Source tracker

Izaak van Dongen

June 4, 2018

1   Some general notes on my sources

Not all of my sources have been cross-referenced, although where possibly I've tried to demonstrate links between papers and their common subject matter. This doesn't really cause any problems though, as most of my sources are mathematical in nature. This means that really any source is either "wrong", or "right" about a theorem proposed by that source. This is pretty different from academic literature in more empirical fields, where cross-referencing is crucial.

Overwhelmingly, peer-reviewed maths papers older than a couple of years will be correct, and my most important sources are all decades old. This binary nature of mathematics also means that most theorems won't really be produced in more than one paper, except in some cases where an alternative proof is offered.

I would also argue that I have intrinsically been doing lots of research while writing my code. I have had to research what does and doesn't work, and write test plans to verify correct behaviour. I have also had to adapt what I've found in papers or online to work within the constraints of the programming language or my project. I have even written programs to simulate biological processes to verify the correct behaviour of my core programs.

2   Evaluation

Source	Source content/usage	Source evaluation
Error detecting and error correcting codes [Hamming, 1950]	Gives a construction of Hamming codes and lays the foundation for many of the concepts around channel-coding that I use in my dissertation.	This, is a paper by R. W. Hamming, who contributed much to modern error-correcting codes. It was published over 50 years ago in a respected paper by a respected author and has not only stood the test of time but also formed much of the basis of modern communication theory and can certainly be trusted.
A mathematical theory of communication [Shannon, 1948]	This is also a useful paper for establishing the theory of communicative coding. It is useful for several definitions and general limits.	This paper, together with [Hamming, 1950] are generally considered to be the seminal works on coding theory. This lays much of the groundwork for communication theory and gives a more general definition of the Hamming Code. Similarly to [Hamming, 1950], this is a paper that can be trusted

Generalized dna bar-code design based on hamming codes [Bystrykh, 2012]	This article is very much relevant to what my project is about. It doesn't seem to be very clear though, and it uses a seemingly non-optimal form of parity. However it provides helpful insight into what actual researchers in the field are doing and have done with these ideas.	It is an article in a reputable, peer reviewed journal by researchers in the field so it can probably be trusted. The journal is slightly less well-known than some others I've made use of.
Introduction to coding theory [Guruswami, 2010]	This is more of a crash-course in coding theory, which was useful in developing a basic understanding but can be built on.	It's not an academic article so it hasn't been peer reviewed and wasn't written with the aim of academic rigour. However it was produced by a mathematician at a prestigious university so is likely to have some merit. It's also specifically aimed at the teaching of the subject rather than academic description which may be useful for my purposes.
Polynomial codes: an optimal design for high-dimensional coded matrix multiplication [Yu, Maddah-Ali and Avestimehr, 2017]	This source provides an explanation of a facet of the polynomial code. This code type didn't end up being used for my project but the source was useful in explaining why not.	This source is very technically detailed, which isn't necessarily bad, but this source feels very dense and can very well be difficult to understand for anyone not familiar with this specific subfield of coding theory. Of course it does have academic merit as it's a published, peer reviewed article.
Families of Hadamard $z_2 z_4 q_8$-codes [del Río and Rifà, 2012]	This is an article relating to a specific subset of Hadamard codes. I didn't end up using these codes but this source provided some broader context for what Hadamard codes are/can be used for.	This is again a highly technical source but is generally reputable aside from that.
Hadamard matrices and their applications [Hedayat and Wallis, 1978]	This is a reasonably low-level paper about Hadamard codes. This was useful for my dissertation as I didn't use any overly complicated Hadamard constructions, and this was a useful source to cite, while also being trustworthy…	It's an academic paper from a trustworthy journal that has stood the test of time, and aside from that doesn't really make any outrageous claims, forming more of a useful and approachable summary of the subject matter.
The search for Hadamard matrices [Golomb and Baumert, 1963]	This paper gives a useful initial summary of the 'basic' construction of Hadamard matrices (of size $2^n \times 2^n$) which I use, corroborating [Hedayat and Wallis, 1978]. It also provides a nice summary of other constructions, which I used to strengthen my case for not using them as they seemed unnecessary.	It's a trustworthy article, and again provides more of a summary. It's written in a relatively friendly manner.

Hadamard matrices and their designs: A coding-theoretic approach [Assmus and Key, 1992]	This paper gives a number of very mathematically involved constructions of Hadamard matrices. I didn't end up using any of these but it provided a useful further background around Hadamard constructions, and could be compared/contrasted with [Golomb and Baumert, 1963]	It's a trustworthy paper from a trustworthy journal and trustworthy authors.
Hadamard designs [Spence, 1972]	This paper is about Hadamard designs for alphabets of size n where $n \neq 2$. This is potentially very interesting with respect to my project as the DNA alphabet size n is 4. This source was used to provide citations in my general discussion about conversion between different alphabets.	It's a trustworthy paper although it is on an apparently quite obscure subject, making it quite a rare occurrence. This doesn't really reflect on the trustworthiness of the source due to the reasons described in section 1.
Lifted polynomials over F_{16} and their applications to dna codes [Oztas and Siap, 2013]	This is a paper about a highly specific and even more advanced class of codes. It involved a lot of field theory, so for someone with an A-level knowledge of maths, will require a lot of further reading. However it does directly talk about applications of codes to DNA, which is useful to validate that this is a field with some merit.	It's a very reputable and technically accurate paper, but again is incredibly dense and required much more knowledge than I have, so I've not been able to effectively leverage it as a citation.
Codes, not ciphers [Baylis, 2010]	This is an article providing description and intruction of coding theory, and contrasting that with cryptography. It is aimed at a 'school …level'. This means that it was useful to gain a quick overview of what coding theory is, but didn't really help with any major lifting.	It's from a very reputable paper, and is also easily accessible. This makes it a good, reliable resource for introduction to the topic.
Error correcting codes: Practical origins and mathematical implications [Pless, 1978]	This is a very short little paper, which provides a short history of the concept of codes. This makes it quite a special little paper, and quite nice to have read to gain a better contextual understanding of the field. However, again, there's not much to cite.	Again, from a highly trustworthy paper and an established mathematician. Occasionally it uses some pretty advanced mathematical analogies, but is certainly a very solid source.
Boole and the algebra of logic [Kneale, 1956]	This is used as a small citation on Boolean algebra, which I needed when I was writing about my decision to use boolean inversion rather than additive inversion for my Hadamard construction. This isn't a very mathematically advanced source, but this is in line with my project as Boolean algebra was only a small part of a subset of my project, and the main focus was not on Boolean algebra.	This isn't really an article but a set of notes as it was published in 'Notes and Records of the Royal Society of London'. While this means it's not the epitome of academic rigour, it's still a useful and factual source on Boolean algebra from the Royal Society.

The degeneracy of the genetic code and Hadamard matrices. [Petoukhov, 2008]	This is a great paper talking about how Hadamard matrices can be applied to and modelled against genetic code. The really cool thing is that this is a paper from a biomechanical perspective, rather than a mathematician's. This highlights the truly awesome thing about working at an intersection of fields. Aside from that, used for some backing up of the application of Hadamard matrices.	This paper is a little wonky in terms of academic merit - it's not published in any well known journal, which might make it seem somewhat questionable. However, it is mathematical in nature, and is mathematically sound. This paper would have nothing to gain from misleading anyone anywa.
Construction of multilevel Hadamard matrices with small alphabet [Trinh and Fan, 2008]	This is a very interesting topic, as it potentially deals with the major limitation of Hadamard matrices - them having an alphabet of 2. Being able to construct a Hadamard matrix with different alphabet (ideally 4) would be incredibly useful. However, this paper doesn't quite go so far as to explicitly guide how that could be possible, so it's just another citation for theoretical extensions to this project.	It is a perfectly sound paper. Only quibble is that it's a little hard to gain access to.
Decoding the hamming code [Ehrenborg, 2006]	This is a tiny article offering a quick introduction to coding theory and Hamming codes. It contextualises the Hamming code in an interesting new way that I quite like and helped me really click with what they could be used for and what they were, so I used this source as an excellent other take on Hamming codes.	Again, despite not being very academic in and of itself it is from a very prestigious university and organisation, and can't really be faulted in any way.
PEP 257: Docstring conventions [Goodger and van Rossum, 2014]	This is part of the Python style guide, used to back up some of my assertions about coding practices in my Definitions section.	Of course, it isn't remotely academic and verges on being an appeal to authority, but it is a well respected style guide with some explicit reasoning behing the decisions being laid out.
WHAT IS CRISPR-CAS9? [Daily Mail, 2015]	This is a short "article" offering a brief introduction of the principles of CRISPR-CAS9. The best part is probably a diagram produced by MIT. It's not particularly informative aside from the most general, basic sense.	It's literally not from a scientific publication and has no academic merit. This is a bad source and does not prove or reinforce anything.

References

Assmus, E. F. and Key, J. D. [1992], 'Hadamard matrices and their designs: A coding-theoretic approach', *Transactions of the American Mathematical Society* **330**(1), 269–293.

URL: *http://www.jstor.org/stable/2154164*

Baylis, J. [2010], 'Codes, not ciphers', *The Mathematical Gazette* **94**(531), 412–425.
URL: *http://www.jstor.org/stable/25759725*

Bystrykh, L. V. [2012], 'Generalized dna barcode design based on hamming codes', *PLOS ONE* .
URL: *http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0036852*

Daily Mail [2015], 'WHAT IS CRISPR-CAS9?'.
URL: *http://www.dailymail.co.uk/sciencetech/fb-5151843/WHAT-CRISPR-CAS9.html*

del Río, Á. and Rifà, J. [2012], 'Families of hadamard z2z4q8-codes', *CoRR* **abs/1211.5251**.
URL: *http://arxiv.org/abs/1211.5251*

Ehrenborg, R. [2006], 'Decoding the hamming code', *Math Horizons* **13**(4), 16–17.
URL: *http://www.jstor.org/stable/25678619*

Golomb, S. W. and Baumert, L. D. [1963], 'The search for hadamard matrices', *The American Mathematical Monthly* **70**(1), 12–17.
URL: *http://www.jstor.org/stable/2312777*

Goodger, D. and van Rossum, G. [2014], 'Pep 257: Docstring conventions', PEP at *https://legacy.python.org/dev/peps/pep-0257*.

Guruswami, V. [2010], 'Introduction to coding theory'.
URL: *http://www.cs.cmu.edu/ venkatg/teaching/codingtheory/notes/notes1.pdf*

Hamming, R. W. [1950], 'Error detecting and error correcting codes', *The Bell System Technical Journal* **26**(2), 147–160.
URL: *http://sb.fluomedia.org/hamming/*

Hedayat, A. and Wallis, W. D. [1978], 'Hadamard matrices and their applications', *The Annals of Statistics* **6**(6), 1184–1238.
URL: *http://www.jstor.org/stable/2958712*

Kneale, W. [1956], 'Boole and the algebra of logic', *Notes and Records of the Royal Society of London* **12**(1), 53–63.
URL: *http://www.jstor.org/stable/530792*

Oztas, E. S. and Siap, I. [2013], 'Lifted polynomials over F_{16} and their applications to dna codes', *Filomat* **27**(3), 459–466.
URL: *http://www.jstor.org/stable/24896375*

Petoukhov, S. V. [2008], The degeneracy of the genetic code and hadamard matrices.
URL: *https://arxiv.org/pdf/0802.3366.pdf*

Pless, V. [1978], 'Error correcting codes: Practical origins and mathematical implications', *The American Mathematical Monthly* **85**(2), 90–94.
URL: *http://www.jstor.org/stable/2321784*

Shannon, C. E. [1948], 'A mathematical theory of communication', *The Bell System Technical Journal* **27**, 379–423, 623–656.
URL: *http://affect-reason-utility.com/1301/4/shannon1948.pdf*

Spence, E. [1972], 'Hadamard designs', *Proceedings of the American Mathematical Society* **32**(1), 29–31.
URL: *http://www.jstor.org/stable/2038298*

Trinh, Q. and Fan, P. [2008], 'Construction of multilevel hadamard matrices with small alphabet', **44**, 1250 – 1252.

Yu, Q., Maddah-Ali, M. A. and Avestimehr, A. S. [2017], 'Polynomial codes: an optimal design for high-dimensional coded matrix multiplication', *CoRR* **abs/1705.10464**.
URL: *http://arxiv.org/abs/1705.10464*