demonstrieren deren Einsetzbarkeit und Nutzen (Abb. 13).

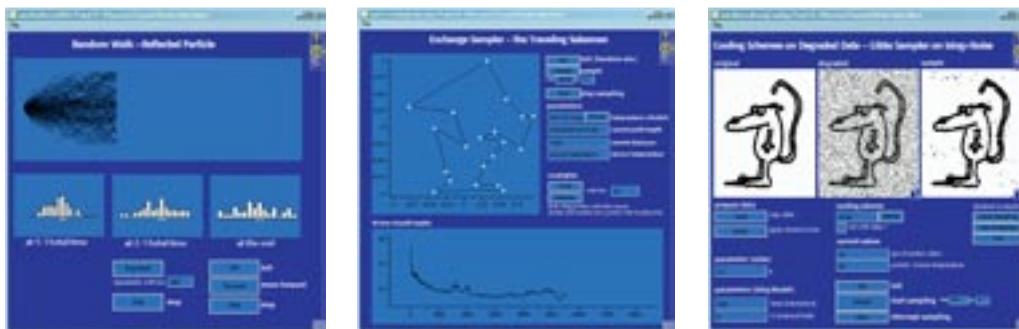


Abb. 13: Reflektierender Random Walk, Annealing auf dem Handelsreisenden-Problem und in der Bayes-Bildanalyse.

## ■ Ausgewählte Veröffentlichungen

Drobyshev, L., Machka, C., Horsch, M., Seltmann, M., Liebscher, V., Hrabé de Angelis, M., Beckers, J.: Specificity assessment from fractionation experiments (SAFE): a novel method to evaluate microarray probe specificity based on hybridisation stringencies. *Nucleic Acids Res.* 31, E1-1 (2003)

Friedrich, F., Winkler, G., Wittich, O., Liebscher, V.: An Elementary Rigorous Introduction to Exact Sampling. *Proceedings of the International Workshop in Applied Probability*. Caracas, Venezuela, 14. 17.1.2002, erscheint bei Kluwer Scientific Publishing (2003)

Martin, A., Zeiser, S.: Stochastic Models for Delta-Notch Regulation. *Satellite Meeting to the 24th European Meeting of Statisticians in Prague*, 14.8.-16.8.2002, München

Winkler, G., Liebscher, V.: Smoothers for Discontinuous Signals. *J. Nonparametric Statistics* 14, 203-222 (2002)

Winkler, G.: *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods. Applications of Mathematics* 27, 2nd ed., Springer, Berlin (2003)