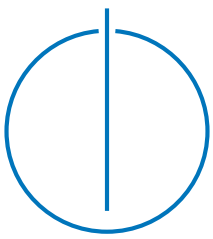
SCHOOL OF COMPUTATION,
INFORMATION AND TECHNOLOGY –
INFORMATICS

TECHNISCHE UNVERSITÄT MÜNCHEN

Bachelor's Thesis in Informatics

Quantum Algorithms for Bioinformatic Applications: Navigating the DNA Multiverse

Rosalind Franklinson



SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY – INFORMATICS

TECHNISCHE UNVERSITÄT MÜNCHEN

Bachelor's Thesis in Informatics

Quantum Algorithms for Bioinformatic Applications: Navigating the DNA Multiverse

Quantenalgorithmik für bioinformatische Anwendungen: Unterwegs im DNA-Multiversum

Author: Rosalind Franklinson

Supervisor: Prof. Dr. Ada Codeheimer, Chair of Quantum Computing and Computational Biology

Advisors: Dr. Alan Bitwise
Zoe Debugsmith, M.Sc.

Submitted: 15 Jan 2024

Declaration of Authorship

I confirm that this bachelor's thesis is my own work and I have documented all sources and material used.

Date Rosalind Franklinson

Zusammenfassung

Das ist von ChatGPT generierter Content, und daher wirklich nicht lustig – sry.

In einem Universum, das nicht allzu weit von unserem entfernt ist, wo Qubits frei umherstreifen und Algorithmen zum Rhythmus der Verschränkung tanzen, begibt sich Sheldon Bytefeld auf eine Quest, um die unerforschten Gebiete der Bioinformatik zu erkunden. Diese Arbeit taucht ein in das faszinierende Reich, in dem Quantencomputing auf die Feinheiten genomischer Daten trifft und darauf abzielt, die Geheimnisse im DNA-Multiversum zu enthüllen.

Abstract

In a universe not too far from ours, where qubits roam free and algorithms dance to the rhythm of entanglement, Sheldon Bytefeld embarks on a quest to explore the uncharted territories of bioinformatics. This thesis delves into the intriguing realm where quantum computing meets the intricacies of genomic data, aiming to unveil the secrets encoded within the DNA multiverse.

Contents

1 intro

Motivation
As xkcd wisely reminds us, “There’s nothing quite so beautiful as watching someone code in a language they’re fluent in.” In this linguistic dance between quantum bits and biological sequences, we strive to achieve fluency in the unique syntax that intertwines the principles of quantum computation with the complex language of life.

Background and related work
Drawing inspiration from xkcd #378, where “Python” is the programming language of choice for researchers plotting world domination, our research aligns itself with the power and elegance of Pythonic quantum algorithms. The literature review explores the quantum landscape, guiding us through the footsteps of pioneers like Schrödinger and Turing, who, if xkcd is to be believed (795), may have shared a cup of coffee in a parallel computational universe.

Outline
In the spirit of xkcd 327, we adopt a robust methodology akin to the “Exploits of a Mom,” ensuring the security and integrity of our quantum bioinformatics framework. Through the careful implementation of quantum gates and bioinformatic algorithms, we navigate the vast space of possibilities, akin to the legendary “Lorenz Attractor” (137).

2 materials & methods

some materials are important: food, WiFi, entertainment, …

2.1 differentiating material materials from immaterial methods

That's important, but I have no idea how to do it. Maybe Figure 1 can help us a tiny bit already?

Also: Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus

Figure 1: Example full-width image, caption above. Most people do not have computer science degrees, and they would probably count as more *method-y* than *material-y*?

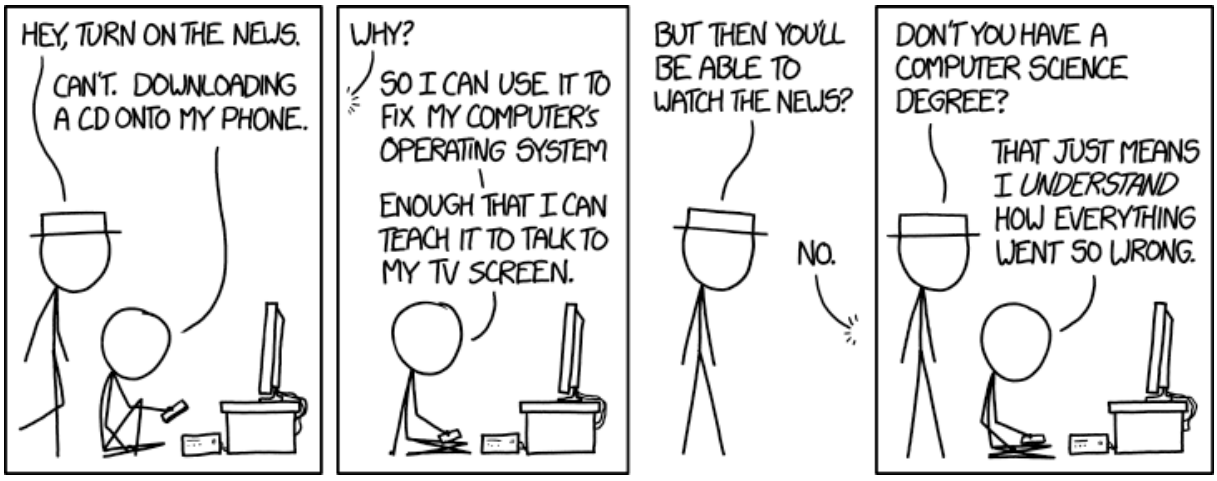


Figure 2: Another example full-width image. Consumers are generally unaware that toddlers fall right into the “tasty food” category for dingos; with anecdotal evidence suggesting that only individuals suffering from excessive slobbering ever successfully repelled the fanged predators. As tasty food, toddlers are clearly a **material**, *not* a method.

saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis, saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus existimo, neque eum Torquatum, qui hoc primus cognomen invenerit, aut torquem illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.

method A

This is a little bit of math and a link to chapter 3, which you can find on page 14:

$$\mathrm{HVAL} = \mathrm{PIDE} - \begin{cases} 100 & \forall L \leq 11 \\ 480 \cdot L^{-0.32[1+\exp(-L/1000)]} & \forall L \leq 450 \\ 19.5 & \forall L > 450 \end{cases}$$

Figure 3:

Example wrapped figure. In fact this is a side-by-side grid of text and this image. Actually a lot easier than in LaTeX! Also flexing the **Figure 3** label is in a different position now.

method B

This world is super chaotic! And the quantum realm isn't any better: May (1976) found this already (and we're quoting them a bit like `citetext` here). But by now we are lucky to have `ColabFold` (Mirdita et al., 2022) and in combination with some more power (see Figure 3), we can go far – for an example look no further than at Bryant et al. (2022). Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et.

2.2 materials

- Food
- WiFI
- coffee

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum

defuturum, quas natura non depravata desiderat. Et quem ad me accedis, saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus existimo, neque eum Torquatum, qui hoc primus cognomen invenerit, aut torquem illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.

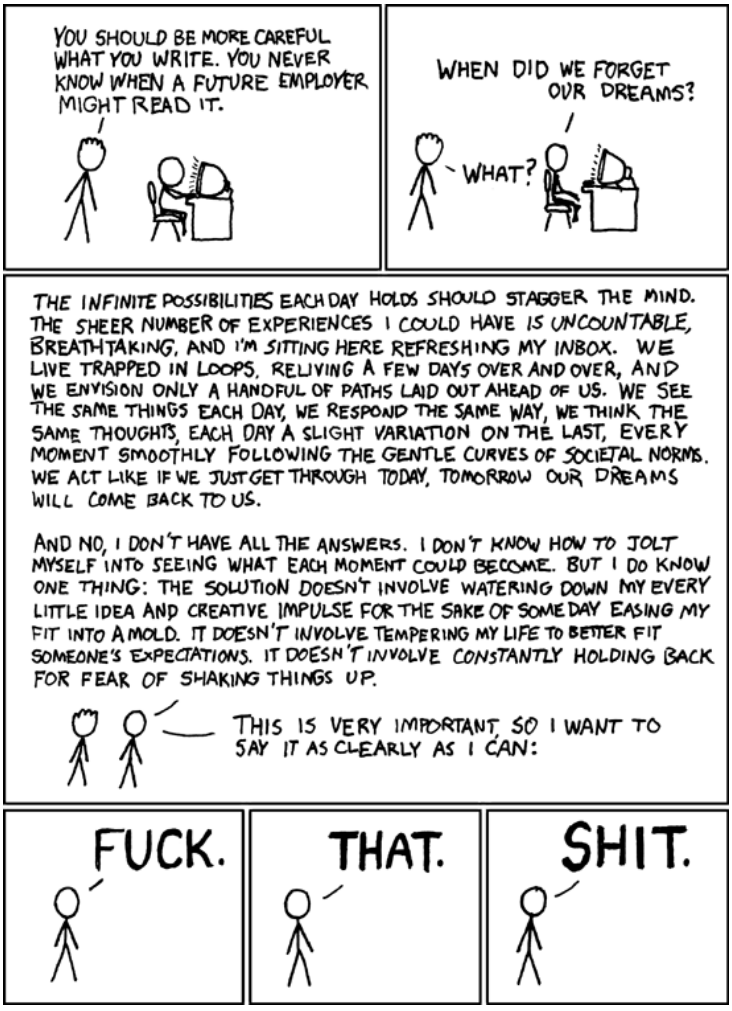


Figure 4: Side caption example. Look it's floating! In addition to `caption-pos` and two `-widths`, you can pass `placement` to `side-cap-fig`; also to `wrap-fig`. But not all combinations work – have a look at the docs if the figure crashes into text or the compiler does. The underlining show rule is scoped, see? Also underlined text; un-underlined link.

3 results

Quantum entanglement meets genomic entanglement as our algorithms traverse the DNA multiverse. Much like the unpredictable nature of xkcd's "Random Number" (221), the results paint a colorful tapestry of possibilities, offering glimpses into the potential of quantum-enhanced bioinformatics applications.

size	source	split	spp.	proteins	PPIs
full	APID	train	18	39393	143171
		validate	10	1750	2279
	HuRI	test	1	4458	19106
small	APID	train	8	8565	15586
		validate	7	818	1184
	HuRI	test	1	430	666

Table 1: PPI data in the sets; disregarding negatives and extra compensation proteins. Finished sets are about twice or 11 × the size.

References

Bryant, P., Pozzati, G., & Elofsson, A. (2022). Improved prediction of protein-protein interactions using AlphaFold2. *Nature Communications*, *13*(1), 1265. https://doi.org/10.1038/s41467-022-28865-w

May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, *261*(5560), 459–467. https://doi.org/10.1038/261459a0

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods*, *19*(6), 679–682. https://doi.org/10.1038/s41592-022-01488-1

List of Figures

List of Tables