

Diss. ETH No. 21173

**Self-Construction in the Context of Cortical Growth:
From One Cell to a Cortex to a Programming
Paradigm for Self-Constructing Systems**

A dissertation submitted to

ETH Zurich

for the degree of
Doctor of Sciences
presented by

Andreas Hauri

Master of Science in Computer Science ETHZ

born 30 September 1981

citizen of
Hirschthal, Switzerland

accepted on the recommendation of

Prof. Dr. Rodney J. Douglas, examiner
Prof. Dr. Christoph von der Malsburg, co-examiner
Prof. Dr. David Willshaw, co-examiner

2013

Abstract

In this thesis we explore the biological development of the neocortex as a model for self-constructing systems. Development begins with a single cell, and expands that cell through cell division and specialization to create a complex organism consisting of trillions (in humans) of cells. We will consider the branch of this process that leads to the neocortex, beginning with precursors in the ventricular zone, and the ganglionic eminences. In biology the instructions for construction of a target organism are encoded in its DNA-code, and these instructions are selectively decoded by a gene regulatory network according to prevailing intra-cellular and extracellular conditions. We approximate this elaborate process with a model that contains the essential elements of biological self-construction, but is sufficiently tractable to be simulated. Because our final goal is to exploit biological construction as a practical engineering technology, we choose a style of simulation that emphasizes physical process in a 3D spatial environment. And, because the compelling feature of biological development is that it is distributed and proceeds without a global controller, our model emphasizes self-construction as the result of locally interacting distributed autonomous agents. The model is simulated using a specialized software platform, Cx3Dp, that is able to simulate the physical development of millions of cells. The biological behavior of these cells is controlled by G-Code, a DNA-like specification language that is able to model cellular behavior as observed in nature. The organization of the G-Code ‘genome’ that is inserted into the cortical precursor cells of the embryonic ventricular zone is obtained by applying the methods of Pfister et al. [146], by which sparse experimental data are used to estimate a model genome able to express a gene-regulatory network (GRN) that controls mouse corticogenesis. The various cell types of the cortical lineage tree are expressions of the attractor states of the GRN. The various states release G-Code encoded cellular functions that cause the cells to divide, migrate, and differentiate. Using these concepts and methods we are able to simulate the self-construction of a sheet of neocortex composed of two areas containing some 200,000 neurons. The simulation recapitulates the majority of experimentally observed features of this process, including the detailed inter- and intra-laminar axonal connection patterns. The simulation platform is parallelized, and its performance scales well with simulation size, so that

the simulation of much larger volumes of cortex are possible. These explorations of cortical development lead us to a new paradigm of engineered self-construction: the Developmental Programming paradigm. At the center of this paradigm is the concept of exponential construction, whereby a single precursor containing a single genome amplifies itself by successive replication and specialization, and so is able to construct a vastly more elaborate product than can be achieved with conventional, more linear fabrication methods. There are three crucial components in the construction process: firstly, the genome that encodes what functions are possible, and provides rules for their expression; secondly the machine-like agents that are instantiations of the genomic functions; and thirdly the container (e.g. cell membrane) that provides the spatial and temporal scope of genome and agent activities. The contents of containers work independently of one another, but can be coupled through local messages. This is a highly parallel and distributed process, in which the computational power of the overall system increases with each container replication. In this thesis we have considered this self-construction process as it is embodied in cortical development. However, it is easy to see how these principles could be extended to novel self-replicating technologies as diverse as genetically engineered cells, nano-robots, DNA-computing systems, and self-configuring software for exascale computers.

Zusammenfassung

In dieser Doktorarbeit untersuchen wir die biologische Entwicklung des Neocortex als Modell für selbstkonstruierende Systeme. Die Entwicklung beginnt mit einer einzigen Zelle, und erweitert diese Zelle durch Zellteilung und Spezialisierung in einen komplexen Organismus, bestehend aus Billionen von Zellen (beim Menschen). Wir betrachten den Teil dieses Prozesses welcher zum Neocortex führt, beginnend mit Präkursor-Zellen in der ventrikulären Zone und den Ganglionic-Eminences. In der Biologie sind die Befehle zum Aufbau eines Zielorganismus in seiner DNS codiert. Diese Befehle werden selektiv durch ein Genregulationsnetzwerk decodiert, je nach den Bedingungen die intrazellulär und extrazellulär vorherrschen. Wir approximieren diesen aufwendigen Prozess mit einem Modell, welches die wesentlichen Elemente des biologischen Selbstkonstruktion enthält, jedoch steuerbar genug ist um es zu simulieren. Da es unser Endziel ist, die biologische Bauweise als praktische Ingenieurtechnik zu nutzen, wählen wir eine Art von Simulation, die den physikalischen Prozess in einer 3 dimensionalen Umgebung betont. Und da die faszinierendste Funktion der biologischen Entwicklung die verteilte Funktionsweise ist, welche ohne globale Steuerung auskommt, betont unser Modell, dass die Selbst-Konstruktion das Ergebnis von lokal interagierenden, verteilten und autonomen Agenten ist. Das Modell wird mit Hilfe der speziellen Software-Plattform, Cx3Dp simuliert, welche in der Lage ist die physikalische Entwicklung von Millionen von Zellen zu simulieren. Das biologische Verhalten dieser Zellen wird durch G-Code gesteuert, eine DNS-ähnliche Beschreibungssprache, die in Lage ist zelluläres Verhalten wie es in der Natur beobachtet wird zu modellieren. Die Organisation des G-Code „Genoms“ wird durch die Anwendung der Methoden von Pfister et al. [146] erlangt und in die kortikalen Präkursor-Zellen der embryonalen ventrikulären Zone eingefügt. Bei dieser Methode wird mittels wenigen experimentellen Daten ein Modell Genom bestimmt, welches ein Genregulationsnetzwerk (GRN) ausdrückt, das die Kortikogenese der Maus steuert. Die verschiedenen Zelltypen der kortikalen Entwicklungslinie sind der Ausdruck von den Attraktor Zuständen des GRN. Die verschiedenen Zustände starten in G-Code kodierte zelluläre Funktionen, welche die Zellen teilen, migrieren und differenzieren lassen. Unter Verwendung dieser Konzepte und Methoden können wir die Selbst-Konstruktion einer neocortikalen Platte bestehend aus zwei Zonen

mit ca. 200'000 Neuronen simulieren. Die Simulationsergebnisse rekapitulieren die Mehrheit der experimentell beobachteten Merkmale dieses Prozesses, einschliesslich der detaillierten inter- und intra- laminaren axonalen Verbindungsmuster. Die Simulationsplattform ist parallelisiert und Ihre Leistungsfähigkeit skaliert gut mit der Simulationsgrösse, so dass gar Simulationen von grösseren kortikalen Volumen möglich sind. Diese Erkundungen der kortikalen Entwicklung führen uns zu einem neuen Paradigma der technischen Selbst-Konstruktion: Das Developmental-Programming-Paradigma. Im Zentrum dieses Paradigmas ist das Konzept von exponentieller Konstruktion, bei welchem ein einziger Präkursorbehälter mit einem einzigen Genom sich selbst Amplifiziert durch Replikation und Spezialisierung, und fähig ist ein weitaus elaborierteres Produkt zu konstruieren als es mit konventionellen linearen Fabrikationsmethoden möglich ist. Hier sind drei wichtige Komponenten im Konstruktionsprozess: Erstens, das Genom, welches die möglichen Funktionen kodiert und Regeln für ihren Ausdruck enthält. Zweitens, die maschinenähnlichen Agenten welche die Instanziierung der genomisch enkodierten Funktionen sind. Und drittens, der Behälter (z.B. Zellmembrane) welche das örtlichen und zeitlichen Ausmass des Genoms und der Agenten Aktivitäten bestimmt. Die Inhalte des Behälters arbeiten unabhängig voneinander, können aber voneinander abhängig gemacht werden mittels Nachrichten. Dies ist ein hoch parallelisierter und verteilter Prozess, welcher die Rechenleistung des Gesamtsystems mit jeder Replikation eines Behälters erhöht. In dieser Doktorarbeit haben wir den Prozess der Selbstkonstruktion, wie er in der kortikalen Entwicklung verkörpert ist, untersucht. Allerdings ist es einfach zu sehen wie diese Prinzipien auf neue selbstreplizierende Technologien erweitert werden könnten, Technologien wie künstliche Zellen, Nanoroboter, DNS-Rechnersysteme und selbstkonfigurierende Software für exascale Computer.