

Relationale Wachstumsgrammatiken: Ein Formalismus zur Spezifikation multiskalierter Struktur-Funktions-Modelle von Pflanzen

Winfried Kurth, Gerhard Buck-Sorlin, Ole Kniemeyer

Pflanzenmodelle: Einige Entwicklungstendenzen

Modelle in den Agrarwissenschaften sollen möglichst aussagekräftige, für die Praxis relevante Ergebnisse liefern. Gleichzeitig wird erwartet, dass sie möglichst auch noch unter sich wandelnden Umweltbedingungen, unter veränderten Anbauformen oder bei Verwendung neuer Sorten ihre Gültigkeit behalten oder zumindest anpassbar sind. Dazu müssen agrarische Modelle in hohem Maße der Komplexität des Systems "Pflanze" und ihrer Umwelt gerecht werden. Es ist daher nicht verwunderlich, dass in den letzten Jahrzehnten auch immer komplexere Modelle konstruiert worden sind.

Drei "Extrem-Typen" von Modellen lassen sich nennen: Zum einen gibt es Modelle, die sich ganz auf chemische und physikalische Prozesse, Energie- und Stoffflüsse konzentrieren. Die räumliche Struktur von Pflanzen oder von agrarischen Beständen wird bei diesem Modelltyp oft stark vernachlässigt oder fließt nur in Form einer groben Kompartimentierung (Spross, Wurzel, Boden, Atmosphäre) in die Modellstrukturen ein. Jedoch wird hier versucht, die *Funktion* von Pflanzen (in Wechselwirkung mit ihrer Umwelt) zu erfassen – in manchen Modellen bis hinunter zu Prozessen auf der zellulären oder gar molekularen Ebene, wenn man etwa an bekannte Photosynthesemodelle denkt (z.B. von Caemmerer & Farquhar 1981). Ein anderer Typ von Modellen ist fokussiert auf die Triebbildung und morphologische Struktur der Pflanze. Phänomene wie wechselseitige Beschattung, Konkurrenz um Raum oder um Bodenressourcen (etwa in der Agroforestry) machen diesen Modelltyp auch für ertragsbezogene Fragestellungen interessant. Eine bedeutende "Dynastie" solcher Strukturmodelle entstand seit den 70er Jahren am agrarischen Forschungsinstitut CIRAD in Montpellier (de Reffye & Snoeck 1976, de Reffye et al. 1995), auch unter dem Einfluss der dortigen bedeutenden Botaniker an der biologischen Fakultät. Allerdings wurden in diesen Modellen wiederum kaum kausale Zusammenhänge und Prozesse berücksichtigt. Ein dritter Modelltyp schließlich arbeitet mit aggregierten Größen über Pflanzenbestände und hat typischerweise die Form eines statistischen Ansatzes; es wird hier auf empirische Erhebungen zurückgegriffen, die auf der räumlichen Skala oberhalb der Auflösungsebene typischer Struktur- oder Prozessmodelle gewonnen wurden. Die drei Modell-Typen lassen sich in einem "Modell-Dreieck" anordnen (Abb. 1).

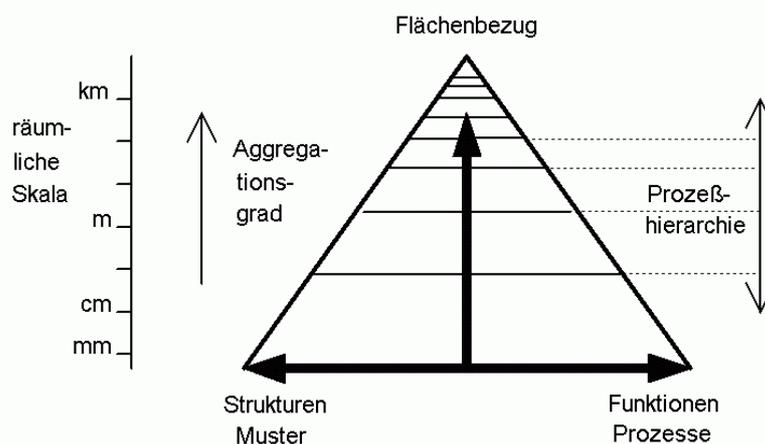


Abbildung 1: Schema der Modelltypen

Es ist klar, dass eine "gesunde Mischung" dieser drei Extreme erstrebenswert ist, um den eingangs erwähnten Anforderungen entgegenzukommen. Auf der Basis-Seite des "Modell-Dreiecks" findet man "Struktur-Funktions-Modelle" (engl. *functional-structural plant models*, FSPM), deren Entwicklung ca. in den letzten 10 Jahren einen bedeutenden Aufschwung erfahren hat (Godin et al. 2004). Die Herausforderung bei diesen Modellen liegt darin, Stoffvorräte und Prozesse auf einzelne pflanzliche Substrukturen zu verteilen und dezentral zu simulieren, wobei zugleich die Wechselwirkungen zwischen diesen Substrukturen zu berücksichtigen und geeignet zu repräsentieren sind. In diesem Beitrag soll ein Formalismus vorgestellt werden, der derartige Repräsentationen räumlicher und zugleich funktionaler Zusammenhänge in sich entwickelnden Organismen oder Beständen unterstützt. Aber

auch die "vertikale Richtung" im "Modell-Dreieck" stellt eine eigene Herausforderung dar: Das Überwinden der Grenzen zwischen Skalenebenen; insbesondere der Versuch, im Simulationsmodell "emergente Phänomene" wiederzufinden, die auf höheren Skalenebenen entstehen als Resultat von Gesetzmäßigkeiten und Prozessabläufen auf niedrigeren Ebenen (Breckling et al. 2005). Wir werden am Beispiel des Modells einer Gerstenpflanze, bei deren Wachstumssimulation und Strukturbildung zugleich auf ein molekulares Netzwerk der Synthese eines Pflanzenhormons zurückgegriffen wird, zeigen, dass unser formales System zur Modellspezifikation auch solche Verknüpfungen zwischen unterschiedlichen Skalenebenen auf konsistente Weise repräsentieren kann.

Die heute gängigen Struktur-Funktions-Modelle von Nutzpflanzen (auch unter der Bezeichnung *Virtual Plants* oder *Virtual Crops* bekannt; Hanan 1997, Wernecke et al. 2000) haben üblicherweise die Form größerer Programmsysteme, die in einer objektorientierten, universellen Programmiersprache (wie C++, Java, Delphi, Simula) implementiert sind (z.B. Perttunen et al. 1996). Die "Objekte" im Sinne der Programmierung sind dabei in der Regel die kleinsten Pflanzenstrukturen (Module), die in dem jeweiligen Modell betrachtet werden – beispielsweise Blätter, Sprosssegmente, Wurzelsegmente; mitunter auch Zellen des umgebenden Bodens oder der Atmosphäre (Breckling 1996). Ein Problem bei diesen "ad hoc" programmierten Struktur-Funktions-Modellen besteht darin, dass ihre Spezifikation sehr unübersichtlich ist. Sie besteht typischerweise aus vielen tausend Zeilen Code und wird nur von wenigen Menschen – oftmals nur vom Autor des Programms – vollständig verstanden. Dies erschwert die Wartbarkeit, Wiederverwendbarkeit und unabhängige Validierung des Modells und kann zudem abschreckend wirken für Biologen oder Agrarwissenschaftler mit nur rudimentären Kenntnissen in objektorientierter Programmierung und Softwaretechnik. Ein Ziel in unseren langjährigen Forschungsanstrengungen, die wir teilweise in Zusammenarbeit mit Agrar- und Forstwissenschaftlern unternommen haben, war daher die Schaffung einer regelbasierten "Metasprache", die auf die Bedürfnisse der Pflanzenmodellierung zugeschnitten ist, von Biologen relativ leicht erlernt werden kann und es erlaubt, eine Modellspezifikation zu ändern, ohne dass in den Computercode der ausführenden Software (der weiterhin in einer objektorientierten Universalsprache geschrieben ist) eingegriffen werden müsste (Abb. 2). Ein solcher "generischer Ansatz" für die Modellspezifikation erhöht auch die Transparenz, Vergleichbarkeit und wechselseitige Kombinierbarkeit von Modellen.

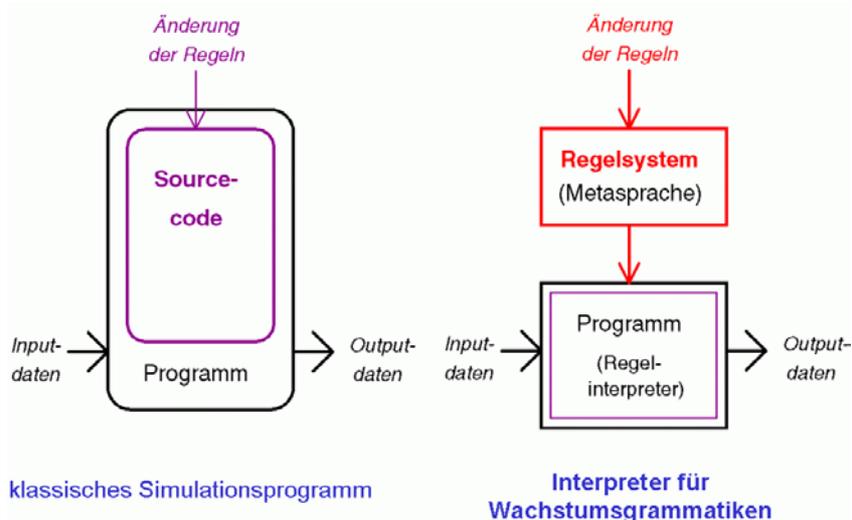


Abbildung 2: **Generischer Simulationsansatz mit Metasprache (rechts) im Vergleich zum klassischen Ansatz (links)**

Ein Formalismus, der dieser Anforderung nach einer Modellspezifikationsprache für die Pflanzenmodellierung nachkommt, existiert bereits und wird weltweit von verschiedenen Pflanzenwissenschaftlern, auch im Agrarbereich, eingesetzt: Es handelt sich um die sogenannten Lindenmayer- oder L-Systeme, benannt nach dem Botaniker Aristid Lindenmayer (Lindenmayer 1968). Es handelt sich um Ersetzungssysteme für Zeichenketten – ähnlich den formalen Grammatiken, die man zur Beschreibung der korrekten Syntax von Programmiersprachen oder natürlichen Sprachen verwendet. Einzelne Symbole in diesen Zeichenketten werden geometrisch interpretiert und entsprechen pflanzlichen Struktureinheiten. In den Ersetzungsregeln lassen sich botanische Gesetzmäßigkeiten (wie z.B. Apikaldominanz, physiologische Alterung von Meristemen, Bildung von Reiterationen) relativ einfach ausdrücken. Mit Hilfe kontextsensitiver Regeln können auch Transportprozesse in der Pflanze und die Wirkung von Pflanzenhormonen nachgebildet werden (Prusinkiewicz & Lindenmayer 1990). Es

existieren etablierte Softwaresysteme zur Interpretation von L-Systemen, z.B. L-Studio (Karwowski & Prusinkiewicz 2004), Graphtal (Streit 1992), Grogra (Kurth 1994).

Jedoch haben L-Systeme noch einige Nachteile, insbesondere hinsichtlich der Transparenz und Einfachheit der Verwendung, wenn es um die Spezifikation komplexer Pflanzenmodelle mit mehreren Skalenebenen und unter Einbeziehung von Prozessen und genetischer Kontrolle geht. Prusinkiewicz (2004) räumt ein, dass die Einbeziehung genetischer Mechanismen in bisherige "virtuelle Pflanzen" (auf L-System-Basis) noch ein offenes Problem darstellt. Kritisch ist an L-Systemen insbesondere, dass sie mit Zeichenketten arbeiten, die vor ihrer Verwendung im Pflanzenmodell zunächst in dreidimensionale geometrische Strukturen transformiert werden müssen (Abb. 3 links).

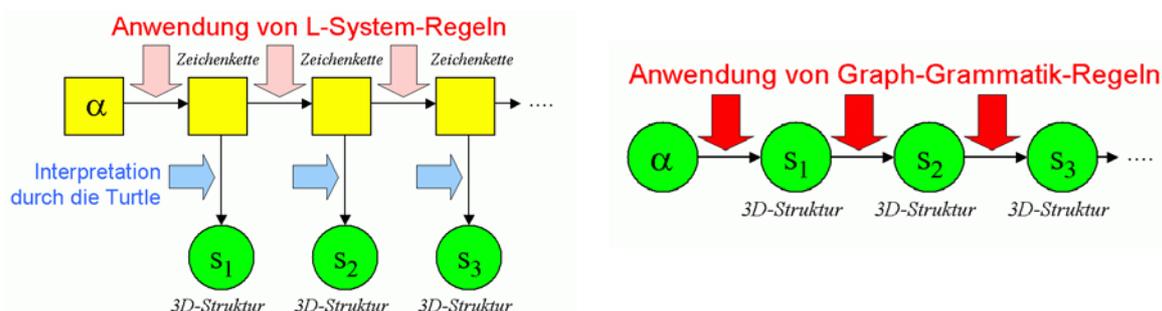


Abbildung 3: **Unterschiedliche Arbeitsweise von klassischen Lindenmayer-Systemen (links) und Graph-Grammatiken (relationalen Wachstumsgrammatiken; rechts) in der Pflanzenstrukturmodellierung**

Indem wir das regelbasierte, mit dem Ersetzen von Substrukturen arbeitende Programmierparadigma (zu Paradigmen der Programmierung im Zusammenhang mit Pflanzenmodellen siehe Kurth 2002) von Zeichenketten auf Graphen (im Sinne der mathematischen Graphentheorie) erweitert haben, in deren Knoten zugleich weitere Informationen, z.B. geometrische Eigenschaften der modellierten Organe, gespeichert werden können, lässt sich der Umweg über die Codierung als Zeichenkette weitgehend vermeiden (Abb. 3 rechts). Nur noch das Regelsystem selbst wird, zum Zweck der einfachen Lesbarkeit und Editierbarkeit, in Textform eingelesen.

Relationale Wachstumsgrammatiken

Die von uns eingeführten "Relationalen Wachstumsgrammatiken" (RGG) gehören zu den Graph-Grammatiken (Kniemeyer et al. 2004). Wie in Abbildung 3 angedeutet, können sie dreidimensionale Strukturen von Pflanzen auf direktere Weise generieren als L-Systeme. Die verwendeten Graphen bestehen aus Knoten und gerichteten Kanten, die diese Knoten verbinden. Dabei können unterschiedliche Kantentypen verwendet werden (z.B. a und b in Abb. 4). Mathematisch handelt es sich um verschiedene *Relationen* auf der Knotenmenge – daher der Name "relationale Wachstumsgrammatiken". Einer dieser Kantentypen kann z.B. für eine Nachfolgerrelation zwischen Internodien einer Sprossachse stehen, ein anderer für eine skalenübergreifende "ist Teil von"-Relation, wieder ein anderer für eine Reaktionsbeziehung zwischen Stoffwechselsubstraten oder für die Relation zwischen einem Gen und einem durch dessen Expression entstehenden Substrat. Auch komplexe Beziehungen, wie die zwischen Genotyp und Phänotyp, können somit prinzipiell mit derselben formalen Einfachheit beschrieben werden wie die topologische Nachbarschaft pflanzlicher Module in klassischen L-Systemen. Zudem ist, anders als bei L-Systemen, die Repräsentation von beliebigen Netzwerken – auch solcher mit Rückkopplungsschleifen – möglich. Als Spezialfälle lassen sich in Form von Graphen auch einfachere Datenstrukturen wie Mengen, Multimengen, Listen und Zeichenketten darstellen; somit ist insbesondere alles, was mit L-Systemen modelliert werden kann, auch durch RGG darstellbar.

Eine relationale Wachstumsgrammatik ist eine Menge von RGG-Regeln. Jede dieser Regeln legt fest, dass ein bestimmter Graph (die linke Regelseite) durch einen anderen Graphen (die rechte Regelseite) zu ersetzen ist. Zusätzlich können Kontexte angegeben werden, deren Vorhandensein in einem Graphen Voraussetzung für die Anwendung der Regel ist, sowie (ebenfalls optional) ein Stück prozeduraler Code, der bei jeder Anwendung der Regel ausgeführt wird. Die Regel, die in Abbildung 4 oben in grafischer Form dargestellt ist, könnte in textueller Form z.B. geschrieben werden als:

$$i -b-> j -a-> k -a-> i \implies j \{P\};$$

darin ist P eine Folge von Anweisungen in der Sprache Java (bzw. genauer in der von uns entwickelten Sprache XL, welche Java erweitert), die in der Grafik nicht dargestellt ist. Die Anwendung der

Regel auf einen Graphen wird im unteren Teil von Abbildung 4 veranschaulicht; sie besteht in diesem Fall aus der Ersetzung dreier Knoten durch einen von ihnen. Wie man sieht, erfordert jede Regelanwendung ein Wiederfinden (*matching*) der linken Regelseite im aktuell vorhandenen Graphen.

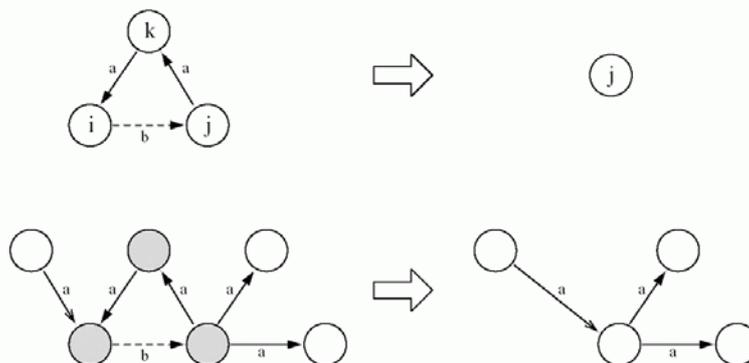


Abbildung 4: Eine RGG-Regel, dargestellt in grafischer Form (oben), und ihre Anwendung auf einen Graphen: Die schattierten Knoten werden ersetzt (unten). Aus Kniemeyer et al. (2004).

Um die Schnittstelle zwischen der regelbasierten Modellbeschreibung und prozeduraler Modellierung von Prozessen noch flexibler zu gestalten, lassen wir neben dem Einbau von Skripten $\{ P \}$ in die Regeln auch zu, dass die Knoten unserer Graphen komplette Objekte im Sinne der objektorientierten Modellierung sind und durch Java-Klassen beschrieben werden. (Insbesondere kann es sich auch um geometrische Objekte handeln – Kugeln, Zylinder, oder z.B. auch Splineflächen.) Dies führt zur Definition der Programmiersprache XL ("extended L-system language"). Neben den bisher vorgestellten Möglichkeiten bietet sie auch noch eine Zusammenfassung von Teilmengen von Regeln in Blöcken an, um die Reihenfolge der Regelanwendung flexibler steuern zu können. Auch gibt es den Regeltyp der "Aktualisierungs-Regel", geschrieben in der Form "linke Seite $::> P$ ", bei deren Anwendung der Graph, der die linke Seite *matcht*, nicht durch einen anderen Graphen ersetzt, sondern lediglich der prozedurale Code P ausgeführt wird, was oft zum Aktualisieren von Knotenparametern benutzt wird. Einige weitere Einzelheiten zu XL und Beispiele findet man bei Kniemeyer (2004) und unter der Webadresse www.grogra.de; eine vollständige Sprachspezifikation wird Bestandteil der zur Zeit in Arbeit befindlichen Dissertation des dritten Autors sein. Dort wird auch die Modellierungs- und Simulationssoftware GroIMP (Growth-grammar related Interactive Modelling Platform) näher beschrieben werden, die der Ausführung von XL-Code dient und zudem Möglichkeiten der interaktiven Visualisierung, der Manipulation erzeugter virtueller Pflanzenstrukturen und der Kopplung mit anderen Softwaretools dient.

Somit basieren relationale Wachstumsgrammatiken auf dem etablierten regelorientierten Modellierungs-Formalismus der L-Systeme, erweitern aber die zugrundeliegenden Datenstrukturen von Zeichenketten auf Graphen, die Zeichenkettenersetzung auf Graph-Ersetzung, und stellen eine Verbindung zum prozeduralen und zum objektorientierten Programmierparadigma her, indem eine konventionelle objektorientierte Programmiersprache (Java) mit eingebettet wird. Durch die Erweiterung auf Graphen unterscheiden sie sich von dem verwandten Ansatz der Sprache "L+C" (Karwowski & Prusinkiewicz 2003), bei dem ebenfalls die Mächtigkeit in der Modellierung durch Einbeziehung einer konventionellen Hochsprache (C++) vergrößert wird, der aber weiterhin auf L-Systemen beruht. Die Nützlichkeit des RGG-Formalismus demonstrieren wir im Folgenden am Beispiel des Modells "BarleyBreeder", welches auf der Plattform GroIMP ausgeführt wird und virtuelle Gerstenpflanzen generiert. Das Modell besteht aus mehreren Teilen, die alle in unserem regelbasierten Formalismus spezifiziert sind:

- eine Regelmenge, die genetische Prozesse (Mutation, Rekombination) beschreibt, welche einen jeder Gerstenpflanze zugeordneten, vereinfachten Genotyp beeinflussen,
- eine Menge von Regeln und Funktionen, die ein regulatorisches Netzwerk für die Synthese eines Pflanzenhormons (bioaktive Gibberellinsäure, kurz GA_1) und dessen Transport in der Gerstenpflanze spezifizieren,
- eine Menge von Regeln zur Morphogenese von Pflanzenorganen (Internodien, Blätter, Blüten),
- schließlich eine Interaktionsmöglichkeit mit dem Benutzer, der zwei Elternindividuen für eine virtuelle Züchtung von Nachkommen bestimmen kann.

Das RGG-Modell wird hier unvollständig skizziert; näheres findet man bei Buck-Sorlin et al. (2005).

Repräsentation eines regulatorischen Netzwerks

Die Vorteile der Einbeziehung regulatorischer Netzwerke in Struktur-Funktions-Modelle von Pflanzen liegen darin, dass auf diese Weise Prozesse des Signaltransports in der Pflanze, Rückkopplungsschleifen der Genexpression und des Metabolismus, durch Transportvorgänge oder Genexpressionsmuster bewirkte zeitliche Verzögerungen und kausale Zusammenhänge in ein kohärentes Bild integriert werden können. Dies kann der Grundlagenforschung helfen, die biochemische Architektur von Pflanzen besser zu verstehen, und wird – auch in Bezug auf agrarische Anwendungen – dazu beitragen, die Reaktionen herkömmlicher und genetisch veränderter Organismen auf Stimuli besser einschätzen zu können.

Metabolische Netzwerke werden in der aktuellen Version der Sprache XL dargestellt, indem jede Reaktion durch eine Regel repräsentiert wird. Eine gerade in Arbeit befindliche, neue Version der Software GroIMP wird es auch ermöglichen, Netzwerke aus einer Datenbank einzulesen, als Graphen zu visualisieren, interaktiv zu editieren und die entsprechenden RGG-Regeln, die die von dem Netzwerk gesteuerte Dynamik der Substratkonzentrationen simulieren, automatisch zu erzeugen (Zhao 2006).

Wir haben uns für unser Beispiel auf einen kleinen Ausschnitt aus dem komplexen Netzwerk hormonaler Reaktionen, die das Wachstum der Gerste beeinflussen, beschränkt. Der Gibberellinsäure-(GA-) Stoffwechselfad, der in unserem Modell repräsentiert wird, stellt einen vereinfachten Ausschnitt aus dem gegenwärtigen Wissensstand über pflanzliche biosynthetische Stoffwechselfade dar. Er besteht aus drei metabolischen Substraten: Bioaktive GA_1 und ihre Vorstufen GA_{19} und GA_{20} . GA_{19} wird im Apikalmeristem und an der Blattbasis produziert; die Menge hängt ab von der Zeit und vom Einfluss bestimmter Gene, deren Fehlfunktion einen Zwergwuchs der Pflanze induziert (Olszewski et al. 2002). Unter der katalytischen Wirkung des Enzyms GA_{19} -Oxidase (mit einer gewissen Steuerung durch die Tageslänge, welche hier vernachlässigt wird) wird GA_{19} über das Zwischenprodukt GA_{20} in die bioaktive GA_1 umgewandelt, wobei der letzte Reaktionsschritt durch das Enzym 3β -Hydroxylase katalysiert wird. GA_1 kontrolliert die Internodienstreckung und wird durch 2β -Hydroxylase zu einem inaktiven Kataboliten abgebaut. Andererseits konkurriert GA_1 mit GA_{19} um die Bindungsregion der GA_{19} -Oxidase, wirkt somit kompetitiv-hemmend auf die Produktion von GA_{20} und somit letztlich auf ihre eigene Produktion. Die reaktionskinetischen Parameter V_{max} und K_m für die Reaktionen im Netzwerk wurden teils aus der Literatur entnommen, teils aus frei zugänglichen Enzym-Datenbanken wie BRENDA (www.brenda.uni-koeln.de), wobei allerdings durch Ungenauigkeiten der Werte und Wissenslücken noch Spielraum für ein Modell-Fitting blieb. Die Parametrisierung ist nur als vorläufig anzusehen, zumal andere wichtige Einflussfaktoren auf das Streckungswachstum unberücksichtigt blieben. Als eine erste, vereinfachende Annahme über den Transport der Substrate innerhalb der Pflanze haben wir einen basipetalen Transport von GA_{19} und einen akropetalen Transport von GA_1 entlang der Sprossachse implementiert. Das Schema des berücksichtigten Teil-Netzwerks der GA-Biosynthese wird in Abbildung 5 gezeigt.

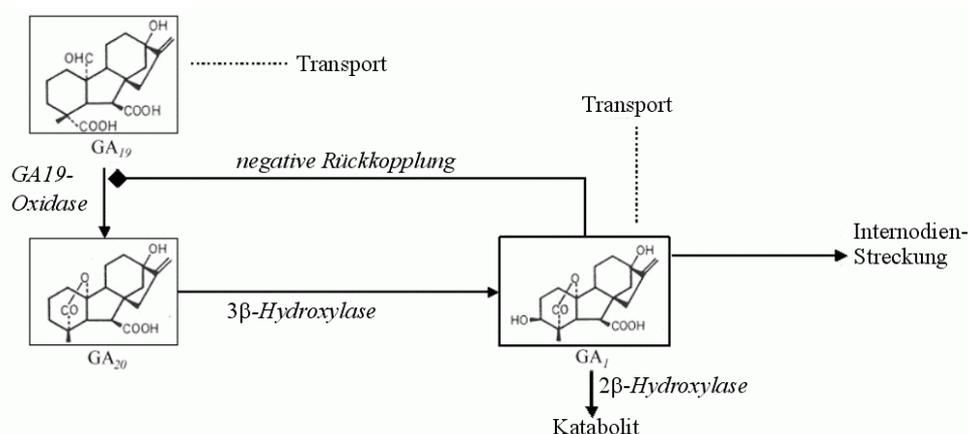


Abbildung 5: Vereinfachtes Schema der Gibberellinsäure- (GA-) Biosynthese, wie es im Gerstenmodell verwendet wurde. Es besteht aus den drei Stoffwechselprodukten GA_{19} , GA_{20} und GA_1 (sowie aus dessen Abbauprodukt) und aus drei Enzymen. GA_1 ist bioaktiv; es kontrolliert unter anderem die Internodienstreckung. GA_{19} und GA_1 werden innerhalb der Pflanze transportiert; GA_1 konkurriert mit GA_{19} um eine Andockstelle am Enzym, das GA_{20} produziert, und wirkt somit kompetitiv hemmend auf seine eigene Produktion zurück. (Aus Buck-Sorlin et al. 2005.)

Die regelbasierte Darstellung der Reaktionsschritte im Netzwerk sei am Beispiel der Reaktion, die GA_{20} liefert und die Feedback-Kontrolle durch GA_1 beinhaltet, gezeigt: Hierbei steht der Graph-Knotentyp "Cell" nicht für eine biologische Zelle, sondern für eine strukturell-morphologische Einheit im Sinne der Modellierung, in welcher sich die sämtlichen Reaktionen des Netzwerks abspielen und die an den Transportprozessen teilhat. Die Substratkonzentrationen sind jeweils in eigenen Knoten repräsentiert, die über einen speziellen Kantentyp (der mit Leerzeichen und eckigen Klammern notiert wird) an "Cell" angebunden sind:

```
Cell [s:GA19] [p:GA20] [f:GA1] ::> competitiveInhibition(s, p, f, Vmax, Km, KI);
```

hierbei ist "competitiveInhibition" eine Java-Methode, die die Numerik einer Reaktion vom Typ "kompetitive Hemmung" auf klassische Weise prozedural implementiert (siehe Buck-Sorlin et al. 2005 für den genauen Code).

Der Transport wird ebenfalls durch einfache "Aktualisierungs-Regeln" implementiert, beispielsweise der basipetale GA_{19} -Transport durch:

```
Cell [a:GA19] -predecessor-> Cell [b:GA19] ::>
{
  double r = c*a.concentration;
  a.concentration -= r*ΔT;
  b.concentration += r*ΔT;
}
```

Hierbei ist c eine Konstante. Die durch einen Doppelpunkt eingeleiteten Zuweisungssymbole $-=$ und $+=$ sind eine Besonderheit von XL: Im Gegensatz zu den entsprechenden Java-Symbolen (ohne den Doppelpunkt) stellen sie eine quasi-parallele Ausführung der entsprechenden Anweisungen sicher, wenn Regeln parallel zur Anwendung kommen. Dies ist eine wichtige Eigenschaft von RGG und steht im Einklang mit der meistens anzutreffenden Parallelität von Prozessen in der Biologie. Die obige Regel ist wie folgt zu interpretieren: In jedem Zeitschritt wird eine bestimmte Menge GA_{19} von jeder Zelle zu ihrer Vorgängerin, die die nächsttiefere Zelle in der Sprossachse ist, weitertransportiert. Die Transportrate r entspricht dabei der Konzentration, multipliziert mit dem konstanten Faktor c . Dieser Transportmechanismus stellt eine starke Vereinfachung dar und wird in späteren Versionen des Modells durch eine realistischere Diffusionsgleichung ersetzt werden.

Die Wachstumsregeln

Der Zweck des GA-Metabolismus auf der makroskaligen Ebene, nämlich die GA_1 -induzierte Internodienstreckung, wird ebenfalls durch eine einzelne Aktualisierungsregel modelliert:

```
i:Internode [s:GA1] ::> i.length += L*s.concentration*ΔT;
```

Diese Regel vergrößert die Länge jedes Internodiums (wobei "Internode" eine Unterklasse von "Cell" ist) um einen Betrag, der proportional zur GA_1 -Konzentration ist. Diese Regel ist Teil eines größeren Regelsatzes zur Gerstenmorphologie, -biometrie und -phänologie, welcher auf früheren Arbeiten des zweiten Autors beruht (Buck-Sorlin & Bachmann 2000, Buck-Sorlin 2002) und hier aus Platzgründen nicht vollständig dargestellt wird. Die zentrale Wachstumsregel sieht in verkürzter Form so aus:

```
m:Meristem ==> Internode [ Leaf ]
  if (Temperatursumme überschreitet Schwellenwert) InflMeristem else m;
```

Hier finden wir die Semantik konventioneller L-Systeme wieder: Bei der Regelanwendung produziert jedes Meristem m ein Internodium, das ein Blatt trägt, und reproduziert sich selbst oder aber ein Blütenmeristem (InflMeristem), falls die Temperatursumme einen bestimmten Schwellenwert überschreitet. Die von dieser Regel erzeugten Internodien unterliegen in den nachfolgenden Schritten dann dem oben beschriebenen, mit dem GA-Metabolismus verknüpften Streckungsmodell.

Vermehrung, Genetik und Züchtungsmodell

Der virtuelle Züchtungsmechanismus des "BarleyBreeder"-Modells wurde durch die "Biomorphe" von Richard Dawkins inspiriert (Dawkins 1996; zur Nachbildung der Original-Biomorphe in XL siehe Knie-meyer et al. 2003). Chromosomen werden durch einen weiteren Knotentyp repräsentiert und tragen Tochterknoten mit ganzzahligen Parametern, die den Allelen bestimmter, ausgewählter Gersten-Gene entsprechen. Wir haben uns für das Modell auf sieben Mendelsche Gene mit zwei oder drei Allelen beschränkt, die Farbe, Form und Größe der Ähren und vegetativer Organe beeinflussen: *cer-ze* (verantwortlich für eine Wachsschicht auf der Epidermis, die eine blaugrüne Färbung aller oberirdischen, vegetativen Pflanzenteile bewirkt), *lks2* (steuert die Ährenlänge), *vrs1* (steuert die Anzahl der Blütenzeilen), *Zeo* (Zwergwuchsgen, das die Längen aller Internodien beeinflusst), *Blp* (beeinflusst

die Färbung der Ähren), *glo-b* (kugelförmiges Korn) und *cul2* (unterdrückt die Bestockung). Rekombinationswahrscheinlichkeiten wurden auf der Grundlage der bekannten Verteilung dieser Gene auf teils unterschiedliche Chromosomen und mittels gemessener Werte festgelegt (siehe Buck-Sorlin et al. 2005). Zur Einbeziehung der Genwirkung in das Modell wird in den meisten Fällen in den Regeln einfach der entsprechende Zahlenwert für das vorliegende Allel abgefragt; d.h. der kausale Pfad über die Zwischenebenen der Genexpression und des Metabolismus wird pragmatisch umgangen. Anders für das Zwergwuchsgen *Zeo*, das direkt mit dem GA-Biosynthesepfad interferiert, indem es die Produktion von GA₁₉ hemmt. In diesem Fall wird der Zahlenwert für das vorliegende Allel in der reaktionskinetischen Gleichung für die GA₁₉-Produktion abgefragt und wirkt somit auf dem Umweg über das metabolische Netzwerk auf die Morphologie der virtuellen Gerste. Die (wahlweise sexuelle oder asexuelle) Vermehrung der Pflanzen wird durch eigene Regeln beschrieben, welche die das Genom beschreibenden Subgraphen zunächst verdoppeln und dabei mögliche (Punkt-) Mutationen und Rekombinationsereignisse (crossing over) berücksichtigen (vgl. Kniemeyer et al. 2003). Beim Start des Modells auf der GroIMP-Software werden dem Benutzer im Ausgabefenster zunächst fünf verschiedene virtuelle Gerstenpflanzen präsentiert, welche die sichtbaren Phänotypen zu fünf unterschiedlichen diploiden Genomen darstellen, deren relevante Allelkombinationen darunter aufgelistet sind. Hieraus können dann per Mausklick ein oder zwei Individuen ausgewählt werden, aus denen das Modell dann die nächste Generation virtueller Gerstenpflanzen generiert (unter Berücksichtigung von Mutationen und, im Falle der sexuellen Vermehrung aus zwei Individuen, von Rekombination). Eine Beispiel-Ausgabe dieses Modells zeigt Abbildung 6. Man beachte, dass insbesondere das Allel "1" für das *Zeo*-Gen tatsächlich zwergwüchsige Individuen erzeugt.

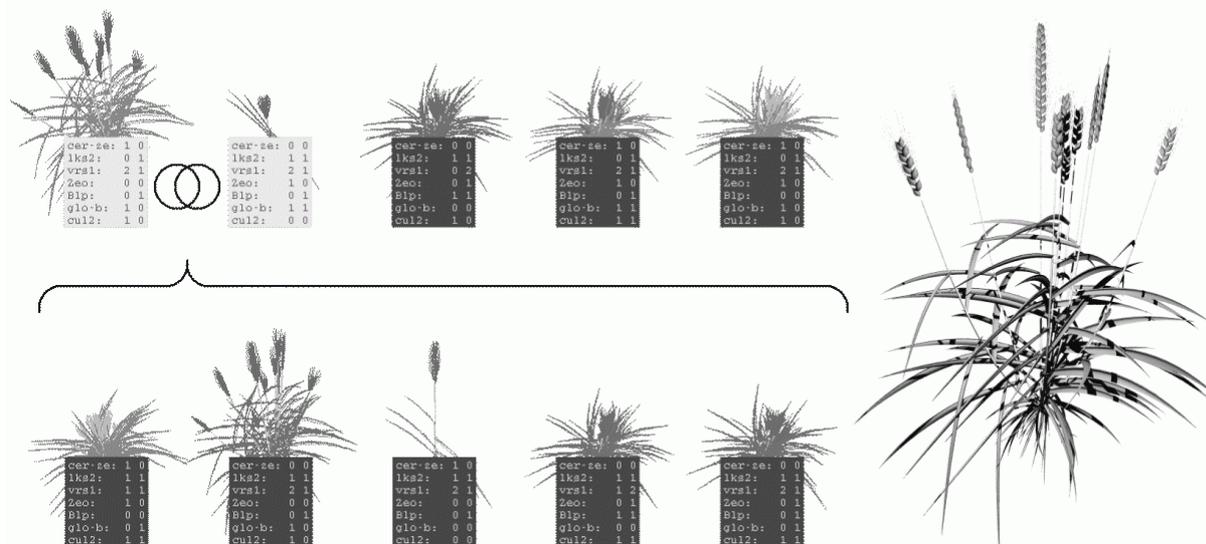


Abbildung 6: Muster-Ausgabe des *BarleyBreeder*-Modells. Oben links: Startpopulation aus fünf virtuellen Gerstenpflanzen, aus denen der Benutzer zwei als Elternindividuen ausgewählt hat; die nächste Generation wird darunter angezeigt. Unter den Pflanzen wird jeweils der Genotyp aufgelistet. Rechts eine mit 3D-Raytracing aufbereitete Ansicht einer Einzelpflanze. (Aus Buck-Sorlin et al. 2005.)

Diskussion

Virtual crops oder *in silico*-Pflanzen haben in den letzten Jahren vermehrte Aufmerksamkeit erfahren. Ihre Verwendung zur Wissensintegration, für Hypothesentests und für verbesserte Prognosen entspricht stark der Denkweise der Systembiologie – letztere ist eigentlich das umfassendere Fach, in das diese Modellansätze hineingehören, obgleich die Aufmerksamkeit der Systembiologie bisher meist auf die molekulare Ebene beschränkt blieb. Wie bereits Minorsky (2003) bemerkt hat, haben solche Modellansätze mittlerweile nicht nur einen Stand erreicht, bei dem sie als ernsthaftes Forschungswerkzeug in den Pflanzenwissenschaften gelten können, sondern sie werden sogar dringend benötigt, um bei der Lösung agronomischer Probleme auf globaler Ebene zu helfen, bei denen rein experimentelle Methoden nicht länger ausreichend sind (z.B. die Notwendigkeit, die Wassernutzungseffizienz von Nutzpflanzen zu erhöhen, oder die Bestimmung der Reaktionen von Pflanzen und Pflanzenbeständen auf erhöhte CO₂-Konzentrationen). Folgen wir weiter Minorsky (2003), so besteht der angemessene Weg, Biologie zukünftig anzuwenden, aus dem ständigen Zusammenwirken pflanzli-

cher Molekularbiologie mit Informatik, vermittelt durch das Arbeiten an virtuellen Pflanzen, um Hypothesen über komplexe Systeme zu prüfen und Versuchsdesign und Messungen zu optimieren.

Das hier kurz vorgestellte Modell ist in der Lage, realistisch aussehende, dreidimensionale oberirdische Pflanzenarchitekturen zu erzeugen, einschließlich derer einiger bekannter Gerstenmutanten. Allerdings ist es nicht durchgehend mit gemessenen Werten der hormonalen Reaktionskinetik parametrisiert (da derartige Daten schwer zu messen sind und für Gerste teilweise nicht vorliegen), und es fehlen im Modell weitere, wichtige Faktoren, die die Morphogenese beeinflussen. Somit muss unser Modell zum gegenwärtigen Zeitpunkt lediglich als ein Prototyp betrachtet werden, der das integrative Potenzial des RGG-Formalismus demonstriert – mit seiner Möglichkeit, metabolische, genetische und morphologische Strukturen im selben Rahmen zu repräsentieren –, kann aber kein komplettes "crop model" mit Prognosefähigkeiten darstellen. Es ist vorgesehen, das Modell schrittweise mit weiteren metabolischen Subsystemen und entsprechenden Genen anzureichern; ein erstes Ziel in unserem gegenwärtigen Projekt ist die Einbeziehung von wichtigen Teilen des Stickstoff-Stoffwechsels, da die Stickstoffnutzungseffizienz offensichtlich ein sehr praxisrelevanter Parameter ist und da wir hoffen, hier mittels der regelbasierten, virtuellen Pflanzen zur Wissensintegration und zum Verständnis des komplexen Systems "Pflanze" beitragen zu können.

Modelle dieser Art könnten in der Zukunft als "Werkzeuge für Pflanzenzüchter" verwendet werden: Als Input bräuchte man die Rekombinations-Distanzen einer beliebigen Menge von Marker-Genen und von Genorten quantitativer Merkmale (*quantitative trait loci*: QTL). Der Züchter würde dann die Größe der Ausgabe-Population und – nebst weiteren Randbedingungen – einen "Ideotyp" festlegen, also einen idealen Phänotyp, auf den durch den Züchtungsprozess hingearbeitet werden soll. Im Idealfall könnte dann das Modell die zu verfolgenden "Züchtungspfade" berechnen, sowie die erforderliche Zahl von Schritten, die man braucht, um das Ziel zu erreichen – und auch die möglichen Schwierigkeiten oder "Flaschenhals"-Situationen, zum Beispiel in dem Fall, dass Zielgene zu dicht an unerwünschten Genen liegen.

Wir sehen relationale Wachstumsgrammatiken als ein zukünftiges Modellierungs- und Forschungswerkzeug in den Händen von Genetikern, Physiologen, vergleichenden Morphologen und Agrarwissenschaftlern.

Danksagung

Wir danken für die Förderung durch die DFG, teilweise im Rahmen der Forschergruppe "Virtual Crops", unter den Projektkennzeichen Ku 847/5 und Ku 847/6-1. Der zweite Autor dankt dem Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK) Gatersleben, insbesondere Dr. Patrick Schweizer, für das Gastrecht.

Literatur

Breckling, B.: An individual based model for the study of pattern and process in plant ecology. An application of object oriented programming. *EcoSys*, 4 (1996), 241-254.

Breckling, B.; Müller, F.; Reuter, H.; Hölker, F.; Fränzle, O.: Emergent properties in individual-based ecological models – introducing case studies in an ecosystem research context. *Ecological Modelling*, 186 (2005), 376-388.

Buck-Sorlin, G. H.: L-system model of the vegetative growth of winter barley (*Hordeum vulgare* L.). In: Polani, D.; Kim, J.; Martinetz, T. (eds.): *Fifth German Workshop on Artificial Life. Abstracting and Synthesizing the Principles of Living Systems*. March 2002, Lübeck. AKA Akadem. Verlagsges., Berlin (2002), S. 53-64.

Buck-Sorlin, G. H.; Bachmann, K.: Simulating the morphology of barley spike phenotypes using genotype information. *Agronomie: Plant Genetics and Breeding*, 20 (2000), 691-702.

Buck-Sorlin, G. H.; Kniemeyer, O.; Kurth, W.: Barley morphology, genetics and hormonal regulation of internode elongation modelled by a relational growth grammar. *New Phytologist*, 166 (2005), 859-867.

Dawkins, R.: *Der blinde Uhrmacher*. dtv, München (1996).

de Reffye, Ph.; Snoeck, J.: Modèle mathématique de base pour l'étude et la simulation de la croissance et de l'architecture du *Coffea robusta*. *Café Cacao Thé*, 20 (1976), 11-32.

de Reffye, Ph.; Houllier, F.; Blaise, F.; Barthélémy, D.; Dauzat, J.; Auclair, D.: A model simulating above- and below-ground tree architecture with agroforestry applications. *Agroforestry Systems*, 30 (1995), 175-197.

Godin, C.; Hanan, J. S.; Kurth, W.; Lacoïnte, A.; Takenaka, A.; Prusinkiewicz, P.; DeJong, T.; Beveridge, C.; Andrieu, B. (eds.): 4th International Workshop on Functional-Structural Plant Models. Abstract Volume. UMR AMAP, Montpellier (2004).

Hanan, J.: Virtual plants – integrating architectural and physiological models. *Environmental Modelling and Software*, 12 (1997), 35-42.

Karwowski, R.; Prusinkiewicz, P.: Design and implementation of the L+C modeling language. *Electronic Notes in Theoretical Computer Science*, 86 (2), 19 S.

Karwowski, R.; Prusinkiewicz, P.: The L-system-based plant-modeling environment L-Studio 4.0. In: Godin et al. (2004), 403-405.

Kniemeyer, O.: Rule-based modelling with the XL/GroIMP software. In: Schaub, H.; Detje, F.; Brüggemann, U. (eds.): *The Logic of Artificial Life. Proceedings of 6th GWAL, Bamberg, April 2004.* AKA Akadem. Verlagsges., Berlin (2004), S. 56-65.

Kniemeyer, O.; Buck-Sorlin, G. H.; Kurth, W.: Representation of genotype and phenotype in a coherent framework based on extended L-systems. *Lecture Notes in Artificial Intelligence*, 2801 (2003), 625-634.

Kniemeyer, O.; Buck-Sorlin, G. H.; Kurth, W.: A graph-grammar approach to artificial life. *Artificial Life*, 10 (2004), 413-431.

Kurth, W.: Growth Grammar Interpreter GROGRA 2.4: A software tool for the 3-dimensional interpretation of stochastic, sensitive growth grammars in the context of plant modelling. Introduction and Reference Manual. *Berichte des Forschungszentrums Waldökosysteme der Universität Göttingen, Ser. B, Band 38* (1994).

Kurth, W.: Spatial structure, sensitivity and communication in rule-based models. In: Hölker, F. (ed.): *Scales, Hierarchies and Emergent Properties in Ecological Models.* Reihe "Theorie in der Ökologie", Bd. 6. Peter Lang, Frankfurt a. M. (2002), 29-46.

Lindenmayer, A.: Mathematical models for cellular interactions in development. *Journal of Theoretical Biology*, 18 (1968), 280-315.

Minorsky, P. V.: Achieving the *in silico* plant. *Systems biology and the future of plant biological research.* *Plant Physiology*, 132 (2003), 404-409.

Olszewki, N.; Sun, T. P.; Gubler, F.: Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell*, 14 (2002), S61-S80.

Perttunen, J.; Sievänen, R.; Nikinmaa, E.; Salminen, H.; Saarenmaa, H.; Väkevä, J.: LIGNUM: A tree model based on simple structural units. *Annals of Botany*, 77 (1996), 87-98.

Prusinkiewicz, P.: Art and science for life: designing and growing virtual plants with L-systems. In: Davidson, C.; Fernandes, T. (eds.): *Nursery Crops: Development, Evaluation, Production and Use.* *Acta Horticulturae*, 630 (2004), 15-28.

Prusinkiewicz, P.; Lindenmayer, A.: *The Algorithmic Beauty of Plants.* Springer, New York (1990).

Streit, Ch.: *Graphical User Manual.* Universität Bern, Bern (1992).

von Caemmerer, S.; Farquhar, G. D.: Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. *Planta*, 153 (1981), 376-387.

Wernecke, P.; Buck-Sorlin, G.; Diepenbrock, W.: Virtual canopies: Combining process- with architectural models: the simulation tool VICA. *Systems Analysis Modelling Simulation (SAMS)*, 39 (2000), 235-277.

Zhao, D.: Simulation und Visualisierung der Struktur und Dynamik metabolischer Netzwerke mit relationalen Wachstumsgrammatiken. Diplomarbeit, Brandenburgische Technische Universität Cottbus, Lehrstuhl Grafische Systeme (2006).