

Herren Professoren

Norbert Elsner
Werner Lehfeldt
Gerd Lüer
Robert Schaback
Kurt Schönhammer

27. Oktober 2005

Liebe Kollegen,

angeblich leben wir im Zeitalter der Information; auf alle Fälle aber ist "Information" nicht nur in aller Munde, sondern auch von großer Bedeutung (wenn nicht gar im Zentrum des Interesses) für nicht wenige wissenschaftliche Disziplinen.

In Kontrast dazu ist der Begriff seltsam schillernd – was den umgangssprachlichen Gebrauch angeht, scheint kaum jemand eine feste Vorstellung davon zu haben, was mit "Information" gemeint sein mag. Die verschiedenen wissenschaftlichen Disziplinen mögen da – jede für sich – konkreter sein, aber sich nicht unbedingt alle auf dasselbe zu beziehen.

Aus dieser Feststellung leitet sich die Frage ab, ob es fruchtbar sein könnte, in ein Gespräch einzutreten, das den Informationsbegriff von verschiedenen Seiten beleuchtet. Obwohl dies sicher schon oft unternommen wurde, habe ich – ohne mir vor gemachtem Versuch sicher zu sein – das Gefühl, zwischen den Disziplinen könne noch einiges zu entdecken sein, was das Potential hat, förderlich auf die verschiedenen Spezialinteressen zurückzustrahlen.

Ich will daher den Versuch machen, Sie als Vertreter von Physik, Informatik, Psychologie, Neurobiologie und Sprachwissenschaften zu einem ersten Gedankenaustausch über das skizzierte Thema zusammenzubringen – nur um herauszubekommen, ob wir genug Substanz in der Sache sehen, um zuversichtlich in eine systematischere Diskussion einzutreten.

Als eigene Vorbereitung habe ich *ad hoc* einige Gedankensplitter zusammengetragen, die sich einem als Molekularbiologen zu dem Thema als erste aufdrängen (siehe unten; Herr Schönhammer und Herr Elsner kennen eine frühere Version dieses Textes bereits).

Als Termin schlage ich den "akademiefreien" Freitagnachmittag (11.11.05) um 17:15h vor. Den Ort des Treffens werde ich noch bekannt geben – mir erschiene die Akademiebibliothek ideal (ob das möglich ist, werde ich aber noch zu eruieren haben). Bitte teilen Sie mir möglichst bald mit, ob Sie an dem Treffen teilnehmen können.

Viele freundliche Grüße

Ihr
Hans-Joachim Fritz

Information und Informationsverarbeitung in der Biologie, eine Stichwortsammlung erster Näherung

1. DNA

- Grundprobleme: Kopiergenauigkeit und Langzeitstabilität – trotz stochastischer Störprozesse. (Weniger als 10^{-10} Fehler pro Monomer und Zellzyklus).
- Redundanz: Minimal und anfällig (eine einzige "Sicherungskopie" mit lediglich ein paar wenigen Atomen pro bit). Dabei kann im Extremfall ein einziger "Bitumfall" (Mutation) tödlich sein (bei wenigen 10^6 bis einigen 10^9 Informationspaketen zu je zwei bit – je nach Organismus).
- Informationserhalt (Korrekturlesen und Reparatur) erzeugt Energiekosten. Dies führt auf das grundsätzliche Problem der Thermodynamik der Information (Maxwell, Szilard, Shannon, Wiener, Landauer und andere).
- Thermophile Organismen (bis 130 °C !) haben mit stark erhöhten Raten chemischer Reaktionen der DNA zu kämpfen, die geeignet sind, die genetische Information zu korrumpieren. Die prämutagene hydrolytische Desaminierung von DNA-Cytosinresten zu Uracil, zum Beispiel, ist bei 105 °C ca. 10^4 mal schneller als bei 37 °C (und schon bei dieser Temperatur ist dieses Problem groß genug, dass es mit redundant angelegten DNA-Reparatursystemen bekämpft werden muss).
- Fragen, die sich an den letzten Punkt anschließen: **(i)** Wie schaffen die Thermophilen Organismen das? **(ii)** Ist der für DNA-Reparatur zu treibende Energieaufwand vielleicht limitierend für die Genomgröße extrem thermophiler Organismen? (Einiges scheint dafür zu sprechen; alternative Erklärungen für die durchweg auffällig niedrigen Genomgrößen sind jedoch nicht ausgeschlossen). **(iii)** Macht das die Entstehung des Lebens in einem heißen Ambiente weniger wahrscheinlich? (Wegen des dann doppelten Handicaps bei noch nicht evolutiv verfeinertem, mithin ungenauem Replikationsapparat).

2. PROTEINBIOSYNTHESE

Die Bezeichnung des Ribosoms als eine numerisch gesteuerte Werkzeugmaschine ist keine Metapher, sondern wörtlich zu nehmen. Es gibt:

- Geordnete, energiegetriebene mechanische Bewegungen.
- Programm-gesteuerten Werkzeugwechsel.
- Klare Trennung von Software und Hardware: Die Maschineninstruktionen werden als lineare und frei wählbare (innerhalb des Standard-Alphabets auch *experimentell* frei wählbare!) Kette von Zeichen auf einem bandförmigen, materiellen Informationsträger eingefüttert. (Pakete zu je acht bit sind an jeweils

einem einzigen, präzise zu bestimmenden Ort auf dem Instruktionsband niedergelegt).

- Einen festen Maschinentakt, durch den das Vollenden eines monoton zu wiederholenden Arbeitszyklus mit dem Vorrücken des Informationsträgers am "Lesekopf" verknüpft ist.

Die oben aufgeführten Punkte sind ein alter Hut; sie in dieser Form noch einmal zusammenzustellen ist aber nützlich, um den davon ganz verschiedenen Sachverhalt bei anderen informationsverarbeitenden Systemen der Biologie zu beleuchten (vgl.u.).

3. GEHIRN

Einige Unterschiede zu Speicherung und Verarbeitung von Information im makromolekularen Stoffwechsel von DNA, RNA und Proteinen:

- Keine Software erkennbar. Statt dessen alles in Hardware angelegt – aber nicht *fest* verdrahtet, sondern plastisch.
- Das Bewahrungsproblem stellt sich hier nicht mit der selben Stringenz wie bei der in DNA angelegten, genetischen Information – dafür aber die wesentlich komplexeren Probleme des *(i)* Erzeugens von Verhalten mit sehr kurzer Reaktionszeit auf einen äußeren Reiz, *(ii)* des Erinnerns, *(iii)* des Lernens, *(iv)* des Reflektierens, *(v)* der Kreativität, *(vi)* des Planens, *(vii)* des Träumens und *(viii)* des (selektiven) Vergessens (und vieles andere mehr – von schwierigerem wie Ich und Wille ganz zu schweigen).
- Replikation, Reparatur und Genexpression betreibt jede Zelle alleine; vom Rest des Organismus und von der Außenwelt wird der Genexpressionsapparat lediglich regulatorisch – und nur in speziellen Teilen – beeinflusst (durch Hormone, Pheromone, Licht, Wärme, vorhandene Nährstoffe und anderes – in Spezialfällen indirekt auch durch neuronale Stimuli). Im Gegensatz dazu sind die im vorigen Punkt aufgezählten Erscheinungen kollektive Leistungen sehr vieler Zellen, das heißt des gesamten Organs Gehirn – oder mindestens großer Teile davon.
- Die genetische Information unterliegt während des Lebenszyklus eines individuellen Lebewesens keiner oder nur sehr geringfügiger und immer zufälliger Veränderung. (Bei somatischen Zellen kann das in besonders interessanten Fällen anders sein. Diese aber sind Einmalartikel und werden deshalb hier nicht als dem **Lebenszyklus** zugehörig betrachtet).
- Die genetische Information ist nicht nur starr (voriger Punkt), ihr Fluss ist zudem unidirektional: Es gibt keinen Mechanismus, individuell gemachte Erfahrungen in DNA "zurückzuschreiben".
- Andererseits überdauert genetische Information das individuelle Leben (Reproduktion vorausgesetzt), während die in einem Gehirn angesammelte

Information mit dem Tod in dem Maße untergeht, wenn sie zuvor nicht per Kommunikation in andere Gehirne transportiert oder auf technischen Datenträgern (Marmor, Papier, Silizium *etc.*) in abrufbarer Form niedergeschrieben wurde.

Fragen:

- Wie wird Information in ein Netzwerk eingespeist, dort gespeichert und von dort abgerufen?
- Wodurch wird der angesammelte Informationsbestand gefährdet und wie wird er gepflegt, d.h. über welche Mechanismen greifen Störprozesse an einem solchen System an und welche Bewahrungswerkzeuge gibt es. (Energiedissipation muss man aus fundamentalen Gründen voraussetzen. Gibt es so etwas wie redundante Anlage in einem Netzwerk? Wenn nicht: Was sonst erzeugt Robustheit? Hat der Verfall der Information wegen ihrer distributiven Anlage eine ganz andere "mechanistische" Qualität als bei der "Ein-bit-ein-Ort"-Lösung? Welche?)
- Wie könnte ein aus technischen Bauelementen konstruiertes neuronales Netz, das heißt eine sich adaptiv selbst umorganisierende, lernfähige Hardware, aussehen? Gibt es das vielleicht schon (oder ist zumindest ein prinzipiell gangbarer Weg dahin plausibel gemacht worden)? (Damit ist mehr gemeint, als einen Rechner zu bauen, bei dem jedes Neuron zwar durch einen eigenen Prozessor repräsentiert ist, bei dem diese Prozessoren aber nach einem festen Schema verschaltet und nur die für jedes Neuron jeweils anzuwendenden *input/output*-Regeln individuell variabel gehalten sind).
- Lehren uns bereits heutige (heißt: in Software auf einer klassischen Von-Neumann-Maschine simulierte) neuronale Netze irgendetwas Biologie-relevantes?
- Gibt es oder (wann) kommt die ("Bio"-)Neuron/Silizium-Kopplung?

4. ENTWICKLUNG

Drei miteinander verwobene Einflüsse, die das Schicksal einer Zelle während der Entwicklung eines Vielzellers bestimmen:

- Eigene Vorgeschichte (bis zurück zum Ei und darüber hinaus bis in stammesgeschichtliche Urzeit).
- Einbettung in Verbund benachbarter Zellen (formale Ähnlichkeit mit Zellulären Automaten).
- Subtil abgestimmte Genregulationskaskaden (Zitat Peter Gruss: "Das Konzert der Gene").

5. VERSCHIEDENES

- Immunsystem (Entstehung von Information durch einen enzymatisch getriebenen, darwinischen Anpassungsprozess auf einer Zeitskala von Tagen und Wochen; das Gedächtnis des Immunsystems; Netzwerk[?])
- Genregulation als Antwort auf Signale von außen, einschließlich Signaltransduktion
- "Systembiologie" (Stoffwechsel ist ein Netzwerk von Reaktionen, dabei insgesamt und in seinen Teilreaktionen abseits des Gleichgewichts; im Netz gibt es Rückkopplungen verschiedenen Vorzeichens und verschiedener Stärke, sowie Vorwärtsverstärkungen verschiedenen Ausmaßes. Von pathologischen Ausnahmesituationen abgesehen, wird dennoch Homöostase erreicht; d.h. es treten weder wilde Oszillationen noch Chaos auf.). Gibt es eine Chance, das ganze System je mit hinreichender Genauigkeit zu parametrisieren, um sinnvolle Modellrechnungen des Globalverhaltens anstellen zu können (so der Anspruch der Systembiologie)? Was sagen Regelungstechniker dazu?
- "Proteinfaltungscode"
- Molekulare Rekonstruktion phylogenetischer Verwandtschaftsbeziehungen aus Sequenzdaten und Parallelen zu den Sprachwissenschaften.
- Informationsverarbeitung durch einen Sozialverbund.
- Theorie und Praxis Mensch-gemachter, selbstreplizierender Objekte/Systeme (Moleküle, Maschinen).
- Rechnen mit DNA.

h.-j.f.

Begonnen: 05_10_06

Version: 05_10_27