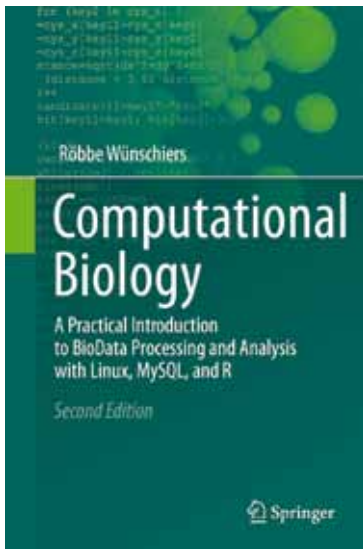


Interview zum Thema: Die Basis der „Computational Biology“ als Buch



Röbbe Wünschiers: *Computational Biology – A Practical Introduction to BioData Processing and Analysis with Linux, MySQL, and R*; 449 Seiten, 80 Illustrationen, 66 Illustrationen in Farbe; Springer-Verlag Berlin Heidelberg; 2. Auflage 2013, ISBN 978-3-642-34749-8; Hardcover 74,85 Euro; e-Book 59,49 Euro.

Der Nobelpreis für Chemie ging in diesem Jahr an Forscher, die sich dem Gebiet der Computerchemie verschrieben haben. Ihre Arbeit betraf bereits Biomoleküle, und sicherlich profitiert auch die Biologie von der rasanten Entwicklung der Computertechnik.

Röbbe Wünschiers hat sein Buch „Computational Biology“, das sich den da verwendeten Datenstrukturen und Programmierumgebungen widmet, nun in der 2. Auflage veröffentlicht. Aus der langen Beziehung des Autors zur CLB – Röbbe Wünschiers schrieb seinen 1. Artikel hier im Jahre 1997 – habe ich ihn als „besten Kenner“ des Buchs um Auskunft darüber gebeten. Ein Teilbereich des Buchs ist den CLB-Lesern übrigens nicht unbekannt: Röbbe Wünschiers veröffentlichte vom November 2003 bis Januar 2005 eine 15-teilige Serie über Unix und verwandte Betriebssysteme in der CLB und ging auch auf die Programmiersprache Awk ein, wie auch in seinem jetzigen Buch. Das führt mich geradewegs zu meiner 1. Frage:

Warum haben Sie seinerzeit das Buch geschrieben? Gab es nicht schon genug Veröffentlichungen zu dem Thema?



Prof. RÖbbe Wünschiers: Natürlich gab es schon beim Erscheinen der 1. Auflage 2004 Bücherberge über Linux und Programmierung – aber nur sehr wenige für den Bereich der Lebenswissenschaften. Mein Ziel

war es schon damals, experimentierenden Wissenschaftlern ein Werkzeug in die Hand zu geben, um mit umfangreichen Datenmengen arbeiten zu können. So habe ich damals selber gearbeitet (experimental & computational biology) und die Methodik habe ich auch in Köln, Bingen und Marburg gelehrt. Der Bedarf war schon damals groß, da mehr und mehr Experimente im Hochdurchsatz durchgeführt wurden und dabei riesige Datenmengen anfielen die mit Excel allein nicht mehr zu bewältigen waren. Das Buch spiegelt in gewisser Weise meinen eigenen Weg wider, wie ich lernte mit diesen Datenmengen zu arbeiten und zu spielen um daraus Wissen zu generieren. Die Programmiersprache AWK erwies sich

dabei als ein Glücksgriff, da sie extrem einfach und klar strukturiert ist und gleichzeitig speziell für die Verarbeitung von textbasierten Inhalten entwickelt wurde. Lebenswissenschaftler arbeiten überwiegend mit Texten: Nukleinsäure- oder Proteinsequenzen; Zahlenkolonnen die Expressionswerte von Proteinen, Verteilungen von Genen in einer Population oder Koordinaten von Atomen in einem Protein angeben; usw.

Was macht die neue, 2. Ausgabe aus? Über den Daumen: Zu welchem Prozentsatz unterscheidet sie sich von der ersten Ausgabe?

Die zweite Auflage ist erheblich überarbeitet und erweitert worden – rund 45% sind neu. Dabei profitiert sie von meinen zwischenzeitlich gesammelten Erfahrungen aus der eigenen Forschung und Lehre, aber auch von meiner Zeit bei der BASF Plant Science als Leiter mehrerer bioinformatischer Großprojekte. Ich war selbst erstaunt, wie vielseitig ich mein Wissen über AWK & Co produktiv einsetzen konnte um schnell mal etwas auszuprobieren. Konzernweite Lösungen wurde dann von Profis programmiert. Ich habe aber auch gelernt, wie wichtig Datenverwaltung und Datenvisualisierung sind. Neu sind daher Kapitel über das Datenbanksystem MySQL und die Datenanalyse und -visualisierungsplattform R. Neu sind auch komplett durchgearbeitete Beispiele die zeigen, wie man mit Linux-Bordmitteln und der Programmiersprache AWK Daten prozessiert, analysiert und visualisiert. Bspw. zeige ich, wie man am Computer herausfinden kann, was EHEC so gefährlich macht. Die neue Auflage profitiert aber auch von den immensen Entwicklung in der Softwarebranche. Heutzutage ist es ein Leichtes eine virtuelle Linux-Maschine auf jedem gängigen Betriebssystem aufzusetzen. Auch das zeige ich – und alles mit frei zugänglicher, kostenloser Software.

In wieweit wird in dem Buch speziell auf Belange von Biologen eingegangen? Gibt es Prozeduren, Algorithmen oder Herangehensweisen informatischer Art, die in der Biologie häufiger auftreten als anderswo und dementsprechend behandelt werden müssen?

Es ist ein Buch von einem Biologen für alle, die in den Lebenswissenschaften arbeiten. Das zeigt sich vor allem in meinen Beispielen. Das Buch ist darauf ausgerichtet zu zeigen, wie man mit großen Datenmengen umgeht und aus ihnen Wissen schöpfen kann. Es geht mir nicht darum zu zeigen, wie neueste bioinformatische Algorithmen umgesetzt werden. Das Buch lehrt Handwerkzeug und ist sehr anwendungsbezogen. Es macht aus dem Leser keinen Bioinformatiker sondern zeigt Wege auf, wie man mit experimentell gewonnen Daten am Computer weiter experimentieren kann.

Ist ein Buch – selbst wenn es wie dieses nun auch als e-Book erhältlich ist – nicht heute eine

überkommene Veröffentlichungsform? Gerade Programmierkenntnisse könnten sich doch interaktiv vermitteln lassen.

Nein, eBook hin, interaktiv her – zu allererst geht es um das Verstehen und Erlernen von Methoden. Den Umgang mit Flüssigkeiten im Hekto- oder Mikrolitermaßstab lerne ich durch Übungen und Wiederholung. Das Hantieren mit Daten am Computer lerne ich auch durch Übungen und Wiederholung. Die Plattform, über welche die Ideen und Inhalte geliefert werden sind letztlich vermutlich egal – oder folgen individuellen Vorlieben. Wichtig ist, dass ich die Dinge übe, anwende, Fehler mache, mich korrigiere. Die Plattform kann ein Film, ein Buch, ein eBook, eine

App oder eine andere interaktive Umgebung sein – sie sollte nur nicht darüber hinweg täuschen, dass kein Weg an der Praxis vorbei führt. Sehr hilfreich finde ich allerdings die Begleitung eines digitalen oder hölzernen Buches über interaktive Medien wie das Web (<http://goo.gl/xGjME>) oder Facebook (<http://goo.gl/Ye7jP>).

Wird es also eine 3. Auflage des Buchs geben?

Das ist gut möglich. Neben weiteren schönen Anwendungsbeispielen würde ich vermutlich Python als Programmiersprache neben AWK vorstellen. Python verdrängt mehr und mehr Perl in der Bioinformatik und bietet spannende Möglichkeiten.

Interview: Rolf Kickuth

Organische Chemie für Fortgeschrittene in äußerst umfangreicher Darstellung

March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure; vollständig überarbeitet und aktualisiert von Michael B. Smit; 2080 Seiten; Verlag John Wiley & Sons; 7. Auflage 2013; ISBN 978-0-470-46259-1; Hardcover; ca. 109,00 Euro.

Das Erscheinen der nunmehr 7. Auflage des erfolgreichen Lehrbuchs March's *Advanced Organic Chemistry* macht die große Akzeptanz dieses Standardwerkes für das Studium der Organischen Chemie auf der Graduierten-Ebene deutlich. Die aus zwei Teilen mit insgesamt 19 Kapiteln bestehende Neuauflage ist nach dem System: Reaktionen, Mechanismen und Struktur aufgebaut. Hinzu kommen eingangs ein Verzeichnis gebräuchlicher Abkürzungen und am Schluss des Buches der Anhang A (36 Seiten) über die Literatur der Organischen Chemie, einschließlich Anleitungen zur Durchführung von Literatur-Recherchen; der Anhang B (25 Seiten) zur Klassifikation von Reaktionen aufgrund der Art der synthetisierten Verbindungen sowie jeweils ein im Vergleich mit anderen Lehrbüchern im Hinblick auf Ausführlichkeit hervorzuhebendes Autoren-Verzeichnis und Sachverzeichnis (zusammen 417 Seiten).

Von den 9 Kapiteln des Teils I (insgesamt 366 Seiten) beschreiben die ersten 5 Kapitel *Strukturen* organischer Verbindungen, beginnend mit den hier vorliegenden Arten der Chemischen Bindung (Kap. 1-3). Als umfangreichstes Kapitel des Teils I beschäftigt sich Kap.4 (86 Seiten) mit der Stereochemie und den Konformationen organischer Verbindungen. Kapitel 5 befasst sich mit den bei organisch-chemischen Reaktionen auftretenden reaktiven Zwischenprodukten: Carbenium-Ionen, Carbanionen, Freien Radikalen, Carbenen und Nitrenen. Die folgenden Kapitel umfassen: Mechanismen organisch-chemischer Reaktionen und Methoden zu deren Bestimmung (6), Photochemie, Sonochemie und Einsatz von Mikrowellen in der Organischen Chemie (7), Säuren und Basen (8) und die Abhängigkeit

der Reaktivität von der Struktur und dem jeweiligen Medium (9).

Teil II (insgesamt 1202 Seiten) beschreibt die in der Organischen Chemie vorherrschenden *Reaktionsmechanismen* mit zunächst 5 Kapiteln über Substitutions-Reaktionen (Kap.10-14 mit insgesamt 486 Seiten). Es folgen umfangreiche Kapitel über Additions-Reaktionen an Verbindungen mit Mehrfachbindungen zwischen C-Atomen sowie zwischen C-Atomen und Hetero-Atomen (15 und 16). Den Abschluss bilden die Kapitel über Eliminierungs- und Umlagerungs-Reaktionen (17 und 18) sowie über Oxidationen und Reduktionen (19). Im Zuge der Darstellung mehrstufiger organischer Synthesen werden auch die einzelnen Reaktionsschritte eingehend erläutert.

Die vorgenommenen Aktualisierungen erfolgten vorrangig im Hinblick auf die in den Jahren 2005 bis 2010 publizierten Fortschritte im Bereich der Organischen Chemie, was zur Aufnahme von mehr als 5500 neuen Hinweisen auf Originalarbeiten und vielfältige Sekundärliteratur geführt hat. Insgesamt enthält das Werk mehr als 20 000 (!) Literaturangaben. Der Absicht des Autors entsprechend wurden Bereiche der Organischen Chemie wie Terpene, Kohlenhydrate, Proteine, zahlreiche organometallische Reagenzien, Kombinatorische Chemie, Polymerisationen und elektrochemische Reaktionen sowie Steroide nur kurz behandelt oder überhaupt weggelassen.

Dieses Lehrbuch ist nicht nur für die in erster Linie angesprochene Zielgruppe, Studierende der Chemie in höheren Semestern, empfehlenswert sondern wird als Nachschlagewerk zweifellos auch für auf dem Gebiet der organischen Synthese tätige Chemiker von großem Wert. *Dr. Dieter Holzner*

